

The Evolution of Pointwise Statistics in Hyperbolic Equations with Random Data*

Alina Chertock[†], Pierre Degond[‡], Amir Sagiv[§], and Li Wang[¶]

Abstract. We consider one-dimensional hyperbolic PDEs, linear and nonlinear, with random initial data. Our focus is the *pointwise statistics*, i.e., the probability measure of the solution at any fixed point in space and time. For linear hyperbolic equations, the probability density function (PDF) of these statistics satisfies the same linear PDE. For nonlinear hyperbolic PDEs, we derive a linear transport equation for the cumulative distribution function (CDF) and a nonlocal linear PDE for the PDF. Both results are valid only as long as no shocks have formed, a limitation which is inherent to the problem, as demonstrated by a counterexample. For systems of linear hyperbolic equations, we introduce the multi-point statistics and derive their evolution equations. In all of the settings we consider, the resulting PDEs for the statistics are of practical significance: they enable efficient evaluation of the random dynamics, without requiring an ensemble of solutions of the underlying PDE, and their cost is not affected by the dimension of the random parameter space. Additionally, the evolution equations for the statistics lead to a priori statistical error bounds for Monte Carlo methods (in particular, kernel density estimators) when applied to hyperbolic PDEs with random data.

Key words. uncertainty propagation, characteristics, hyperbolic, Young measure

MSC codes. 35L65, 35R60, 60H30, 65C20, 62G07

DOI. 10.1137/25M1777955

1. Introduction. In mathematical modeling, the inputs of an otherwise deterministic model are often random or uncertain. It is then natural to study the *quantity of interest* (model output) as a random variable. This task, known as uncertainty propagation (or a forward problem in uncertainty quantification), is prevalent in many scientific and engineering fields. In particular, such approaches have been successfully applied to hyperbolic partial differential equations (PDEs) and conservation laws; see, e.g., [AM17, JP17, MS12, MSŠ13, MJM13, Peh20, PIN09, PDL09, TLMNE10]. Such hyperbolic equations, in turn, are fundamental

*Received by the editors July 15, 2025; accepted for publication (in revised form) February 19, 2026; published electronically July 22, 2026.

<https://doi.org/10.1137/25M1777955>

Funding: The work of the first author was supported in part by NSF grant DMS-2208438. The work of the third author was supported in part by NSF grant DMS-2508811 and Binational Science Foundation Research grant 2022254. The work of the fourth author was partially supported by NSF grants DMS-1846854 and DMS-2513336 and the Simons Fellowship.

[†]Department of Mathematics, North Carolina State University, Raleigh, NC 27695 USA (chertock@math.ncsu.edu).

[‡]Institut de Mathématiques de Toulouse, CNRS & Université de Toulouse, 31062 Toulouse Cedex 9, France (pierre.degond@gmail.com).

[§]Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ 07102 USA (amir.sagiv@njit.edu).

[¶]School of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA (liwang@umn.edu).

models in such diverse fields as gas dynamics, shallow water equations, magneto-hydrodynamics in plasma physics, car and pedestrian traffic, and nonlinear elasticity.

Common computational approaches to uncertainty propagation utilize an ensemble of solutions to the underlying model, thereby approximating the statistical properties of the quantity of interest (see section 1.2). Moreover, even if the quantity of interest is a property of the solution of a PDE at some point in time, such methods will usually disregard the *dynamics* of that quantity in time.

In this work, we study the dynamics of measures associated with the solutions of (deterministic) hyperbolic PDEs with random initial data. Concretely, consider a family of one-dimensional hyperbolic equations with random initial conditions,

$$(1.1) \quad \begin{cases} \partial_t u(t, x; \omega) = c(x) \partial_x [a(u)] , \\ u(0, x, \omega) = u_{(0,x)}(\omega) . \end{cases}$$

Here $u : [0, T] \times \mathbb{R}_x \times \Omega \rightarrow \mathbb{R}^n$ for some $T > 0$, the parameter space $\Omega \subseteq \mathbb{R}^d$ is equipped with a Borel probability measure μ , for every $x \in \mathbb{R}$ the initial data $u_{(0,x)} : \Omega \rightarrow \mathbb{R}^n$ is a measurable function, $c : \mathbb{R}_x \rightarrow \mathbb{R}^n$, and $a(u)$ is either the identity matrix or a strictly convex function (see the respective sections for details). We study the *pointwise statistics* $\phi_{t,x}$: a measured-valued function (a Young measure [B⁺95, Ped97]), which for every fixed $x \in \mathbb{R}$ and $t \geq 0$ is defined as

$$(1.2a) \quad \phi_{t,x} \equiv u(t, x; \cdot)_{\#} \mu ,$$

where we recall that the *pushforward probability measure* of μ by the solution operator $u(t, x; \cdot)$ satisfies

$$(1.2b) \quad \phi_{t,x}(A) = \mu(\{\omega \in \Omega \mid u(t, x; \omega) \in A\})$$

for all Borel measurable sets $A \subseteq \mathbb{R}^n$.¹ We ask the following:

QUESTION 1.1. *Can one determine $(t, x) \mapsto \phi_{t,x}$, the pointwise statistics at time $t \geq 0$, solely based on their initial value $x \mapsto \phi_{0,x}$?*

Fundamentally, it is worth noting that for general initial value problems with random data, the answer to Question 1.1 is *negative*: the random field $u(t, x; \cdot)$ contains information regarding the correlations between different points in space and time, whereas the pointwise statistics $\phi_{t,x}$ does not. Put differently, even full knowledge of $x \mapsto \phi_{t,x}$ tells us nothing about the *joint* probability of $u(t, x; \omega)$ and $u(t, y; \omega)$ for any $x \neq y$. Since the mapping between $u(t, x; \omega)$ to $\phi_{t,x}$ is not injective, knowledge of $\phi_{0,x}$ should not be sufficient to predict $\phi_{t,x}$.

Notwithstanding, and using the fact that $u(t, x; \omega)$ is constrained by being the solution of a hyperbolic PDE such as (1.1), one can still answer Question 1.1 affirmatively. The importance of the hyperbolicity of the underlying ODE can be seen, for example, by our use of the method of characteristics (see, e.g., [Eva22, sect. 3.2]) in proving the main results. To date, we *do not know of any other category of PDEs* for which the answer to Question 1.1 is positive.

¹In the literature, $\phi_{t,x}$ is also known as the “law” of the solution. To avoid ambiguity, we refrain from this term.

1.1. Main results. In the *linear scalar* (possibly heterogeneous) transport equation

$$\partial_t u(t, x; \omega) = c(x) \partial_x u,$$

the probability density function (PDF) of $\phi_{t,x}$, denoted by $f(t, x, \cdot)$, satisfies *the same* PDE (Theorem 2.1):

$$\partial_t f(t, x, U) = c(x) \partial_x f.$$

This description provides a positive answer to Question 1.1: since f satisfies a linear transport equation, the initial data $f(0, \cdot, \cdot)$ completely determines $f(t, \cdot, \cdot)$ for all later times. We provide two new proofs for this result: one using a weak formulation of the transport equation (weak in the random variable ω , not in the space variable x) and the other using the method of characteristics.

In the *scalar nonlinear case*,

$$\partial_t u(t, x; \omega) = \partial_x (a(u)), \quad a'' > 0,$$

the story becomes more complicated: in the *absence of shocks*, the cumulative distribution function (CDF) $F(t, x, U) \equiv \phi_{t,x}(-\infty, U)$ satisfies a family of *linear* equations (Theorem 3.1)

$$\partial_t F + a'(U) \partial_x F = 0,$$

while the PDF satisfies a nonlocal PDE (Theorem 3.2). However, *when a shock forms, the answer to Question 1.1 is negative*. We show an explicit example where knowledge of $\phi_{t,x}$ before the shock formation is *not sufficient* to determine $\phi_{t,x}$ for later times (section 6.3).

For a broad class of *systems of linear hyperbolic equations*, the answer to Question 1.1 is negative (section 4): $\{\phi_{t,x}\}_x$ cannot be determined from $\{\phi_{0,x}\}_x$, due to the importance of pointwise correlations between the different components of the vector-valued random fields (Proposition 4.1). Instead, for N -dimensional systems we introduce the *N -points statistics* $\phi_{t,x_1,\dots,x_N}$ and show that it can be propagated in time and that its PDF satisfies an $(N+1)$ -dimensional linear transport equation (Proposition 4.5).

Finally, in section 5 we apply our results to the *statistical* study of *Monte Carlo* methods in uncertainty quantification (UQ): using the evolution equations for $\phi_{t,x}$ derived in the previous sections, we show that the PDF $f(t, x, U)$ maintains the regularity of the initial PDF $f(0, x, U)$ (with analogous results for the CDF). We use that regularity result to prove *a-priori statistical convergence rates* for a popular nonparametric estimator of $f(t, x, U)$, kernel density estimators (KDEs); see [Was06]. Hence, even without solving the evolution equations for the PDF and CDF, their mere existence provides insight into the accuracy of the more standard Monte Carlo methods.

1.2. Motivation. Computing $\phi_{t,x}$ is a quintessential task in the field of UQ, often known as *uncertainty propagation*: given a deterministic model with random parameters or inputs (usually a PDE), one is interested in the statistical characterization of the model outputs, e.g., of the solution $u(t, x; \omega)$. While traditionally the UQ literature focuses on moment estimation [GHO⁺17, SDK00, Xiu10], in many applications it is necessary to obtain the “full statistics,”

i.e., the PDF or CDF of the quantity of interest [CGH05, CNP⁺18, DFS20, LMMKH10, NDM⁺09, PGS⁺19, Sag20, Sag22, SDF17, GZ07]. Standard approaches to uncertainty propagation can be divided into two categories:

- *Monte Carlo methods.* Here i.i.d. samples of $u(t, x; \omega)$ are computed and used in a nonparametric density estimator [Tsy09, Was06]. Broadly speaking, these methods are robust and theoretically sound, but tend to converge slowly (in the number of samples, i.e., of PDE solves), and are therefore often computationally expensive.
- *Surrogate models.* Here the function $\omega \mapsto u(t, x; \omega)$ is approximated by a simpler function, e.g., a polynomial or a spline. While constructing the surrogate model may be computationally expensive, evaluating the surrogate model is relatively computationally inexpensive. Hence, the statistics associated with the surrogate model can be obtained to arbitrary precision. The accuracy of surrogate models relies on the smoothness of the map $\omega \mapsto u(t, x; \omega)$, which can only be a-priori guaranteed for certain equations (for such examples, see, e.g., [CDS10, JS22]). Another fundamental challenge of surrogate models is the curse of dimensionality: without special care (e.g., the design of sparse grids [BNTT11]), the computational cost of a surrogate model with fixed accuracy would increase exponentially in the dimension of Ω .

Uncertainty propagation in *hyperbolic equations* has garnered a lot of attention over the years, with hundreds of works using both Monte Carlo methods, e.g., [MS12, MSŠ13, MJM13], and surrogate models, e.g., [PIN09, PDL09, TLMNE10]. See [AM17, JP17] and the references therein for a broader and systematic review.

From a *practical point of view*, any description of the dynamics of $\phi_{t,x}$ (i.e., a positive answer to Question 1.1) *circumvents* the need for a large ensemble of solutions of the underlying PDE (1.1). From this perspective, our results show that (in some cases) the pointwise statistics $\phi_{t,x}$ can be studied directly, without resorting to the costly processes of sampling from or approximating the random field $u(t, x; \omega)$.

1.3. Existing literature. Let us first mention two other notable analogous settings in which Question 1.1 has been explored. First, consider an *ordinary* differential equation (ODE) of the form $\dot{\mathbf{y}}(t) = Q(\mathbf{y})$ and where the initial conditions $\mathbf{y}(0) = \mathbf{y}_0 \in \mathbb{R}^n$ are randomly distributed according to the probability measure ρ_0 . Define the solution operator (flow) for time $t \geq 0$ as $Y_t \mathbf{y}_0 \equiv \mathbf{y}(t)$. Then $\mathbf{y}(t)$ is distributed according to $\rho(t) \equiv (Y_t)_\# \rho_0$, which satisfies the following $(n+1)$ -dimensional continuity equation [San15, Chapter 4]:

$$(1.3) \quad \partial_t \rho(t, y) + \vec{\nabla}_y \cdot (\rho Q) = 0, \quad \rho(0, \cdot) = \rho_0.$$

For presentation purposes, it suffices to consider an absolutely continuous $\rho(t)$ and to identify it with its PDF, $\rho(t, \cdot) \in C([0, \infty) \rightarrow L^1(\mathbb{R}^n))$.² We now see that $\rho(t, y)$ is not the pointwise statistics $\phi_{t,x}$.³ Therefore, Question 1.1 and our main results are not an immediate generalization of (1.3).

²For a general measure, the derivatives need to be taken in an appropriate sense; see [San15, Chapter 4].

³The analogue of $\phi_{t,x}$ in the ODE settings would be the coordinatewise statistics $\phi(t, j)$ for $1 \leq j \leq n$, defined as the corresponding marginal, e.g., $\phi_{t,1}(A) \equiv \rho(t)(A \times \mathbb{R}^{n-1})$ for all measurable $A \subseteq \mathbb{R}$. The analogue of $\rho(t)$ in the PDE settings can be defined as follows: $U_t u_0(\cdot) \equiv u(t, \cdot)$ is the flow/solution operator of the hyperbolic initial value problem (1.1). Then ρ_0 would be the measure on the function space describing the distribution of $u_{0,\cdot}(\omega)$ is, and $\rho(t) \equiv (U_t)_\# \rho_0$. This is still a probability measure on function spaces, as opposed to $\phi_{t,x}$.

The second analogous settings are those of *stochastic* PDEs of the form

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t.$$

Here the randomness in the solution is due to the equation itself, rather than the initial data, and the pointwise statistics $\phi_{t,x}$ satisfy a deterministic PDE known as the Fokker–Planck equation (FPE). Perhaps the simplest and most well-known settings are those of driftless homogeneous white noise ($\mu \equiv 0$ and $\sigma(t, X_t) \equiv X_t$), in which the FPE reduces to the heat equation. For further details, see, e.g., [Pav14, Ris96].

Turning to the present settings of *hyperbolic PDEs*, evolution equations for the PDF and CDF, such as (2.2) and (3.2), were previously derived in an insightful line of works under the name “the method of distribution.” See, e.g., [BT16, BT20, TG15, YBTT20, YTT22] and the references therein. These works provide algorithms for various use cases, with a focus on practical applications.

This work makes several substantial contributions, with the long-term goal of developing a fundamental understanding of Question 1.1. First, our methods of proof are new. In particular, our use of the method of characteristics throughout this work helps identify what is unique about hyperbolic equations that yield evolution equations for $\phi_{t,x}$. Our measure-transport perspective also provides a clear method for deriving the dynamics of one-point statistics. Second, by providing definitive negative answers to Question 1.1 in the cases of shock formation and linear systems, we identify the limitations of the pointwise statistics: notably, for linear systems, we identify the N -points statistics $\phi_t(x_1, \dots, x_N)$ as the fundamental object that can be propagated in time. Finally, we establish a linkage between the evolution equations for the CDF and PDF, their regularity, and statistical error analysis of Monte Carlo methods, which we believe could be a helpful approach to studying Monte Carlo methods in UQ more broadly.

Related to this work is the concept of *statistical solutions* of hyperbolic conservation laws [FLM17, FLMW20]. The original motivation in [FKMT17] is the well-posedness of systems of N hyperbolic conservation laws $\partial_t u + \vec{\nabla}_x \cdot f(u) = 0$ in dimension $d > 1$. The idea was to generalize the PDE by embedding the initial data of the PDE in the space of Young measures and then seeking the corresponding measure-valued solution. In detail, the initial data $u_0 : \mathbb{R}^d \rightarrow \mathbb{R}^N$ is identified with the (atomic) Young measure $x \mapsto \delta_{u_0(x)}$, which is the initial data for a corresponding *measure-valued* Cauchy problem

$$(1.4) \quad \partial_t \langle \nu(t, x), \text{id} \rangle + \nabla_x \langle \nu(t, x), f \rangle = 0,$$

where a measure acts on a function by $\langle \nu(t, x), g \rangle = \int g(U) d\nu(t, x)(U)$. A main result of [FKMT17] is that (1.4) with “deterministic” (or atomic) initial condition admits an entropy-minimal solution (in an appropriate sense).

In the settings of this manuscript, (1.1), solutions to (1.4) with $\nu(0, x) = \phi_{0,x}$ are far from unique: for example, even for the simplest linear and scalar conservation law, $u_t + u_x = 0$, (1.4) reads as

$$\partial_t \mathbb{E}_\omega[u(t, x; \omega)] + \partial_x \mathbb{E}_\omega[u(t, x; \omega)] = 0;$$

i.e., it can only predict the mean, not the full measure $\phi_{t,x}$.

In [FLM17], this nonuniqueness issue is solved by embedding (1.4) into an *infinite hierarchy* of evolution equations not only for $\phi_{t,x}$ but also to the k -points statistics, e.g., for $k = 2$ the quantity $\phi_{t,x,y}$ defined in section 4. On the one hand, the authors of [FLM17] prove that this infinite hierarchy of pointwise correlations is guaranteed to have a solution in time. On the other hand, it is an infinite hierarchy (and so cannot be numerically computed), and it is indeed shown to be equivalent to the evolution of the entire random field $u(t, x; \omega)$. From both a practical and a fundamental perspective, it is interesting that this hierarchy can be truncated after a finite number of points: in this paper, $k = N$ for N linear hyperbolic equations, and $k = 1$ in nonlinear equations without shocks.

1.4. Discussion and consequences. The positive answers to Question 1.1 lead to several consequences which, to the best of our knowledge, were not previously noted.

Spatial correlations are not necessary. The pointwise statistics $\{\phi_{t,x}\}_{(t,x) \in \mathbb{R}_t \times \mathbb{R}_x}$ is a two-dimensional parametric family of probability measures on $\mathbb{R}$, *but not a random field*, i.e., not a probability measure of some function space from $\mathbb{R}_t \times \mathbb{R}_x$ to $\mathbb{R}_U$. In particular, $\{\phi_{t,x}\}_{(t,x)}$ does not contain information about spatial correlations between different points in space and time. For example, even full knowledge of $\{\phi_{t,x}\}_{(t,x) \in \mathbb{R}_t \times \mathbb{R}_x}$ cannot be used to compute $\mu\{\omega \mid u(t, x_1; \omega) > u(t, x_2; \omega)\}$ for $x_1 \neq x_2$ and some $t > 0$.

Hence, Theorem 2.1 shows that while an initial condition is expressed in terms of its dependence on ω , it is only the PDF $\phi_{0,x}$ that determines $\phi_{t,x}$ later on. Therefore, rather than tracking the (potentially infinite-dimensional) random field $u(t, x; \omega)$ and *derive* $\phi_{t,x}$, one can simply evolve in time the 2-dimensional scalar function $t \mapsto F(t, x; U)$.

Canonical choice of (Ω, ω) and dimensionality reduction. Suppose that, in addition to (1.1), we are given another set of initial data

$$\tilde{u}(0, x, \gamma) = \tilde{u}_{(0,x)}(\gamma),$$

where γ is in *another* probability space (Γ, μ_γ) . Our results then guarantee that if

$$\phi_{(0,x)} = u_{0,x}(\omega)_\# \mu = \tilde{u}_{0,x}(\gamma)_\# \mu_\gamma \quad \forall x \in \mathbb{R},$$

then the corresponding PDFs $\phi_{t,x}$ will remain equal for all times.

We can thus propose a *canonical* choice of ω and $\tilde{u}(0, x; \omega)$: *the inverse CDF*. Set $\Omega = [0, 1]$ with the uniform probability (Lebesgue) measure, and set

$$u_{0,x}(\omega) = F_{(0,x)}^{-1}(\omega) \equiv \inf\{U \mid F(U) \geq \omega\},$$

the left-continuous inverse of the CDF (or quantile function). It is standard that $(F_{(0,x)}^{-1})_\# \lambda = \phi_{(0,x)}$; see, e.g., [San15, Chapter 2].

This canonical choice of initial condition serves as a dimensionality reduction technique, which can be used as a preprocessing step for surrogate models: if all we are interested is in computing $\phi_{t,x}$ at later times, then it is sufficient to consider the one-dimensional construction above, thus allowing us to always construct surrogate models in one-dimensional settings.⁴

⁴We emphasize that this approach to dimension reduction will not provide information about correlations between different points in space, since the resulting random fields u and $\tilde{u}$ may very well be different.

Comparison with existing algorithms. For a given $t_*, U_* > 0$ and $x_* \in \mathbb{R}$, computing $f(t_*, x_*, U_*)$ can be done by (i) computing $f(0, x, U_*)$ directly for all $x \in \mathbb{R}$, and (ii) solving the PDE (2.2) once, up to time t_* . An analogous procedure can be proposed for computing the CDF $F(t_*, x_*, U_*)$. How does this approach compare with existing algorithms for finding $f(t, x, U)$?

The accuracy of the proposed algorithm is as good as the accuracy of the PDE solver. The computational downside is that for *every* choice of $U_* \in \mathbb{R}$, one needs to repeat the PDE solve anew.

In comparison, a *Monte Carlo method* would draw $N \geq 1$ i.i.d. samples $\omega_1, \dots, \omega_N \sim \mu$, each of which solves the original PDE (2.1) to obtain an approximation of $\{u(t, x; \omega_j)\}_{j=1}^N$. Here the complexity is N -times that of the PDE solver, but we immediately get an estimate for $f(t, x, U)$ for *all* (t, x, U) . The accuracy is dominated by two factors: the statistical error of the MC method, which *at best* scales at the parametric rate $\mathcal{O}(N^{-1/2})$; and the numerical error of the PDE scheme.

Similarly, *surrogate models* (e.g., collocation gPC), use N solutions of the original PDE (2.1) to obtain $\{u(t, x; \omega_j)\}_{j=1}^N$ and thereby obtain an approximation of $f(t, x, U)$ for all x and U . As noted above, the accuracy of such methods is preferable to that of Monte Carlo methods in moderate dimensions but deteriorates exponentially with the dimension of Ω without any special treatment [GHO⁺17, SDK00].

Regularity theory and statistical error estimates for Monte Carlo methods.

In those cases where the PDF $f(t, x, U)$ or CDF $F(t, x, U)$ is governed by a PDE, we can deduce from it their regularity for all time $t > 0$. Furthermore, regularity theory could yield an a-priori statistical error analysis for Monte Carlo methods when applied to these equations. See section 5 for details.

1.5. Open questions.

- (1) Theorem 3.1, which concerns the CDF associated with nonlinear scalar conservation laws, is only valid outside of the “shock envelope,” the regions in the (t, x) plane where the solution exists globally with probability 1. This limitation is demonstrated to be a fundamental one by a numerical example (section 6.3). One intriguing approach to overcoming this limitation is that of [BT20], where a PDE for the CDF is deduced even in the presence of shocks; however, it contains an unknown source term that the authors learn from data. In [FLM17], this issue is bypassed rigorously by considering the *infinite* hierarchy of multi-point statistics $\{\phi_{t, x_1, \dots, x_k}\}_{k \geq 1}$. Since these infinite hierarchies cannot be computed in full, the question of their computational approximation is then studied in [FLMW20]. Whether or not some knowledge of the shocks can be incorporated into a *finite* statistical object whose evolution can be described in closed form remains an open question.
- (2) Theorem 4.2 hinges on Assumption 4.2: for linear systems of evolution equations, the speed of a propagation matrix, $D(x)$ in (4.1), can be diagonalized by an x -independent matrix. The source of the issue appears to be technical and is discussed in Remark 4.4. Whether an evolution equation for $\phi_{t, x, y}$ associated with the general system (4.1) *without limiting assumptions* remains open.
- (3) Scalar nonlinear equations $u_t + \partial_x a(u) = 0$ with $a'' > 1$ can be rewritten in relaxation form: the limit $\varepsilon \rightarrow 0$ of the system (after a change of coordinates) [JX95]

$$(1.5) \quad \begin{cases} p_t + p_x = -\frac{1}{\varepsilon} \left(\frac{p-q}{2} - a \left(\frac{p+q}{2} \right) \right) \equiv -b(p, q), \\ q_t - q_x = b(p, q), \end{cases}$$

where we solve for unknown $p(t, x)$ and $q(t, x)$. This seemingly alternative approach to propagating $\phi_{t,x}$ in the presence of shocks, however, proves to be challenging. As noted in the previous open problem, even for a fixed $\varepsilon > 0$, two-point statistics associated with (1.5), which we denote by $\phi_{t,x,y}^\varepsilon$, cannot be obtained from Theorem 4.2—due to the mixed sources, i.e., the p and q dependence of the right-hand side in (1.5). If one was able to propagate in time $\phi_{t,x,y}^\varepsilon$, then the challenge would be to understand the limit $\lim_{\varepsilon \rightarrow 0} \phi_{t,x,y}^\varepsilon$ and see whether it is related to the original nonlinear problem.

2. Scalar linear transport equation. We begin with linear, scalar, variable-coefficient (heterogeneous) hyperbolic equations. Though the results here could be derived as a special case of the nonlinear case (section 3), we consider these separately for presentation purposes: the proofs here are simpler, and there are almost no technical caveats. Concretely, consider $u : \mathbb{R}_t \times \mathbb{R}_x \times \Omega \rightarrow \mathbb{R}$ satisfying⁵

$$(2.1) \quad \begin{cases} \partial_t u(t, x; \omega) = c(x) \partial_x u, \\ u(0, x, \omega) = u_{(0,x)}(\omega) \end{cases}$$

for some $c \in C^1(\mathbb{R})$.

Theorem 2.1. Consider (2.1) with $c \in C^1(\mathbb{R})$. Assume that $f_0(x, U)$ is piecewise C^1 in x for every $U \in \mathbb{R}$. Then $f(t, x, U)$, the PDF of the pointwise statistics $\phi_{t,x}$, satisfies the following equation:

$$(2.2) \quad \begin{cases} \partial_t f(t, x; U) = c(x) \partial_x f(t, x; U), \\ f(0, x, U) = f_0(x, U), \end{cases}$$

where $f_0(x, U)$ is the PDF of $\phi_{0,x}$.

While a version of this result appeared in [TG15], we provide two new proofs in section 2.1.

Remark (explicit example). Consider $\Omega = [0, 1]$ with the uniform Lebesgue measure, and consider $u_t = u_x$ with initial data $u_{(0,x)}(\omega) = (1 + \omega)e^{-x^2}$. At $t = 0$, the measure $\phi(0, x)$ is the uniform probability measure on $I_0(x) \equiv [e^{-x^2}, 2e^{-x^2}]$ for every $x \in \mathbb{R}$. Therefore, $f(0, x; U) = |I_0(x)|^{-1} \chi_{I_0(x)}(U)$, where we denote by $\chi_{[a,b]}$ the identifier function of an interval $[a, b]$. Since the solutions are traveling wave $u(t, x; \omega) = u(0, x+t; \omega)$, we have that $f(t, x, U) = |I_0(x+t)|^{-1} \chi_{I_0(x)}(U)$. Indeed, $f(t, x, U)$ satisfies (2.2), as expected by Theorem 2.1.

2.1. Two proofs of Theorem 2.1.

Proof of Theorem 2.1. Fix $t, x \in \mathbb{R}$. Since $\phi_{t,x}$ is the pushforward measure of μ by the evolution operator $u(t, x; \cdot)$, then for every $h \in C_c^\infty(\mathbb{R})$ we have that

$$(2.3) \quad \int_{\mathbb{R}} f(t, x, U) h(U) dU = \int_{\Omega} h(u(t, x; \omega)) d\mu(\omega).$$

⁵A sufficient regularity condition for existence and uniqueness is that $u_{0,x}(\omega)$ is piecewise C^1 in x and $c \in C^1(\mathbb{R})$. We do not pursue here the question of minimal regularity, such as in [Eva22, Chapter 7].

Differentiating both sides of (2.3), we get on the left-hand side

$$\partial_t \int_{\mathbb{R}} f(t, x, U) h(U) \, dU = \int_{\mathbb{R}} \partial_t f(t, x, U) h(U) \, dU$$

and on the right-hand side

$$\begin{aligned} \partial_t \int_{\Omega} h(u(t, x; \omega)) \, d\mu(\omega) &= \int_{\Omega} h'(u(t, x; \omega)) \partial_t u(t, x; \omega) \, d\mu(\omega) \\ \text{[using (2.1)]} \quad &= \int_{\Omega} h'(u(t, x; \omega)) c(x) \partial_x u(t, x; \omega) \, d\mu(\omega) \\ &= c(x) \int_{\Omega} h'(u(t, x; \omega)) \partial_x u(t, x; \omega) \, d\mu(\omega) \\ &= c(x) \partial_x \int_{\Omega} h(u(t, x; \omega)) \, d\mu(\omega) \\ \text{[using (2.3)]} \quad &= c(x) \partial_x \int_{\mathbb{R}} f(t, x, U) h(U) \, dU \\ &= \int_{\mathbb{R}} c(x) \partial_x f(t, x, U) h(U) \, dU. \end{aligned}$$

Hence,

$$(2.4) \quad \int_{\mathbb{R}} \partial_t f(t, x, U) h(U) \, dU = \int_{\mathbb{R}} c(x) \partial_x f(t, x, U) h(U) \, dU$$

for every $h \in C_c^\infty$ and therefore by density for every $h \in L^1$. Hence, the integrands are equal, and we have the PDE (2.2).

Note that, since the initial condition is piecewise C^1 in the x -variable, the PDE (2.2) itself guarantees that $f(t, \cdot; \cdot)$ remains in the same space for all $t \geq 0$, and hence the derivation is valid. ■

The alternative proof to Theorem 2.1 uses the method of characteristics.

An alternative proof for Theorem 2.1. Fix $\omega \in \Omega$ such that $x \mapsto u_{0,x}(\omega)$ is in C^1 . Then the solution to (2.1) can be written, by the method of characteristics, as

$$u(t, x; \omega) = u_0(\xi_x(t); \omega),$$

where $\xi_x(t)$ solves the ODE

$$(2.5) \quad \begin{cases} \dot{\xi}_x(t) = -c(\xi_x), \\ \xi_x(0) = x. \end{cases}$$

Fix $t \geq 0$, and define $q_t(x)$ to be the map $x \mapsto \xi_x(t)$. From ODE theory for existence and uniqueness, q_t is injective, and therefore $u(t, x; \omega) = u_0(q_t^{-1}(x); \omega)$. Hence, for any measurable $A \subseteq \mathbb{R}$

$$\begin{aligned} \phi_{t,x}(A) &= \mu(\{\omega \in \Omega \mid u(t, x; \omega) \in A\}) \\ &= \mu(\{\omega \in \Omega \mid u_0(q_t^{-1}(x); \omega) \in A\}) = \phi_{t,q_t^{-1}(x)}(A), \end{aligned}$$

and therefore the PDF satisfies $f(t, x, U) = f_0(q_t^{-1}(x), U)$ for all $t \in \mathbb{R}$. ■

2.2. Extension to randomness in coefficient. We now extend the initial randomness in (2.1) to include randomness in the coefficient as well. While similar results have been obtained in [TG15], in this brief section, we show how our methods provide an alternative derivation. We consider the following stochastic hyperbolic equation:

$$(2.6) \quad \begin{cases} \partial_t u(t, x; \omega) = c(\omega) \partial_x u, \\ u(0, x, \omega) = u_{(0,x)}(\omega). \end{cases}$$

Let

$$(2.7) \quad \phi_{t,x} := (u(t, x; \cdot), c(\cdot))_{\#} \mu,$$

and if $\phi_{t,x}$ is absolutely continuous, let $f(t, x, U, C)$ be its PDF. Then, for every $h \in C_c^\infty(\mathbb{R} \times \mathbb{R})$,

$$\int_{\mathbb{R} \times \mathbb{R}} f(t, x; U, C) h(U, C) dU dC = \int_{\Omega} h(u(t, x; \omega), c(\omega)) d\mu(\omega).$$

Relation (2.7) can be viewed as an extension of (1.2). Then the following can be proven in an analogous way to the proof of Theorem 2.1.

Corollary 2.2. *Consider (2.6), and assume that $x \mapsto f_0(x, U, C)$ is piecewise C^1 for all $U, C \in \mathbb{R}$. Then the PDF $f(t, x; U, C)$ satisfies the following linear transport equation:*

$$(2.8) \quad \begin{cases} \partial_t f(t, x; U, C) = C \partial_x f(t, x; U, C), \\ f(0, x, U, C) = f_0(x, U, C). \end{cases}$$

3. Nonlinear equations. Consider a nonlinear scalar conservation law with random initial conditions,

$$(3.1) \quad \begin{cases} u_t + a(u)_x = 0, \\ u(0, x, \omega) = u_{(0,x)}(\omega), \end{cases}$$

where $a : \mathbb{R} \rightarrow \mathbb{R}_+$ is strictly convex. We define $\phi_{t,x}$ as in (1.2) and as before define the corresponding CDF as $F(t, x, U) \equiv \phi_{t,x}((-\infty, U])$ and the corresponding PDF as $f(t, x, U) = \partial_U F(t, x, U)$. Throughout, we assume that up to time $T \geq 0$, no shocks have formed. This assumption allows for the methods of characteristics to be used to derive a linear hyperbolic evolution equation for the CDF (Theorem 3.1). While it is possible to differentiate the resulting CDF $F(t, x, U)$ to obtain $f(t, x, U)$, we also provide a direct derivation of the nonlocal linear equation that governs the propagation of the PDF (Theorem 3.2). Both methods of proof fall short of dealing with shock formation (Remark 3.3), and thus we complement our theoretical results by a numerical example which demonstrates why there *could be no propagator* for $\phi_{t,x}$ in the presence of shocks (section 6.3).

3.1. CDF evolution in the absence of shocks.

Theorem 3.1. *Consider (3.1), and assume that μ -almost everywhere (i) $u_{(0,x)}(\omega) > 0$, and (ii) there exists $T \leq \infty$ such that a unique strong solution of (3.1) exists for time $t \in [0, T]$. Then, for all such t ,*

$$(3.2) \quad \begin{cases} \partial_t F(t, x, U) + a'(U) \partial_x F = 0, \\ F(0, x, U) = F_{(0,x)}(U). \end{cases}$$

Equivalently, the evolution of the CDF obeys $F(t, x, U) = F(0, x - a'(U)t, U)$.

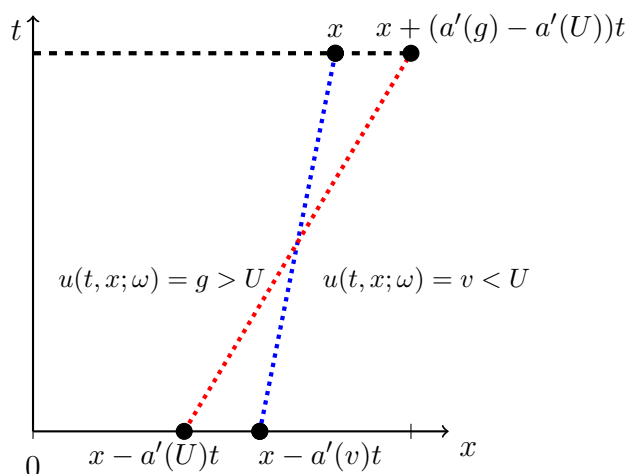


Figure 1. Illustration of the contradiction in the proof of Theorem 3.1. Tracing the line from (x, t) to $(x - a'(v)t, 0)$ (dotted blue), and using the fact that $a'(U) > a'(v)$ due to convexity, we see that if $u(0, x - a'(U)t) = g > U$, it would violate the principle that characteristic lines do not intersect (dotted red line).

Proof. Define the sublevel sets of $u(t, x; \omega)$ by

$$S(t, x, U) \equiv \{\omega \in \Omega \mid u(t, x; \omega) < U\}$$

for every $U \in \mathbb{R}$. We can then rewrite the CDF as

$$F(t, x, U) = \int_{S(t, x, U)} d\mu(\omega).$$

It is therefore sufficient to show that

$$(3.3) \quad S(t, x, U) = S(0, x - a'(U)t, U).$$

First, take any $\omega \in S(t, x, U)$. Then, by definition, there exists $v \in [0, U)$ such that $u(t, x; \omega) = v$. By the method of characteristics, $u(0, x - a'(v)t; \omega) = v$. Now, due to the strict convexity of a , then $a'(U) > a'(v)$, and so $x - a'(U)t < x - a'(v)t$. Suppose that $\omega \notin S(0, x - a'(U)t, U)$; this would mean that $u(0, x - a'(U)t; \omega) = g > U$. But then $u(t, x + (a'(g) - a'(U))t; \omega) = g$, but that would mean that characteristic lines have crossed *before* the time t , contrary to the assumption that no shocks have formed. See Figure 1. Therefore, $\omega \in S(0, x - a'(U)t, U)$, and we have proven that $S(t, x, U) \subseteq S(0, x - a'(U)t, U)$.

To prove inclusion in the opposite direction, fix any $\omega \in S(0, x - a'(U)t, U)$, i.e., $u(0, x - a'(U)t; \omega) = v < U$. Define $x_* \equiv x + (a'(v) - a'(U))t$. By the method of characteristics, $u(t, x_*; \omega) = v$. Since a' is strictly increasing, $x_* < x$. On the other hand, if $u(t, x; \omega) = V > U$, but then $x - a'(V)t < x - a'(U)t$, we would have that the characteristics cross, contradicting the assumption of no shocks being formed. Hence, we have $S(t, x, U) \supseteq S(0, x - a'(U)t, U)$, and (3.3) is proven. ■

3.2. PDF evolution using a nonlocal PDE. Here we will derive an equation for the PDF $f(t, x, U)$ directly from the conservation law (3.1).

Theorem 3.2. Consider the scalar conservation law with random initial condition, (3.1). Assume that the PDF $f(t, x, U)$ exists for all $t \geq 0$ and $x \in \mathbb{R}$ and that the probability of shocks forming on $[0, t]$ is zero. Then $f(t, x, U)$ satisfies

$$(3.4) \quad \partial_t f + \frac{\partial}{\partial U} \left(a'(U) \int_{-\infty}^U \partial_x f dv \right) = 0.$$

By integrating (3.4) with respect to U , one can readily obtain (3.2), the evolution equation for the CDF. On the one hand, it is a slightly weaker result, since it *assumes* the existence of the PDF, i.e., that $\phi_{(0,x)}$ is absolutely continuous for all $x \in \mathbb{R}$. However, we present it separately because the method of proof here is substantially different from that of Theorem 3.1.

Proof.

$$\begin{aligned} \partial_t \int_{-\infty}^{\infty} f(t, x, U) h(U) dU &= \partial_t \int_{\Omega} h(u_{(t,x)}(\omega)) d\mu(\omega) \\ &= - \int_{\Omega} h'(u_{(t,x)}) a'(u_{(t,x)}(\omega)) \partial_x u_{(t,x)}(\omega) d\mu(\omega). \end{aligned}$$

Now let

$$(3.5) \quad \psi'(u) = a'(u) h'(u);$$

then (3.5) becomes

$$\begin{aligned} \partial_t \int_{-\infty}^{\infty} f(t, x, U) h(U) dU &= - \int_{\Omega} \psi'(u_{(t,x)}(\omega)) \partial_x u_{(t,x)}(\omega) d\mu(\omega) \\ &= - \partial_x \int_{\Omega} \psi(u_{(t,x)}(\omega)) d\mu(\omega) \\ &= - \partial_x \int_{-\infty}^{\infty} f(t, x, U) \psi(U) dU \\ &= - \partial_x \int_{-\infty}^{\infty} f(t, x, U) \left(\int_{-\infty}^U a'(v) h'(v) dv \right) dU \\ &= - \partial_x \int_{-\infty}^{\infty} \psi'(v) \int_v^{\infty} f(t, x, U) dU dv \\ &= - \partial_x \int_{-\infty}^{\infty} \psi'(v) \left(1 - \int_{-\infty}^v f(t, x, U) dU \right) dv \\ &= \partial_x \int_{-\infty}^{\infty} h'(v) a'(v) \int_{-\infty}^v f(t, x, U) dU dv \\ &= - \partial_x \int_{-\infty}^{\infty} h(v) \frac{d}{dv} \left(a'(v) \int_{-\infty}^v f(t, x, U) dU \right) dv. \end{aligned}$$

Therefore, $f(t, x, U)$ solves (3.4). ■

Remark 3.3 (why do we assume that there are no shocks?). Our analysis of nonlinear hyperbolic equations with random data, (3.1), is valid only up to times where no shocks have

formed (with nonzero probability). We claim that this is an inherent limitation of the problem, rather than the result of a technical difficulty.

First, let us examine why shock is detrimental to our methods of proof. In Theorem 3.1, the method of characteristics is used to trace the evolution of the sublevel set $S(t, x, U)$ in time. When characteristic lines cross and a shock forms, such an argument can no longer hold (even if $u(0, x - a'(U)t; \omega) = U$, whether or not $u(t, x; \omega) = U$ depends on where, in (t, x) , the shock is), which in turn requires full knowledge of $x \mapsto u(0, x; \omega)$. But the latter information is not encoded in $\phi_{0,x}$. The importance of multi-point (in x) correlations is reminiscent of the work of [FLM17], where (for systems of nonlinear hyperbolic conservation laws) a hierarchy of k -points statistics for all $k \geq 1$ is needed to determine the statistical solution. In the proof of Theorem 3.2, already in (3.5) we require $\partial_x u(t, x; \omega)$ to exist in a strong sense, an assumption that fails in the presence of a shock.⁶

Beyond that, *the no-shock condition is fundamental to Question 1.1*. Our numerical results in section 6.3 demonstrate that two distinct initial random fields $u(0, x; \omega)$ and $v(0, x; \omega)$ may have the same initial pointwise statistics (i.e., $\phi_{0,x}^v = \phi_{0,x}^u$ for all $x \in \mathbb{R}$), but as shocks form at different points in (t, x, ω) space for u and v , the pointwise statistics differ as well. Therefore, in general there can be no operator or algorithm which takes the measure-valued function (or Young measure) $x \mapsto \phi_{0,x}$ and returns $x \mapsto \phi_{t,x}$ if shocks form with nonzero probability for some time $t \in [0, T)$.

4. Linear hyperbolic systems. Consider a linear 2×2 hyperbolic system in one space variable:

$$(4.1) \quad \begin{cases} \tilde{z}_t(t, x; \omega) + D(x)\tilde{z}_x = 0, \\ \tilde{z}(0, x, \omega) = \tilde{z}_{(0,x)}(\omega), \end{cases}$$

where $x \mapsto D(x)$ is a smooth map from $\mathbb{R}$ to the space of 2×2 matrices, and each $D(x)$ is real diagonalizable.

4.1. The failure of one-point statistics. In these settings, one can define the pointwise statistics $\phi_{t,x}$ in a way analogous to the scalar case

$$\phi_{t,x}(A) \equiv \mu(\{\omega \mid \tilde{z}(t, x; \omega) \in A\})$$

for every measurable $A \subseteq \mathbb{R}^2$. Unfortunately, $\phi_{t,x}$ is in general *not uniquely determined by the initial pointwise statistic* $x \mapsto \phi_{0,x}$. Hence, there can be no hope to find the time-propagator of $\phi_{t,x}$ (or its density and CDF). We show this using the following explicit counterexample.

Proposition 4.1. *In a vector-valued systems such as (4.1), the initial joint pointwise statistics $\{\phi_{0,x}\}_{x \in \mathbb{R}}$ are not sufficient to determine $\{\phi_{t,x}\}_{x \in \mathbb{R}}$ for $t > 0$.*

⁶A weak form of that derivation does not seem to advance us any further, unless the shock is known to form almost surely at a fixed location. We find these assumptions too stringent and therefore do not pursue this direction here.

Table 1

Two different sets of random initial conditions, $(u^{(1)}, v^{(1)})$ and $(u^{(2)}, v^{(2)})$, prescribed at $x = \pm 1$ for all possible values of ω .

ω	0	1
$u^{(1)}(0, 1; \omega)$	1	0
$u^{(1)}(0, -1; \omega)$	0	0
$v^{(1)}(0, 1; \omega)$	0	0
$v^{(1)}(0, -1; \omega)$	1	0
$u^{(2)}(0, 1; \omega)$	1	0
$u^{(2)}(0, -1; \omega)$	0	0
$v^{(2)}(0, 1; \omega)$	0	0
$v^{(2)}(0, -1; \omega)$	0	1

Proof. Consider the one-dimensional wave equation in system form:

$$\begin{cases} u_t(t, x; \omega) + u_x = 0, \\ v_t(t, x; \omega) - v_x = 0, \\ u(0, x, \omega) = u_{(0, x)}(\omega), \\ v(0, x, \omega) = v_{(0, x)}(\omega). \end{cases}$$

To show that the initial joint statistics $\phi_{0, x}$ do not determine the $\phi_{t, x}$ at later times, it is enough to provide two distinct sets of random initial conditions for which the initial joint pointwise statistics are the same but the joint pointwise statistics at $t = 1$ are different.

Take $\Omega = \{0, 1\}$ with the uniform measure ($\mu(0) = \mu(1) = 1/2$). We define the initial conditions, $(u^{(1)}, v^{(1)})$ and $(u^{(2)}, v^{(2)})$ in Table 1. Note that we only specify the initial conditions at $x = \pm 1$. This is because these are the only points we need to arrive at a contradiction, and any value they assume at $x \neq \pm 1$ would not change the argument.

The corresponding pointwise measures can be directly computed to show that

$$\phi_{0, \pm 1}^{(1)} = \phi_{0, \pm 1}^{(2)},$$

as both measures assign a probability $1/2$ to each of the two points $(1, 0), (0, 0) \in \mathbb{R}^2$ at $x = 1$ and probability $1/2$ to $(0, 1), (0, 0) \in \mathbb{R}^2$ at $x = -1$. Hence, we established that the initial pointwise statistics of both sets of initial conditions are the same.

Now observe that the solutions to both equations are simple traveling waves in the opposite directions, i.e., $u(t, x; \omega) = u(0, x - t; \omega)$ and $v(t, x; \omega) = v(0, x + t; \omega)$ for all $t \geq 0$, $x \in \mathbb{R}$, and $\omega \in \Omega$. Hence,

$$u^{(1)}(1, 0; 0) = v^{(1)}(1, 0; 0) = 1, \quad u^{(1)}(1, 0; 1) = v^{(1)}(1, 0; 1) = 0,$$

and so $\phi_{1, 0}^{(1)}(1, 1) = \phi_{1, 0}^{(1)} = \frac{1}{2}$. At the same time, the analogous calculation shows that $\phi_{1, 0}^{(2)}(1, 1) = 0$, which shows that $\phi_{1, 0}^{(1)} \neq \phi_{1, 0}^{(2)}$, which concludes our proof. ■

4.2. Two-point statistics. The counterexample above demonstrates the issue with propagating $\phi_{t,x}$ but, at the same time, suggests the appropriate fix: in vector-valued systems, two-point correlations (in space) are significant. This motivates the following definition: denote $\vec{z} = (z_1, z_2)^\top$, and by a small abuse of notation, define the two-point statistics as

$$(4.2) \quad \phi_{t,x,y}(A) \equiv \mu \left(\left\{ \omega \in \Omega \mid \begin{pmatrix} z_1(t, x; \omega) \\ z_2(t, y; \omega) \end{pmatrix} \in A \right\} \right)$$

for all measurable $A \subseteq \mathbb{R}^2$. Alternatively, we can define the measure $\phi_{t,x,y}$ as a functional acting on test functions $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$:

$$\phi_{t,x,y}[\psi] = \int_{\Omega} \psi(z_1(t, x; \omega), z_2(t, y; \omega)) \, d\mu(\omega).$$

In what follows, we make the following assumption.

Assumption 4.2. The matrix $D(x)$ in (4.1) admits an x -independent diagonalization; i.e., there exist an invertible matrix P and two functions, $c_1(x, u)$ and $c_2(x, v)$, such that for all $x \in \mathbb{R}$

$$PD(x)P^{-1} = \begin{pmatrix} c_1(x) & 0 \\ 0 & c_2(x) \end{pmatrix}.$$

Remark 4.1. Consider the diagonalized system

$$(4.3) \quad \begin{cases} u_t + c_1(x)u_x = 0, \\ v_t + c_2(x)v_x = 0, \\ u(0, x, \omega) = u_{(0,x)}(\omega), \quad v(0, x, \omega) = v_{(0,x)}(\omega), \end{cases}$$

and suppose that we are given the two-point statistics *for the new coordinates u and v* , which for the sake of this remark we shall denote by $\tilde{\phi}_{t,x,y}$. The one-point statistics $\phi_{t,x}$ of $\vec{z}$ can be obtained, for any $A \subseteq \mathbb{R}^2$, using the change of variables

$$\begin{aligned} \phi_{t,x}(A) &\equiv \mu \{ \vec{z}(t, x; \omega) \in A \} \\ &= \mu \left\{ (u(t, x; \omega), v(t, x; \omega))^\top \in P^{-1}A \right\} \\ &= \tilde{\phi}_{t,x,x}(P^{-1}(A)), \end{aligned}$$

where $P^{-1}(A) = \{P^{-1}\vec{\zeta} \mid \vec{\zeta} \in A\}$.

Hence, under Assumption 4.2, we need only to study $\phi_{t,x,y}$ for the decoupled system (4.3).

Theorem 4.2. Consider (4.3) with Assumption 4.2. Let $\phi_{t,x,y}$ be the two-points statistics, defined for every $A \subseteq \mathbb{R}^2$,

$$\phi_{t,x,y}(A) \equiv \mu \left(\left\{ \omega \in \Omega \mid (u(t, x; \omega), v(t, y; \omega))^\top \in A \right\} \right),$$

and $f(t, x, y, U, V)$ its PDF. If $f(0, x, y, \cdot, \cdot) = f_0(x, y, \cdot, \cdot) \in L^1(\mathbb{R}^2)$, then

$$\partial_t f(t, x, y, U, V) + c_1(x)\partial_x f + c_2(y)\partial_y f = 0,$$

for all $t \geq 0$ and $(x, y) \in \mathbb{R}^2$, with the appropriate initial conditions.

Proof. Letting $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a test function, we write

$$\begin{aligned}
 \partial_t \iint_{\mathbb{R}^2} f(t, x, y, U, V) \psi(U, V) \, dU \, dV &= \partial_t \int_{\Omega} \psi(u(t, x; \omega), v(t, y; \omega)) \, d\mu(\omega) \\
 &= \int_{\Omega} \left[\partial_U \psi(u(t, x; \omega), v(t, y; \omega)) \partial_t u(t, x; \omega) + \partial_V \psi(u(t, x; \omega), v(t, y; \omega)) \partial_t v(t, y; \omega) \right] d\mu(\omega) \\
 &= \int_{\Omega} \left[\partial_U \psi(u(t, x; \omega), v(t, y; \omega)) (-c_1(x) \partial_x u(t, x; \omega)) \right. \\
 (4.4) \quad &\quad \left. + \partial_V \psi(u(t, x; \omega), v(t, y; \omega)) (-c_2(y) \partial_y v(t, y; \omega)) \right] d\mu(\omega) \\
 &= -c_1(x) \partial_x \int_{\Omega} \psi(u(t, x; \omega), v(t, y; \omega)) \, d\mu(\omega) - c_2(y) \partial_y \int_{\Omega} \psi(u(t, x; \omega), v(t, y; \omega)) \, d\mu(\omega) \\
 &= \iint_{\mathbb{R}^2} (-c_1(x) \partial_x - c_2(y) \partial_y) f(t, x, y, U, V) \psi(U, V) \, dU \, dV.
 \end{aligned}$$

Since this is true for all test functions, this completes the proof. ■

What can be said about the general case, i.e., if one omits Assumption 4.2? Since $D(x)$ is diagonalizable, there exists an x -dependent diagonalizing matrix $P(x)$ satisfying

$$P(x)D(x)P^{-1}(x) = \text{diag}(c_1(x), c_2(x)) \equiv \Lambda(x).$$

Defining $(u, v)^{\top} \equiv P(x)\tilde{z}$, then (4.1) can be rewritten as

$$(4.5a) \quad \begin{cases} u_t(t, x; \omega) + c_1(x)u_x = g_1(x, u, v), \\ v_t(t, x; \omega) + c_2(x)v_x = g_2(x, u, v), \\ u(0, x, \omega) = u_{(0,x)}(\omega), \\ v(0, x, \omega) = v_{(0,x)}(\omega), \end{cases}$$

where the linear source terms are

$$(4.5b) \quad \begin{pmatrix} g_1(x, u, v) \\ g_2(x, u, v) \end{pmatrix} \equiv \Lambda(x)P'(x)P^{-1}(x) \begin{pmatrix} u \\ v \end{pmatrix}.$$

While we do not know how to treat the general case, there is a relaxation of Assumption 4.2 we can consider.

Proposition 4.3. Consider (4.5), and assume that $\Lambda(x)P'(x)P^{-1}(x)$ is a diagonal matrix. Then the PDF of the two-points statistics (see (4.2)) satisfies the following PDE:

$$\partial_t f(t, x, y, U, V) + c_1(x) \partial_x f + c_2(y) \partial_y f - \partial_U (f \cdot g_1(x, U)) - \partial_V (f \cdot g_2(y, V)) = 0.$$

Proof. In expanding $\int_{\Omega} \partial_U \psi \partial_t u \, d\mu(\omega)$ and $\int_{\Omega} \partial_V \psi \partial_t v \, d\mu(\omega)$, two new terms are added to (4.4):

$$\begin{aligned}
 (4.6) \quad &\int_{\Omega} \partial_U \psi(u(t, x; \omega), v(t, y; \omega)) g_1(x, u(t, x; \omega), v(t, x; \omega)) \, d\mu(\omega) \\
 &+ \int_{\Omega} \partial_V \psi(u(t, x; \omega), v(t, y; \omega)) g_2(y, u(t, y; \omega), v(t, y; \omega)) \, d\mu(\omega).
 \end{aligned}$$

Using the assumption that $\Lambda(x)P'(x)P^{-1}(x)$ is diagonal, (4.6) simplifies to

$$\begin{aligned} & \int_{\Omega} \partial_U \psi(u(t, x; \omega), v(t, y; \omega)) g_1(x, u(t, x; \omega)) d\mu(\omega) \\ & + \int_{\Omega} \partial_V \psi(u(t, x; \omega), v(t, y; \omega)) g_2(y, v(t, y; \omega)) d\mu(\omega) \\ & = \iint_{\mathbb{R}^2} \partial_U \psi(U, V) g_1(x, U) f(t, x, y, U, V) dU dV \\ & + \iint_{\mathbb{R}^2} \partial_V \psi(U, V) g_2(y, V) f(t, x, y, U, V) dU dV \\ & = - \iint_{\mathbb{R}^2} \psi(U, V) \partial_U [g_1(x, U) f(t, x, y, U, V)] dU dV \\ & - \iint_{\mathbb{R}^2} \psi(U, V) \partial_V [g_2(y, V) f(t, x, y, U, V)] dU dV, \end{aligned}$$

where the last passage is due to integration by parts and the vanishing of the boundary terms. Finally, since these passages hold for all test functions ψ , we get the PDE above. ■

Remark 4.4. The proof of Proposition 4.3 also reveals the difficulty with (4.5) when $\Lambda(x)P'(x)P^{-1}(x)$ is not diagonal: since g_1 depends on V (and analogously g_2 depends on U), (4.6) depends on $v(t, x; \omega)$ in a nontrivial way. Given the definition of the two-points statistics $\phi_{t,x,y}$ (and its PDF f), it does not convey information about the joint probability of $u(t, x; \omega)$, $v(t, y; \omega)$ and $v(t, x; \omega)$, and therefore we cannot proceed further.

Our proofs above generalize without any change to the case of N -systems of linear hyperbolic equations.

Proposition 4.5. Consider a system of $N > 1$ linear hyperbolic equations with random initial data

$$(4.7) \quad \begin{cases} \vec{z}_t(t, x; \omega) + D(x) \vec{z}_x = 0, \\ \vec{z}(0, x, \omega) = \vec{z}_{(0,x)}(\omega), \end{cases}$$

where $D(x)$ can be diagonalized by an x -independent matrix P or, equivalently a system of the form

$$\begin{cases} u_t^i(t, x; \omega) + c_i(x) u_x^i = 0, \\ u^i(0, x, \omega) = u_{(0,x)}^i(\omega), \end{cases} \quad 1 \leq i \leq N.$$

Define the N -points pointwise statistics as the Young measure satisfying, for every measurable $A \subseteq \mathbb{R}^N$,

$$\phi_{t,x_1,\dots,x_N}(A) \equiv \mu \left(\left\{ \omega \in \Omega \mid (u^1(t, x_1; \omega), \dots, u^N(t, x_N; \omega))^T \in A \right\} \right).$$

Then the density $f(t, x_1, \dots, x_N, U_1, \dots, U_N)$ satisfies the PDE

$$(4.8) \quad \partial_t f(t, x_1, \dots, x_N, U_1, \dots, U_N) + \sum_{i=1}^N c_i(x_i) \partial_{x_i} f = 0.$$

The appearance of an $(N+1)$ -dimensional PDE in Proposition 4.5 is not an artifact of the method, but rather unavoidable, as demonstrated by the counterexample in section 4.1. From a computational perspective, solving (4.8) becomes increasingly challenging as N grows, due to the well-known curse of dimensionality. For moderate dimensions, such as $N = 2, 3, 4$, which arise in many applications, (4.8) can be efficiently treated using deterministic approaches, including sparse grid discretizations; see, e.g., [Gri01, BG04, GV10, Lub15]. For very large values of N , however, scalable and highly parallelizable Monte Carlo, particle-based, or conversely surrogate model approaches to (4.7) provide the only practical computational alternatives; see, e.g., [Spo91, JPT00, Gol16, KL07].

5. Regularity of the PDF and density estimation. Kernel density estimators (KDEs) are a class of nonparametric statistical methods for density estimation. As mentioned in the introduction, for random initial value problems such as (1.1), using KDEs (or any other Monte Carlo-based method) is a standard technique, especially when the dimension of the parameter space, Ω , is large. Here we ask the following.

QUESTION 5.1. *Given N samples from μ , and solving (1.1) to obtain $u(t, x; \omega)$ for each such sample, what is the expected accuracy of a KDE in estimating $f(t, x; U)$ or $F(t, x; U)$?*

Let us first briefly recall what KDEs are (see [Ize91, Tsy09, Was06] for broader expositions). Let $p \in L^1(\mathbb{R})$ be a PDF, let $K \in L^1(\mathbb{R})$ be a nonnegative kernel function, and choose a bandwidth parameter $h > 0$. Then, given $N \geq 1$ identically and independently distributed (i.i.d.) samples $y_1, \dots, y_N \sim p$, the associated KDE is

$$(5.1) \quad \hat{p}(y) = \hat{p}^{N,K,h}(y) \equiv \frac{1}{N} \sum_{j=1}^N K\left(\frac{y - y_j}{h}\right)$$

for all $y \in \mathbb{R}$. The expected error of $\hat{p}$ is a standard result [Tsy09, Theorem 1.1], presented here in simplified form.

Theorem 5.1 (Tsybakov [Tsy09]). *Let p be a PDF with $\beta \geq 1$ continuous derivatives and $\|p^{(\beta)}\|_\infty < L$; choose a kernel $K \in L^2(\mathbb{R})$ with $\int_{\mathbb{R}} K(u) du = 1$ and any $\alpha > 0$. Given N i.i.d. samples from the target, choose $h = h(N; \alpha) = \alpha N^{-1/(2\beta+1)}$. Then*

$$(5.2) \quad \sup_{y \in \mathbb{R}} \mathbb{E} |p(y) - \hat{p}(y)|^2 \leq C N^{-\frac{2\beta}{2\beta+1}}, \quad C = C(\alpha, \beta, L, K) > 0,$$

where the expectations are with respect to the realizations of the samples. Furthermore, for any fixed $\alpha > 0$ the error rate in (5.2) as a function of N is optimal.

Hence, to apply (5.2) to our settings, we need to know the regularity of $f(t, x, U)$ (or $F(t, x, U)$) a priori. Given a general initial value problem with random initial conditions, however, it is not trivial to a-priori guarantee the regularity of the PDF or CDF of the solution. Such guarantees can only be obtained for particular equations; see, e.g., [CDS10, JS22] for results on the regularity of the solution with respect to the parameters. We show next that for hyperbolic conservation laws, Theorems 2.1 and 3.1 provide the necessary regularity guarantees, and therefore a quantitative answer to Question 5.1, which only depends on $\phi_{0,x}$.

Consider first the linear scalar settings of (2.1). The following regularity statement is a direct corollary of Theorem 2.1.

Lemma 5.2. *Under the settings of Theorem 2.1, if $f_0(x, \cdot) \in C^\beta(\mathbb{R}_U)$ for some $\beta \in \mathbb{N}$ and for all $x \in \mathbb{R}$, then $f(t, y, \cdot) \in C^\beta(\mathbb{R}_U)$ for all $t, y \in \mathbb{R}$.*

Proof. Note that solutions to (2.2) are traveling waves in t and x . While the speed $c(x)$ might not be constant, it does not depend on U . Hence, for every $t, x \in \mathbb{R}$ there exists $x_0 = x_0(t, x) \in \mathbb{R}$ such that $f(t, x, U) = f_0(x_0, U)$. Note further that $x_0 = x_0(t, x)$ is determined by the method of characteristics, and therefore (by regularity theory of ODEs) it is a continuous function of x for every fixed $t \geq 0$. Therefore, while $\|f(t, x, \cdot)\|_{C^r(\mathbb{R}_U)}$ can generally depend on t , it cannot blow up in finite time. ■

With this regularity result and the statistical theory of (5.2), we can establish the statistical approximation error of applying KDEs to compute $f(t, x, U)$ in the linear scalar settings of (2.1).

Proposition 5.3. *Suppose $f_0(\cdot, U) \in C^1(\mathbb{R}_x)$ for all $U \in \mathbb{R}$ and that there exists $\beta \in \mathbb{N}$ and $L > 0$ such that*

$$\sup_{x \in \mathbb{R}} \sup_{U \in \mathbb{R}} \left| \partial_U^{(\beta)} f_0(x; U) \right| < L.$$

Let $\hat{f}(t, x, \cdot)$ be the KDE of $f(t, x, \cdot)$ using $N \geq 1$ i.i.d. samples from μ and solving (2.1) for each sample; see (5.1). Then, for every $x, U \in \mathbb{R}$ and every $T \geq 0$, there exists $C(T) \geq 0$ such that

$$\sup_{x \in \mathbb{R}, t \in [0, T]} \sup_{U \in \mathbb{R}_U} \mathbb{E} |f(t, x; U) - \hat{f}(t, x; U)|^2 \leq C(T) N^{-\frac{2\beta}{2\beta+1}}.$$

Proof. For any fixed $t \geq 0$ and $x \in \mathbb{R}$, an upper bound on $\sup_{U \in \mathbb{R}_U} \mathbb{E} |f(t, x; U) - \hat{f}(t, x; U)|^2$ is a straightforward application of (5.2) to $p = f(t, x, \cdot)$, combined with the regularity result in Lemma 5.2. For $t = 0$, a uniform bound in x is a consequence of the fact that $\partial_U^{(\beta)} f_0(x, U)$ is uniformly bounded in x as well. The uniform bound over all $t \in [0, T]$ is a consequence of the fact that the solution $f(t, x; U)$ maintains its regularity for all $t \geq 0$, and satisfies an upper bound of the form

$$\sup_{x \in \mathbb{R}} \sup_{U \in \mathbb{R}} \left| \partial_U^{(\beta)} f(t, x; U) \right| < L(t),$$

for some (potentially increasing but always finite) $L : [0, \infty) \rightarrow [0, \infty)$. This concludes the proof. Whether or not the interval $[0, T]$ can be taken to be $[0, \infty)$ depends on whether $L(t) \rightarrow \infty$ as $t \rightarrow \infty$, which in turn depends on the particulars of the medium heterogeneity function $c(x)$ in (2.1). ■

Now we turn to the nonlinear settings of (3.1). Here it is preferable to discuss CDF estimation (i.e., the statistical approximation of $F(t, x, U)$ based on samples of $u(t, x; \omega)$), since the regularity theory of F is more straightforward.

Lemma 5.4. *Assume the conditions of Theorem 3.1, and suppose that F_0 is $\beta \in \mathbb{N}$ differentiable in both x and U . Then there exists $C > 0$, independent of t , such that*

$$\sup_{x \in \mathbb{R}} \sup_{U \in \mathbb{R}} \|\partial_U^\beta F(t, x, U)\|_\infty \leq Ct^\beta$$

for all $t \geq 0$, for which Theorem 3.1 is valid.

Proof. Since $F(t, x, U) = F(0, x - a'(U)t, U)$, and denoting the partial derivatives of F with respect to the second and third coordinates by F_x and F_U , respectively, we have

$$\begin{aligned}\partial_U F(t, x, U) &= \partial_U F(0, x - a'(U)t, U) \\ &= -F_x(0, x - a'(U)t, U)a''(U)t + F_u(0, x - a'(U)t, U).\end{aligned}$$

Hence, if $F(0, \cdot, \cdot)$ is differentiable in x and U , then so is $F(t, \cdot, \cdot)$. Similarly, by induction, after β derivatives, $\partial_U^\beta F(t, x, U)$ is a sum of powers of t : the largest one is of the form

$$\partial_x^\beta F(0, x - a'(U)t, U) (a''(U)t)^\beta,$$

which completes the upper bound for large times $t \gg 1$. ■

The Kernel estimator of the CDF $\hat{F}(y)$ is the cumulative integral of the KDE $\hat{p}$ (see (5.1)) on the interval (∞, y) . The theory of these estimators is a consequence of the theory of KDEs, and we refer the reader to [BP09, Wan91] and the references therein for details. In much the same way as Proposition 5.3, we obtain the following.

Proposition 5.5. *Let $\hat{F}(t, x, \cdot)$ be the Kernel estimator of $F(t, x; \cdot)$ using $N \geq 1$ i.i.d. samples from μ and solving (3.1) for each sample. Then, for every $x, U \in \mathbb{R}$ and every $T \geq 0$, there exists $C > 0$ such that*

$$\sup_{x \in \mathbb{R}, t \in [0, T]} \sup_{U \in \mathbb{R}_U} \mathbb{E} |F(t, x; U) - \hat{F}^N(t, x; U)|^2 \leq CT^\beta N^{-\frac{\beta+1}{2\beta+1}},$$

under a choice of optimal Kernel width as specified in [BP09, Wan91].

6. Numerical examples.

6.1. Linear scalar case. Let us first demonstrate that, for a linear hyperbolic equation, the evolution of the pointwise statistics, $\phi_{t,x}$, through the corresponding PDFs $f(t, x, U)$, follows an evolution law which is (i) independent of the particular underlying random field $u(t, x; \omega)$ and (ii) described by Theorem 2.1.

6.1.1. Independence of spatial correlations. Here we consider two different random initial value problems, satisfying the same PDE. In both cases, $\phi_{0,x}$ are the same, but $u(0, x; \omega)$ is different from $v(0, x; \omega)$.

$$(6.1) \quad \begin{cases} \partial_t u(t, x; \omega) + c(x) \partial_x u = 0, & c(x) = 0.1x^2 + 1, \quad x \in [-3, 4], \\ u(0, x, \omega) = \omega(g(x+1) + g(x-1)), & g(x) = \max\{1 - (3x)^2, 0\}, \quad \omega \sim \text{unif}[0, 1], \end{cases}$$

$$(6.2) \quad \begin{cases} \partial_t v(t, x; \omega) + c(x) \partial_x v = 0, & c(x) = 0.1x^2 + 1, \quad x \in [-3, 4], \\ v(0, x, \omega) = \omega g(x+1) + (1-\omega)g(x-1), & g(x) = \max\{1 - (3x)^2, 0\}, \quad \omega \sim \text{unif}[0, 1]. \end{cases}$$

Not only do (6.1) and (6.2) look different, but they also represent different random fields. For example, the probability that $v(0, x; \omega) = v(0, -x; \omega)$ is zero for $x \in (-4/3, -2/3) \cup (2/3, 4/3)$, whereas $u(0, x; \omega) = u(0, -x; \omega)$ almost surely for these same values of x .

Nevertheless, denoting the pointwise statistics corresponding to u and v by $\phi_{t,x}^u$ and $\phi_{t,x}^v$, one can verify that $\phi_{0,x}^u = \phi_{0,x}^v$ for every $x \in \mathbb{R}$. Indeed, denoting the respective CDFs by F^u and F^v , we therefore have a single function $F_0: \mathbb{R}_x \rightarrow L^1(\mathbb{R}_U)$ with $F_0 = F^u(0, \cdot, \cdot) = F^v(0, \cdot, \cdot)$. Hence, by Theorem 3.1, we expect that F^v and F^u both satisfy the following PDE:

$$(6.3) \quad \begin{cases} \partial_t F(t, x; U) + c(x) \partial_x F(t, x; U) = 0, \\ F(0, x, U) = F_0(x, U). \end{cases}$$

The initial condition for the CDF is

- for $x \in [-\frac{2}{3}, \frac{2}{3}]$ or $x \leq -\frac{4}{3}$ or $x \geq \frac{4}{3}$:

$$F(0, x, U) = \begin{cases} 0, & U < 0, \\ 1, & U > 0; \end{cases}$$

- for $x \in [-\frac{4}{3}, -\frac{2}{3}]$:

$$F(0, x, U) = \begin{cases} 0, & U < 0, \\ \frac{U}{g(x+1)}, & 0 \leq U \leq g(x+1), \\ 1, & U \geq g(x+1); \end{cases}$$

- for $x \in [\frac{2}{3}, \frac{4}{3}]$:

$$F(0, x, U) = \begin{cases} 0, & U < 0, \\ \frac{U}{g(x-1)}, & 0 \leq U \leq g(x-1), \\ 1, & U \geq g(x-1). \end{cases}$$

The initial data for the PDF in (2.2) can be derived accordingly by taking the derivative of $F(0, x, U)$ with respect to U .

In practice, we solve (6.1) and (6.2) using the Monte Carlo method in w with 10^5 samples and reconstruct the PDF and CDF using a KDE for every x .⁷ For (6.1) and (6.2), the spatial and temporal derivatives are treated using the upwind method, with numerical parameters $\Delta x = 0.01$, $\Delta t = 10^{-3}$. For (6.3), we also use an upwind scheme with a slope limiter, applying the same discretization parameters. We discretize in U with $\Delta U = 0.0151$.

The results are collected in Figures 2 and 3. Figure 2 shows the CDF obtained from the three models individually, and the last plot displays the error between (6.1) and (6.2). The L^1 error between (6.1) and (6.3) is 0.0438. Figure 3 further compares the mean and variance across the three models, (6.1), (6.2), and (2.2), with errors recorded in the caption. In all, we observe an excellent match between the Monte Carlo computed pointwise statistics of u and v , and with the directly-evolved $f(t, x; U)$, as predicted.

6.1.2. Two-dimensional example. As noted in the introduction, having direct evolution equations for the pointwise statistics, $\phi_{t,x}$, implies that even if the underlying parameter space

⁷We use the MATLAB built-in function `ksdensity`.

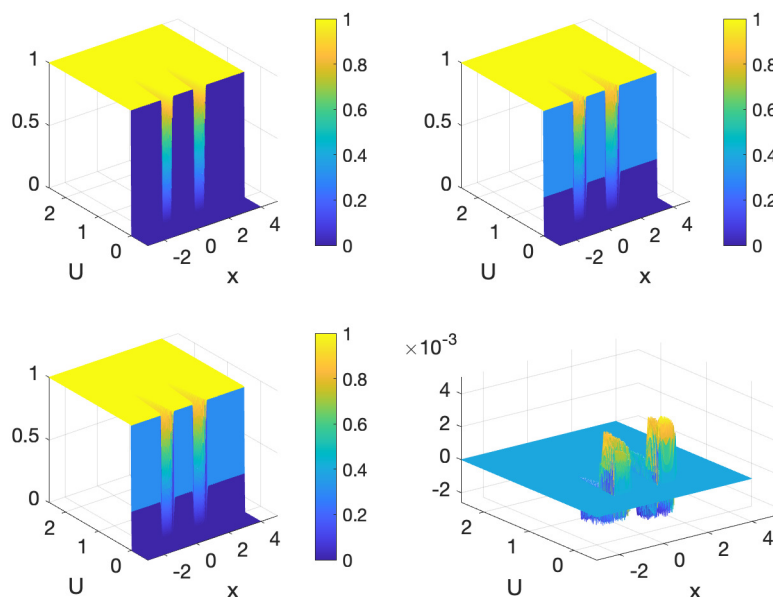


Figure 2. Scalar, linear hyperbolic equations: CDF $F(0.2, x, U)$ corresponding to, from left to right, top to bottom (6.3), (6.1), and (6.2). The last plot is the error between (6.1) and (6.2). The L^1 error between the solution to (6.3) and (6.1) is 0.0438.

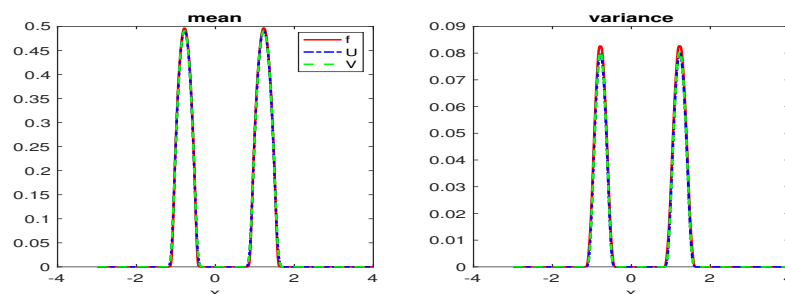


Figure 3. Same settings as in Figure 2. Comparison of mean and variance. The L^1 error in mean between U and V is $5.7607e-04$; and between U and f it is 0.0136. Likewise, the L^1 error in mean between U and V is $4.8364e-04$ and between V and f is 0.0044.

Ω is high-dimensional, we can solve the (fixed-dimensional) equations for the PDF or CDF instead. To see this numerically, consider

$$(6.4) \quad \begin{cases} \partial_t u(t, x; \vec{\gamma}) = c(x) \partial_x u, & c(x) = 1 - 0.1x^2, \\ u(0, x, \vec{\gamma}) = u_{(0,x)}(\vec{\gamma}) := \gamma_1 g_1(x) + \gamma_2 g_2(x), \end{cases}$$

where $\vec{\gamma} = (\gamma_1, \gamma_2) \in \mathbb{R}^2$ is distributed according to the normal distribution with mean $(0, 1)^\top$ and covariance matrix $\text{diag}(1, 0.25)$ and where $g_1(x) = e^{10(x+1)^2} + 1$ and $g_2(x) = e^{10(x-1)^2} + 1$. By Theorem 2.1, f should satisfy (2.2) with $f_{(0,x)}(U) = \mathcal{N}(g_2(x), g_1(x)^2 + 0.5^2 g_2(x)^2)$ for any x in the computational domain.

Figure 4 compares the PDFs obtained by a Monte Carlo simulation of (6.4) and by solving the PDF evolution equation (2.2). Similarly, Figure 5 compares the mean and variance

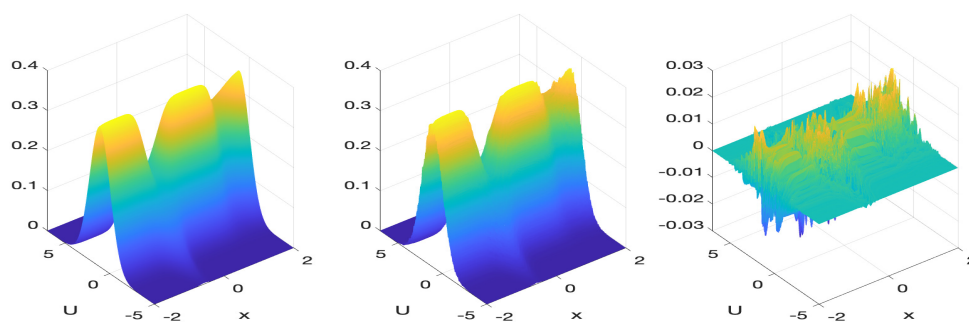


Figure 4. The PDF $f(t=0.4, x, U)$ for (6.4). Left is computed using (2.2) and middle is using (6.4) with 10^5 samples. The right plot is the difference between the two.

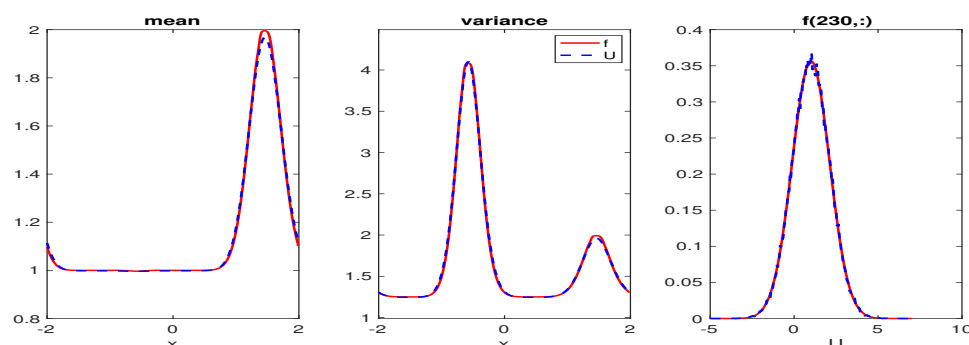


Figure 5. Same settings as in Figure 4. Comparison of mean, variance, and PDF (at $x=0.29$) at $t=0.4$, using a direct solution of (2.2) (red, solid) and Monte Carlo simulations of (6.4) (blue, dashed). The L^1 errors between the mean and variance are 0.0266 and 0.0554, respectively.

obtained from these two models, showing good agreement. The L^1 errors between them are 0.0266 and 0.0554, respectively. The right plot in Figure 5 displays $f(0.4, 0.29, U)$, where we observe a good agreement between the results. To explain these errors, first note that in our method, since $f(0, x, U)$ can be computed in closed form, the only error is that associated with solving the PDE for f , (2.2). In this one-dimensional PDE, this error can be controlled relatively well with the numerical parameters we report below. Hence, the source of the error is rather due to the Monte Carlo KDE estimator. The error of such nonparametric statistical estimators converges quite slowly in the number of samples (see Theorem 5.1). Indeed, we also observe numerically that increasing the number of samples does dramatically reduce the overall error (results not shown).

In our examples, we set $x \in [-2, 3]$ and apply a periodic boundary condition. For (2.2), we use $\Delta x = 0.01$, $\Delta t = 0.002$, and $\Delta U = 0.012$. For the Monte Carlo method applied to (6.4), we use $\Delta x = 0.01$, $\Delta t = 0.0025$, and 10^5 samples. The density for (6.4) is reconstructed using `ksdensity` with kernel width 0.02.

6.2. Nonlinear rarefaction waves. To demonstrate the validity of Theorems 3.1 and 3.2, we consider the Burgers' equation in the *absence of shock*; i.e., the initial data are set such that a rarefaction wave emerges almost surely:

$$(6.5) \quad \begin{cases} \partial_t u + \frac{1}{2} \partial_x (u^2) = 0, \\ u(0, x, \omega) = \omega(g_1(x-1) + g_2(x+1)), \quad \omega \sim \text{unif}[0, 1], \end{cases}$$

where

$$(6.6) \quad g_1(x) = \frac{1}{2}(1 + \tanh(40x)), \quad g_2(x) = \frac{1}{2}(\tanh(40x) - 1) = g_1(x) - 1.$$

Here, as in the linear case of section 6.1.1, we will show that only $\phi_{0,x}$ matters to determine $\phi_{t,x}$ for $t > 0$. Hence, we consider a different set of random data with the same initial pointwise statistics:

$$(6.7) \quad \begin{cases} \partial_t v + \frac{1}{2} \partial_x (v^2) = 0, \\ u(0, x, \omega) = \omega g_2(x+1) + (1-\omega)g_1(x-1), \quad \omega \sim \text{unif}[0, 1]. \end{cases}$$

One can check directly that the CDFs for $u(0, x, \omega)$ and $v(0, x, \omega)$ are identical for each x , up to exponentially small terms. To calculate this common CDF, $F_0(x; U)$, define $[a, b] = [-0.25, 0.25]$ as the interval on which $g_1(x-1) + g_2(x+1) \approx 0$, where we approximate $g_1(x-1) \approx 0$ for $x < a$ and $g_2(x+1) \approx 0$ for $x > b$. Therefore, by direct computation,

- for $x \in [a, b]$:

$$F_0(x, U) = \begin{cases} 0, & U < 0, \\ 1, & U > 0; \end{cases}$$

- for $x < a$:

$$F_0(x, U) = \begin{cases} 1, & U \geq 0, \\ 1 - \frac{U}{g_2(x+1)}, & 0 \geq U \geq g_2(x+1), \\ 0, & U < g_2(x+1); \end{cases}$$

- for $x > b$:

$$F_0(x, U) = \begin{cases} 0, & U < 0, \\ \frac{U}{g_1(x-1)}, & 0 \leq U \leq g_1(x-1), \\ 1, & U \geq g_1(x-1). \end{cases}$$

An alternative way of computing $F_0(x, U)$ is

$$(6.8) \quad F_0(x, U) = \begin{cases} \min \left\{ \max \left\{ \frac{U}{g_1(x-1) + g_2(x+1)}, 0 \right\}, 1 \right\}, & g_1(x-1) + g_2(x+1) > 0, \\ \max \left\{ 1 - \max \left\{ \frac{U}{g_1(x-1) + g_2(x+1)}, 0 \right\}, 0 \right\}, & g_1(x-1) + g_2(x+1) < 0, \\ \mathbf{1}_{U \geq 0}, & g_1(x-1) + g_2(x+1) = 0, \end{cases}$$

where for any set B we denote by $\mathbf{1}_B$ the characteristic function of B .

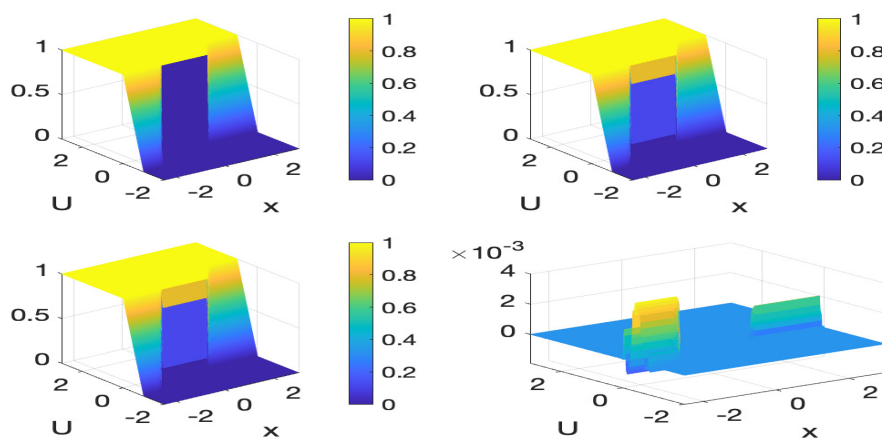


Figure 6. Comparison of CDF of random rarefaction solutions at time $t=0.8$. From top to bottom, left to right, the results are obtained by solving (3.2) (for F), (6.5), and (6.7), respectively. The bottom right plot is the difference between (6.5) and (6.7), and the L^1 error between (3.2) and (6.7) is 0.0179.

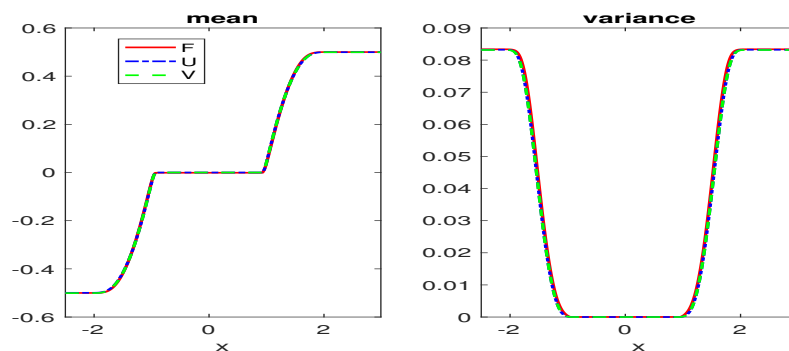


Figure 7. Same settings as Figure 6. Comparison of mean and variance at time $t=0.8$ between three models: u (6.5) (dashed blue), v (6.7) (dashed green), and the directly-evolved CDF, F , according to (3.2) (solid red).

From the expression of $F_0(x, U)$ in (6.8), we see that F_0 is close to a function which is not differentiable everywhere, and therefore $f(t, x; U) = \partial_U F(t, x; U)$ is very large at this point. Therefore, it is more prudent to evolve the CDF equation according to (3.2) and then compute the PDF from it using central-difference numerical differentiation.

Both the CDF and the resulting mean and covariance are compared to direct Monte Carlo simulations of (6.5) and (6.7), as shown in Figures 6 and 7. The Monte Carlo simulations use 10^5 samples. As in the linear case discussed previously, the results are nearly indistinguishable. The discretization parameters used are $\Delta x = 0.01$, $\Delta t = 0.0025$, and $\Delta U = 0.01$.

We now turn to demonstrating the nonlocal PDE (3.4) for evolving the PDF f . First, as noted, differentiate the CDF obtained above with respect to U to obtain (a numerical approximation of) $f(t, x; U)$. Alternatively, we compute $f_0(x, U)$ by differentiating $F_0(x; U)$ (see (6.8)) and then we evolve the numerically computed PDF according to (3.4) and compare the two approximations of f at $t=0.8$. The scheme by which we solve (3.4) is described in Appendix A.

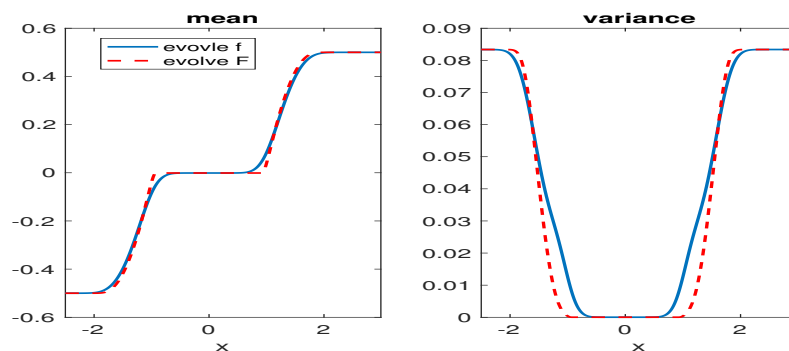


Figure 8. For the rarefaction wave settings of (6.5), we compare the mean and variance at $t = 0.8$ as obtained through evolving the PDF using (3.4) (blue, solid) and by evolving the CDF using (3.2) (red, dashed).

The mean and covariance for these two solutions are shown in Figure 8. It is observed that the “differentiate and evolve” approach (based on (3.4)) is slightly more diffusive than the “evolve then differentiate” (based on $F(t, x; U)$) approach.

Remark 6.1 (evolving the CDF vs the PDF: practical considerations). Computationally, evolving the CDF is often preferable for the following reasons: Estimating the initial CDF from samples is statistically more stable than estimating a density. Furthermore, the CDF satisfies a simple linear transport equation, whereas the evolution equation for the PDF is nonlocal and requires solving in parallel for many values of U . Last, the CDF is well-defined even when the measure $\phi_{t,x}$ has singular components.

On the other hand, certain applications require the PDF directly, for instance when densities are used for visualization or as a basis for sampling algorithms. A prominent example is Markov chain Monte Carlo, where the acceptance step depends on evaluating the target density (up to normalization); see, e.g., [BGJM11, RC10]. In such cases, evolving the PDF avoids numerical differentiation of a CDF, which may introduce additional error. Both approaches are valid, and the choice depends on the intended use.

6.3. Nonlinear case: Shock. As in the previous examples, we study the effect of spatial correlations by comparing two distinct initial conditions that share identical pointwise statistics $\phi_{0,x}$. We will demonstrate that, although the pointwise statistics agree before the creation of the shock, they diverge afterwards.

Specifically, the equation for $u(t, x; \omega)$ reads as

$$(6.9) \quad \begin{cases} \partial_t u + \frac{1}{2} \partial_x (u^2) = 0, \\ u(0, x, \omega) = \omega \left(g_1 \left(x - \frac{1}{2} \right) + g_2 \left(x + \frac{1}{2} \right) \right), \quad \omega \sim \text{unif}[0, 1], \end{cases}$$

where

$$g_1(x) = \begin{cases} 0, & x < -\frac{1}{4}, \\ -2x + \frac{1}{2}, & -\frac{1}{4} \leq x < \frac{1}{4}, \\ -1, & x \geq \frac{1}{4}, \end{cases} \quad g_2(x) = g_1(x) + 1.$$

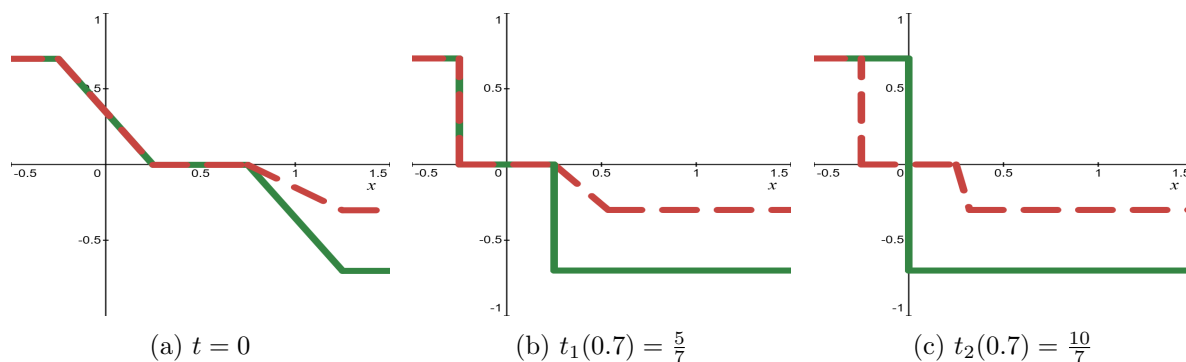


Figure 9. Illustration of $u(t, x; \omega)$ (solid green) and $v(t, x; \omega)$ (dashed red) for $\omega = 0.7$; see (6.9) and (6.10), respectively. Left: $t = 0$. Middle: $t_1(0.7) = 5/7$. Here $u(t_1(\omega), x; 0.7)$ developed two shocks, whereas $v(t_1(\omega), x; 0.7)$ developed only one. Right: $t_2(0.7) = 10/7$. Here the two shocks of u have merged, whereas v still has only one shock.

The equation for $v(t, x; \omega)$ reads as

$$(6.10) \quad \begin{cases} \partial_t v + \frac{1}{2} \partial_x (v^2) = 0, \\ v(0, x, \omega) = \omega g_2 \left(x + \frac{1}{2} \right) + (1 - \omega) g_1 \left(x - \frac{1}{2} \right), \quad \omega \sim \text{unif}[0, 1]. \end{cases}$$

What do typical solutions u or v look like? $u_0(x; \omega)$ and $v_0(x; \omega)$ both consist of two consecutive step-like structures that develop a shock at finite time; see Figure 9, left panel. By explicit calculation (using the method of characteristics), the left step develops into a shock for both u and v at time $t_1(\omega) = (2\omega)^{-1} \in [1/2, \infty)$. For u , the right step also develops into a shock at $t_1(\omega)$. For v , however, the shock develops at $t'_1(\omega) = (2(1 - \omega))^{-1}$ (see Figure 9, middle panel). Once formed, the right shocks of u and v travel (left) at speeds of $-\omega/2$ and $-(1 - \omega)/2$, respectively, whereas the left shocks travel (right) at a speed of $\omega/2$. Therefore, at $t_2(\omega) = \omega^{-1} \in [1, \infty)$ the two shocks of $u(t, x; \omega)$ merge at $x = 0$, where they remain for all $t \geq t_2(\omega)$, i.e., with probability 1. For v , however, the shocks merge at different times and will travel to the right at speed 1.

Notwithstanding the differences in the dynamics, the initial CDF corresponding to both u_0 and v_0 , denoted for simplicity by $F_0(x, v)$, is given by

- for $x \in [-0.25, 0.25]$:

$$F_0(x, U) = \begin{cases} 0, & U < 0, \\ 1, & U > 0; \end{cases}$$

- for $x > 0.25$:

$$F_0(x, U) = \begin{cases} 1, & U \geq 0, \\ 1 - \frac{U}{g_1(x - \frac{1}{2})}, & 0 \geq U \geq g_1(x - \frac{1}{2}), \\ 0, & U < g_1(x - \frac{1}{2}); \end{cases}$$

- for $x < -0.25$:

$$F_0(x, U) = \begin{cases} 0, & U < 0, \\ \frac{U}{g_2(x + \frac{1}{2})}, & 0 \leq U \leq g_2(x + \frac{1}{2}), \\ 1, & U \geq g_2(x + \frac{1}{2}). \end{cases}$$

Again, we compare the directly-computed (using Monte Carlo) CDF and PDF corresponding to u and v with the directly evolved versions according to (3.2) and (3.4), respectively. In the following, we compare the solutions at three distinct time points relative to the times of shock formation, as shown in Figure 9 for illustration. In all these simulations, the Monte Carlo simulations use 2×10^4 samples, and the discretization parameters used are $\Delta x = 0.01$, $\Delta t = 0.0025$, and $\Delta U = 0.01$.

Figures 10 and 11 present the results at $t = 0.4$, prior to the formation of any shocks, i.e., $t < \min_{\omega} t_1(\omega) = t_1(1) = 0.5$. Since no shocks have formed, the three solutions agree as expected (section 3).

At $t = 0.8$ (see Figures 12–14), u has developed two shocks for all $\omega > 5/8$ and none otherwise. On the other hand, v developed exactly one shock, whose location depends on the value of ω . Here the analysis of section 3 is no longer valid, and indeed, as Remark 3.3 implies,

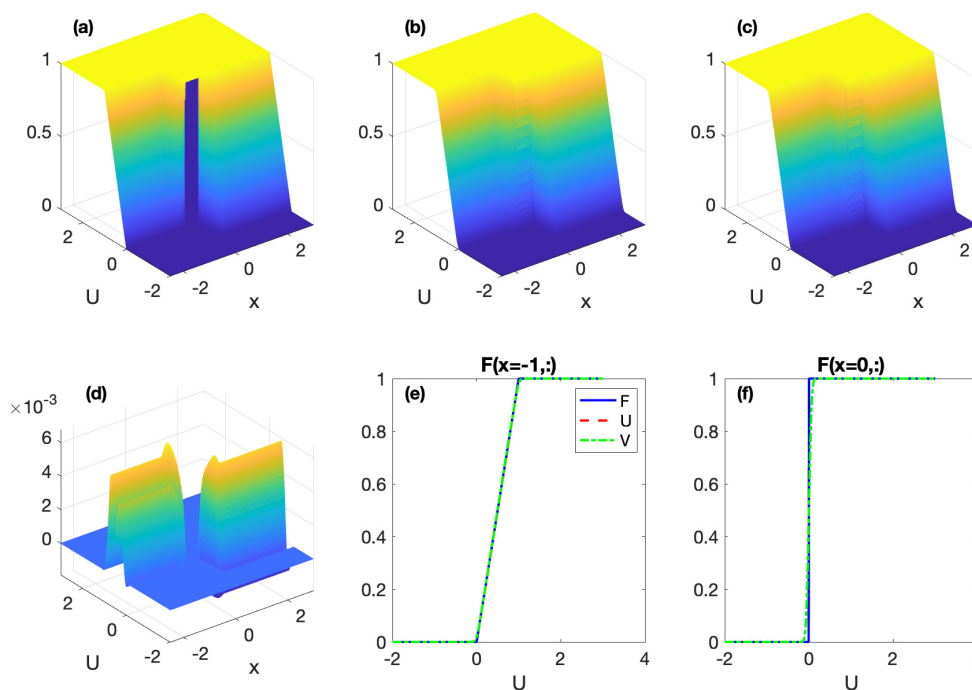


Figure 10. Plots (a)–(c) show the results at $t = 0$ of solving (3.2) for F and of performing direct Monte Carlo simulations of (6.5) and (6.7), respectively. Plot (d) displays the difference between the solutions of (6.5) and (6.7). Although the 3D plot may visually suggest a significant discrepancy (based on color) between plots (a) and (b) or (c), the actual L^1 error between (a) and (b) is only 0.0468. To further illustrate this point, plots (e) and (f) show the CDFs at slices $x = -1$ and $x = 0$, respectively, with the direct solution of F (solid blue), u (dashed red), and v (dashed green).

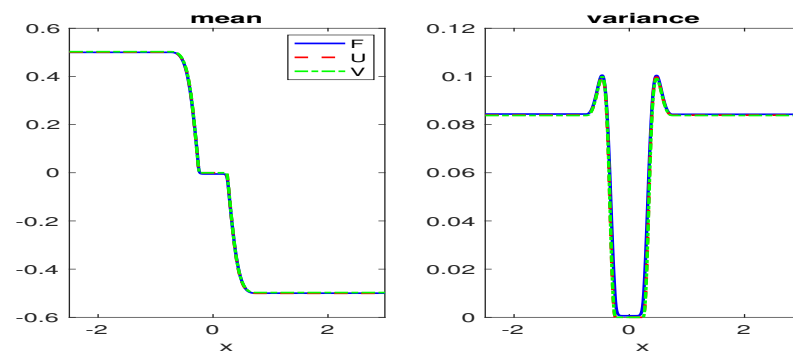


Figure 11. Same settings as Figure 10, comparison of mean (left) and variance (right): u (6.9) (dashed red), v (6.10) (dashed green), and the direct evolution of the CDF via (3.2) (solid blue).

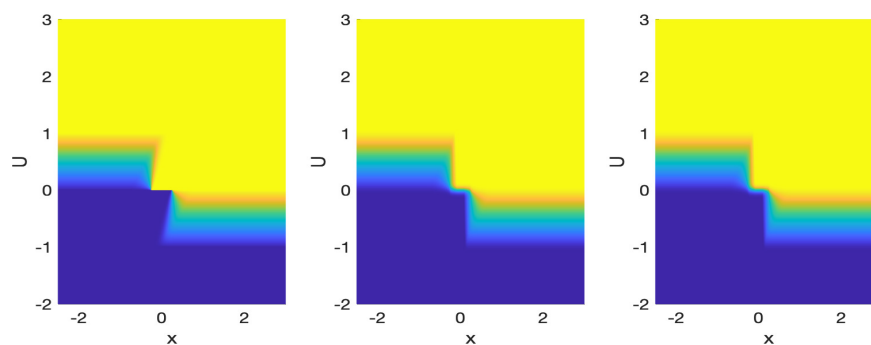


Figure 12. Settings of section 6.3 at $t = 0.8$. Comparison of the CDFs obtained from the directly evolved CDF (3.2) (left), u (6.9) (middle), and v (6.10) (right).

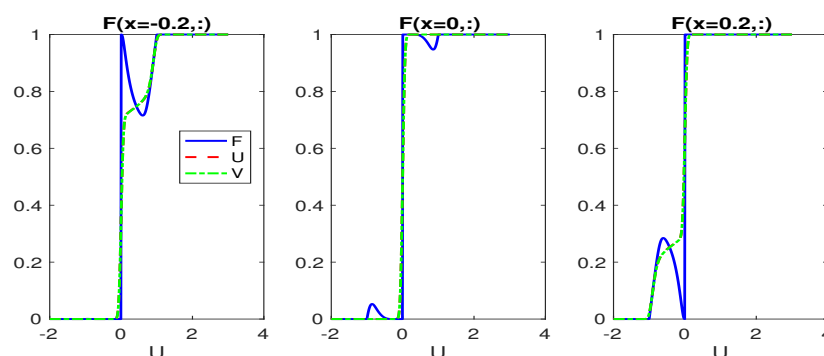


Figure 13. Settings of section 6.3 at $t = 0.8$. Cross-sections of Figure 12 at three different locations, $x = -0.2$, $x = 0$, and $x = 0.2$, with direct solution of F (solid blue), u (dashed red), and v (dashed green).

the directly-computed PDF $f(t, x; U)$ does not agree with those of u and v . What may be surprising is that the directly-computed (via Monte Carlo simulations) pointwise statistics of u and v agree. This is due to a symmetry between u and v : for $x < 0$, they behave exactly the same at $\min_{\omega} t_1(\omega) < t < \min_{\omega} t_2(\omega)$. Furthermore, while they differ for $x > 0$, there is a

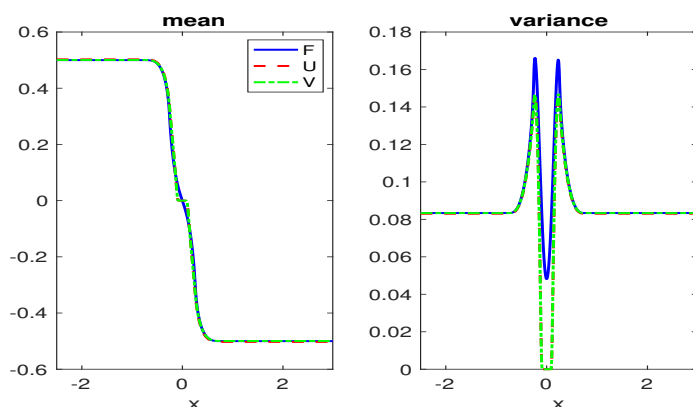


Figure 14. Same settings and legend as Figure 13. Comparison of mean and variance.

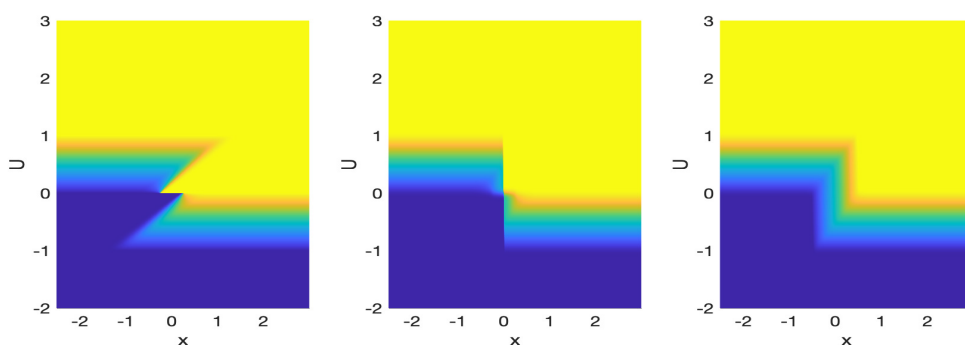


Figure 15. Settings of section 6.3 at $t = 2.0$. Comparison of CDF obtained from new model (3.2) (left), u (6.9) (middle), and v (6.10) (right).

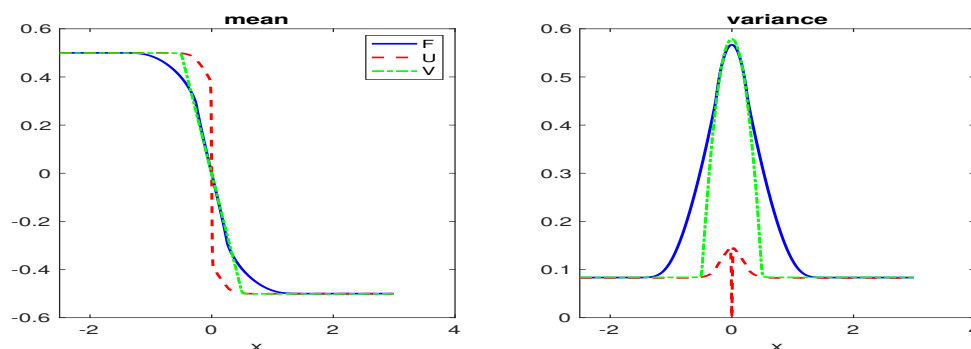


Figure 16. Settings of section 6.3 at $t = 2.0$. Comparison of mean and variance for the directly solved F (solid blue), u (dashed red), and v (dashed green).

one-to-one correspondence between $u(t, x; \omega)$ and $v(t, x; 1 - \omega)$ for these x values, and therefore the pointwise statistics agree.

Finally, we consider $t = 2 > t_2(\omega)$ for $\omega > 1/2$, when the two shocks have collided and merged for u (see Figure 9, right panel). As can be seen in Figures 15 and 16, the pointwise

statistics of u and v already disagree here (since the shocks have not merged for v , or indeed might not have formed yet).

Appendix A. Numerical scheme for the nonlocal equation (3.4). First, reformulate (3.4) into a conservative form in x and then adopt the central upwind scheme developed in [KNP01]. More precisely, rewrite (3.4) into

$$(A.1) \quad \partial_t f + \partial_x \left[a''(U) \int_{-\infty}^U f(t, x, v) dv + a'(U) f(U) \right] = 0.$$

Let

$$F(f) = \frac{\partial}{\partial U} \left[a'(U) \int_{-\infty}^U f(t, x, v) dv \right],$$

and denote the flux; then (A.1) reduces to

$$(A.2) \quad \partial_t f + \partial_x F(f) = 0,$$

which can be considered as a conservation law for f . In this regard, many shock-capturing schemes can be employed. We select the central upwind scheme for its simplicity and high-order accuracy.

Choose the computational domain $x_L \leq x_R$, $U_L \leq U \leq U_R$, and $t \leq 0$, and denote

$$f_{jk}^n = f(t_n, x_j, U_k), \quad 0 \leq n \leq N_t, \quad 0 \leq j \leq N_x, \quad 0 \leq k \leq N_U,$$

where $t_n = n\Delta t$, $x_j = x_L + j\Delta x$, $U_k = U_L + k\Delta U$. Here Δt , Δx , and ΔU are time steps and mesh sizes in x and U , respectively.

Remark. We note that (3.4) itself does not require any a priori knowledge of the range of U , since the integral is defined over $(-\infty, U)$ for every $U \in \mathbb{R}$. In numerical computations, however, one must choose a finite computational domain $[U_L, U_R]$. This can be consistently estimated from the range of the initial data: values of U which are not attained at $t = 0$ (or attained with probability zero) will not be attained for all times.

Then (A.2) can be discretized as

$$(A.3) \quad \frac{f_{jk}^{n+1} - f_{jk}^n}{\Delta t} + \frac{H_{j+1/2,k} - H_{j-1/2,k}}{\Delta x} = 0,$$

where

$$H_{j+1/2,k} = \frac{a_{j+1/2}^+ F(f_{j+1/2,k}^-) - a_{j+1/2}^- F(f_{j+1/2,k}^+)}{a_{j+1/2}^+ - a_{j+1/2}^-} + \frac{a_{j+1/2}^+ a_{j+1/2}^-}{a_{j+1/2}^+ - a_{j+1/2}^-} (f_{j+1/2,k}^+ - f_{j+1/2,k}^-)$$

and

$$a_{j+1/2}^+ = \max \left\{ \left(\frac{\partial F}{\partial f} \right)_{j+1/2,k}^n, 0 \right\}, \quad a_{j+1/2}^- = \min \left\{ \left(\frac{\partial F}{\partial f} \right)_{j+1/2,k}^n, 0 \right\}.$$

Here $f_{j+1/2}^\pm$ are constructed using the minmod slope limiter. That is,

$$f_{j+1/2,k}^- = f_{j,k} + \frac{\Delta x}{2}(f_x)_{j,k}, \quad f_{j+1/2,k}^+ = f_{j+1,k} - \frac{\Delta x}{2}(f_x)_{j+1,k},$$

with

$$(f_x)_{j,k} = \text{minmod} \left(\theta \frac{f_{j,k} - f_{j-1,k}}{\Delta x}, \frac{f_{j+1,k} - f_{j-1,k}}{2\Delta x}, \theta \frac{f_{j+1,k} - f_{j,k}}{\Delta x} \right),$$

and

$$\text{minmod}(z_1, z_2, \dots) = \begin{cases} \min_j \{z_j\}, & z_j > 0 \quad \forall j, \\ \max_j \{z_j\}, & z_j < 0 \quad \forall j, \\ 0 & \text{otherwise.} \end{cases}$$

Note that without the slope limiter, i.e., $(f_x)_{j,k} \equiv 0$ for any j , and $a_{j+1/2}$ keeps its sign, then the scheme (A.3) reduces to the upwind scheme.

Acknowledgments. L. Wang would like to thank Professors Shi Jin and Daniel Tartakovsky for fruitful discussions.

REFERENCES

- [AM17] R. ABGRALL AND S. MISHRA, *Uncertainty quantification for hyperbolic systems of conservation laws*, in Handbook of Numerical Analysis, 2017, pp. 507–544.
- [BG04] H.-J. BUNGARTZ AND M. GRIEBEL, *Sparse grids*, Acta Numer., 13 (2004), pp. 147–269, <https://doi.org/10.1017/S0962492904000182>.
- [BGJM11] S. BROOKS, A. GELMAN, G. L. JONES, AND X.-L. MENG, *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC, 2011, <https://doi.org/10.1201/b10905>.
- [BNTT11] J. BÄCK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Stochastic spectral Galerkin and Collocation Methods for PDEs with random coefficients: A numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations: Selected Papers from the ICOSA-HOM'09 Conference, Trondheim, Norway, 2011, pp. 43–62.
- [BP09] A. BERG AND D. POLITIS, *CDF and survival function estimation with infinite-order kernels*, Electron. J. Statist., 3 (2009), pp. 1436–1454, <https://doi.org/10.1214/09-EJS396>.
- [BT16] F. BOSO AND D. M. TARTAKOVSKY, *The method of distributions for dispersive transport in porous media with uncertain hydraulic properties*, Water Resour. Res., 52 (2016), pp. 4700–4712, <https://doi.org/10.1002/2016WR018745>.
- [BT20] F. BOSO AND D. M. TARTAKOVSKY, *Data-informed method of distributions for hyperbolic conservation laws*, SIAM J. Sci. Comput., 42 (2020), pp. A559–A583, <https://doi.org/10.1137/19M1260773>.
- [B⁺95] E. J. BALDER, *Lectures on Young Measures*, Cahiers de Mathématiques de la Décision 9517, 1995.
- [CDS10] A. COHEN, R. DEVORE, AND C. SCHWAB, *Convergence rates of best N-term Galerkin approximations for a class of elliptic SPDEs*, Found. Comput. Math., 10 (2010), pp. 615–646, <https://doi.org/10.1007/s10208-010-9072-2>.
- [CGH05] Q.-Y. CHEN, D. GOTTLIEB, AND J. S. HESTHAVEN, *Uncertainty analysis for the steady-state flows in a dual throat nozzle*, J. Comput. Phys., 204 (2005), pp. 378–398, <https://doi.org/10.1016/j.jcp.2004.10.019>.
- [CNP⁺18] I. COLOMBO, F. NOBILE, G. PORTA, A. SCOTTI, AND L. TAMELLINI, *Uncertainty quantification of geochemical and mechanical compaction in layered sedimentary basins*, Comput. Methods Appl. Mech. Engrg., 328 (2018), pp. 122–146, <https://doi.org/10.1016/j.cma.2017.08.049>.

- [DFS20] A. DITKOWSKI, G. FIBICH, AND A. SAGIV, *Density estimation in uncertainty propagation problems using a surrogate model*, SIAM/ASA J Uncertain. Quantif., 8 (2020), pp. 261–300, <https://doi.org/10.1137/18M1205959>.
- [Eva22] L. C. EVANS, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, 2022.
- [FKMT17] U. S. FJORDHOLM, R. KÄPPELI, S. MISHRA, AND E. TADMOR, *Construction of approximate entropy measure-valued solutions for hyperbolic systems of conservation laws*, Found. Comput. Math., 17 (2017), pp. 763–827, <https://doi.org/10.1007/s10208-015-9299-z>.
- [FLM17] U. S. FJORDHOLM, S. LANTHALER, AND S. MISHRA, *Statistical solutions of hyperbolic conservation laws: Foundations*, Arch. Ration. Mech. Anal., 226 (2017), pp. 809–849, <https://doi.org/10.1007/s00205-017-1145-9>.
- [FLMW20] U. S. FJORDHOLM, K. LYE, S. MISHRA, AND F. WEBER, *Statistical solutions of hyperbolic systems of conservation laws: Numerical approximation*, Math. Models Methods Appl. Sci., 30 (2020), pp. 539–609, <https://doi.org/10.1142/S0218202520500141>.
- [GHO⁺17] R. GHANEM, D. HIGDON, AND O. HOUMAN, EDs., *Handbook of Uncertainty Quantification*, Vol. 6, Springer, New York, 2017.
- [Gol16] F. GOLSE, *On the dynamics of large particle systems in the mean field limit*, in Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity, Lect. Notes Appl. Math. Mech. 3, Springer, 2016, pp. 1–144, <https://doi.org/10.1007/978-3-319-26883-5>.
- [Gri01] M. GRIEBEL, *The curse of dimensionality in numerical analysis*, Acta Numer., 10 (2001), pp. 1–73.
- [GV10] M. GRIEBEL AND J. O. R. VOGT, *Sparse grid discretizations of the transport equation*, Numer. Math., 114 (2010), pp. 1–26.
- [GZ07] B. GANAPATHYSUBRAMANIAN AND N. ZABARAS, *Sparse grid collocation schemes for stochastic natural convection problems*, J. Comput. Phys., 225 (2007), pp. 652–685, <https://doi.org/10.1016/j.jcp.2006.12.014>.
- [Ize91] A. J. IZENMAN, *Review papers: Recent developments in nonparametric density estimation*, J. Am. Stat. Assoc., 86 (1991), pp. 205–224, <https://doi.org/10.1080/01621459.1991.10475021>.
- [JP17] S. JIN AND L. PARESCHI, *Uncertainty Quantification for Hyperbolic and Kinetic Equations*, Springer, 2017.
- [JPT00] S. JIN, L. PARESCHI, AND G. TOSCANI, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numer. Anal., 38 (2000), pp. 913–936, <https://doi.org/10.1137/S0036142998347978>.
- [JS22] T. JAHNKE AND B. STEIN, *A multilevel stochastic collocation method for Schrödinger equations with a random potential*, SIAM/ASA J. Uncertain. Quantif., 10 (2022), pp. 1753–1780, <https://doi.org/10.1137/21M1440517>.
- [JX95] S. JIN AND Z. XIN, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Commun. Pure Appl. Math., 48 (1995), pp. 235–276, <https://doi.org/10.1002/cpa.3160480303>.
- [KL07] O. KOCH AND C. LUBICH, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454, <https://doi.org/10.1137/050639703>.
- [KNP01] A. KURGANOV, S. NOELLE, AND G. PETROVA, *Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton–Jacobi equations*, SIAM J. Sci. Comput., 23 (2001), pp. 707–740, <https://doi.org/10.1137/S1064827500373413>.
- [LMMKH10] O. P. LE MAÎTRE, L. MATHELIN, O. M. KNIO, AND M. Y. HUSSAINI, *Asynchronous time integration for polynomial chaos expansion of uncertain periodic dynamics*, Discrete Contin. Dyn. Syst. Ser. A, 27 (2010), pp. 199–226, <https://doi.org/10.3934/dcds.2010.28.199>.
- [Lub15] C. LUBICH, *Time integration in high dimensions*, Acta Numer., 24 (2015), pp. 209–298.
- [MJM13] F. MÜLLER, P. JENNY, AND D. W. MEYER, *Multilevel Monte Carlo for two phase flow and Buckley–Leverett transport in random heterogeneous porous media*, J. Comput. Phys., 250 (2013), pp. 685–702, <https://doi.org/10.1016/j.jcp.2013.03.023>.

- [MS12] S. MISHRA AND C. SCHWAB, *Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data*, Math. Comp., 81 (2012), pp. 1979–2018, <https://doi.org/10.1090/S0025-5718-2012-02574-9>.
- [MSŠ13] S. MISHRA, C. SCHWAB, AND J. ŠUKYS, *Multi-level Monte Carlo finite volume methods for uncertainty quantification in nonlinear systems of balance laws*, in Uncertainty Quantification in Computational Fluid Dynamics, Lect. Notes Comput. Sci. Eng. 92, Springer, 2013, pp. 225–294, <https://doi.org/10.1007/978-3-319-00885-1>.
- [NDM⁺09] H. N. NAJM, B. J. DEBUSSCHERE, Y. M. MARZOUK, S. WIDMER, AND O. P. LE MAÎTRE, *Uncertainty quantification in chemical systems*, Internat. J. Numer. Methods Engrg., 80 (2009), pp. 789–814, <https://doi.org/10.1002/nme.2551>.
- [Pav14] G. A. PAVLIOTIS, *Stochastic Processes and Applications*, Texts Appl. Math. 60, Springer, 2014.
- [PDL09] G. POËTTE, B. DESPRÉS, AND D. LUCOR, *Uncertainty quantification for systems of conservation laws*, J. Comput. Phys., 228 (2009), pp. 2443–2467, <https://doi.org/10.1016/j.jcp.2008.12.018>.
- [Ped97] P. PEDREGAL, *Parametrized Measures and Variational Principles*, Springer Science & Business Media, 1997.
- [Peh20] B. PEHERSTORFER, *Model reduction for transport-dominated problems via online adaptive bases and adaptive sampling*, SIAM J. Sci. Comput., 42 (2020), pp. A2803–A2836, <https://doi.org/10.1137/19M1257275>.
- [PGS⁺19] G. PATWARDHAN, X. GAO, A. SAGIV, A. DUTT, J. GINSBERG, A. DITKOWSKI, G. FIBICH, AND A. L. GAETA, *Loss of polarization of elliptically polarized collapsing beams*, Phys. Rev. A, 99 (2019), 033824, <https://doi.org/10.1103/PhysRevA.99.033824>.
- [PIN09] P. PETTERSSON, G. IACCARINO, AND J. NORDSTRÖM, *Numerical analysis of the Burgers' equation in the presence of uncertainty*, J. Comput. Phys., 228 (2009), pp. 8394–8412, <https://doi.org/10.1016/j.jcp.2009.08.012>.
- [RC10] C. P. ROBERT AND G. CASELLA, *Monte Carlo Statistical Methods*, 2nd ed., Springer, 2010.
- [Ris96] H. RISKEN, *Fokker-Planck Equation*, Springer, 1996.
- [Sag20] A. SAGIV, *The Wasserstein distances between pushed-forward measures with applications to uncertainty quantification*, Commun. Math. Sci., 18 (2020), pp. 707–724, <https://doi.org/10.4310/CMS.2020.v18.n3.a6>.
- [Sag22] A. SAGIV, *Spectral convergence of probability densities for forward problems in uncertainty quantification*, Numer. Math., 150 (2022), pp. 1165–1186, <https://doi.org/10.1007/s00211-022-01281-4>.
- [San15] F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, Progr. Nonlinear Differential Equations Appl. 55, Birkhäuser, 2015.
- [SDF17] A. SAGIV, A. DITKOWSKI, AND G. FIBICH, *Loss of phase and universality of stochastic interactions between laser beams*, Opt. Express, 25 (2017), pp. 24387–24399, <https://doi.org/10.1364/OE.25.024387>.
- [SDK00] B. SUDRET AND A. DER KIUREGHIAN, *Stochastic Finite Element Methods and Reliability: A State-of-the-Art Report*, University of California, 2000.
- [Spo91] H. SPOHN, *Large Scale Dynamics of Interacting Particles*, Springer, 1991.
- [TG15] D. M. TARTAKOVSKY AND P. A. GREMAUD, *Method of distributions for uncertainty quantification*, in Handbook of Uncertainty Quantification, Springer, 2015, pp. 1–22.
- [TLMNE10] J. TRYOEN, O. LE MAITRE, M. NDJINGA, AND A. ERN, *Intrusive Galerkin methods with upwind-ing for uncertain nonlinear hyperbolic systems*, J. Comput. Phys., 229 (2010), pp. 6485–6511, <https://doi.org/10.1016/j.jcp.2010.05.007>.
- [Tsy09] A. B. TSYBAKOV, *Nonparametric Estimators*, Springer, 2009.
- [Wan91] S. WANG, *Nonparametric estimation of distribution functions*, Metrika, 38 (1991), pp. 259–267, <https://doi.org/10.1007/BF02613619>.
- [Was06] L. WASSERMAN, *All of Nonparametric Statistics*, Springer Science & Business Media, 2006.
- [Xiu10] D. XIU, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.

- [YBTT20] H. J. YANG, F. BOSO, H. A. TCHELEPI, AND D. M. TARTAKOVSKY, *Method of distributions for quantification of geologic uncertainty in flow simulations*, Water Resour. Res., 56 (2020), e2020WR027643, <https://doi.org/10.1029/2020WR027643>.
- [YTT22] H. J. YANG, H. A. TCHELEPI, AND D. M. TARTAKOVSKY, *Method of distributions for two-phase flow in heterogeneous porous media*, Water Resour. Res., 58 (2022), e2022WR032607, <https://doi.org/10.1029/2022WR032607>.