Prediction of Shock Formation From Boundary Measurements

Helmuth Naumer
Department of Electrical and Computer Engineering
and The Coordinated Science Laboratory
University of Illinois Urbana–Champaign
Urbana, IL, USA
hnaumer2@illinois.edu

Farzad Kamalabadi
Department of Electrical and Computer Engineering
and The Coordinated Science Laboratory
University of Illinois Urbana–Champaign
Urbana, IL, USA
farzadk@illinois.edu

2023 IEEE Statistical Signal Processing Workshop (SSP) | 978-1-6654-5245-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/SSP53291.2023.10207929

***Abstract*—We propose a method for the Bayesian prediction of shocks in scalar partial differential equations (PDEs) representing conservation equations from noisy observations of the boundary conditions. By considering the implicit transformation from boundary conditions to shocks, we construct an arrival process interpretation of shocks as well as an associated arrival rate function. We then introduce a Monte Carlo method to approximate the arrival rate of shocks based on the probability of a sufficiently large range of values in an epsilon ball conditioned on noisy boundary measurements. We illustrate the method with simulations of Burgers' equation with initial conditions set by Brownian motion. Despite the non-smooth boundary, our proposed method constructs a sparse and readily interpretable probabilistic structure of shock arrival and propagation.**

***Index Terms*—Non-linear PDEs, Forecasting, Arrival Processes, Inverse Problems**

I. Introduction

Discontinuities in the solutions of partial differential equations (PDEs), also known as shocks, can represent sudden changes in behavior and often have significant implications in their interactions with their surroundings. Shocks occur most commonly in systems of conservation equations, and in many such systems can be identified through the method of characteristics when the boundary conditions are known. Despite a long history of both research and applications, systems of conservation equations are still not fully understood mathematically [1]. Furthermore, due to their extremely localized structure, shocks are notoriously difficult to capture through numerical solvers. Perhaps due to these limitations, the literature is exceedingly sparse in the statistical estimation of shocks beyond abstract work regarding stationary distributions in specific PDEs [2]–[5].

Theoretical work on shock formation in systems of conservations equations has a long history. Shocks form in PDEs when local solutions are not unique, and various entropy conditions are required to determine the physical solution. Many tools have been developed to analyze conservation equations in 1D, resulting in an understanding of sensitivity to source terms [6]. Burgers' equation, a minimal example of a system which forms shocks, has previously been shown to be "supersensitive" to boundary conditions, where changes to the positions of the shocks are $O(1)$ as the size of the perturbation goes to zero [7]. While one dimensional systems tend to be well-understood, higher dimensional systems are much less well developed and the conditions for shock formation are much more complicated [8], [9].

Statistical estimation of PDE solutions is a computationally expensive problem due to the infinite-dimensionality of the space of solutions. The high resolutions required for grid based methods with shocks make the estimation of the density of PDE solutions particularly challenging. One of the most significant developments in recent years around solving non-linear PDEs based on noisy measurements emerged from the area of physics informed neural networks (PINNs) [10]. While the idea of using a neural network to solve a differential equation dates back to at least 1998 [11], there has been a significant resurgence of interest in the techniques [12]. PINNs are a particularly good candidate for identifying shocks due to the lack of the requirement of an explicit discretization of the space. Despite their advantages, they typically lack formal probabilistic interpretations of their results.

An alternative to the limitations of fixed discretizations is a dynamic multigrid method. These methods can work well for shocks, as they adaptively increase the resolution in the neighborhood of the shock itself. Such work is still actively in development, even for one dimensional problems [13]. While originally used for solving known PDEs, they have found applications in tomography [14], [15], sensor network interpolation [16], and general non-linear inverse problems [17].

In this work, we propose a method for the Bayesian prediction of shock formation from noisy measurements of boundary conditions. The method interprets shocks in PDEs as an arrival process and utilizes a Monte Carlo approximation scheme in conjunction with filtering algorithms to estimate an associated arrival rate function. We demonstrate the technique numerically on Burgers' equation, a canonical example of a PDE which forms shocks. Our simulations show that even with distributions of non-smooth boundary conditions, the proposed arrival rate may remain sparse and readily interpretable.

This work was supported in part by the National Science Foundation under Grant 1936663.

II. Problem Formulation

The goal in this work is to characterize the induced stochastic process of shock formation in deterministic PDEs with random boundary conditions. Among PDEs, shocks occur in conservation equations which are commonly written in a form relating the flux to the partial derivative of the solution with respect to time. In this work, we consider scalar-valued conservation equations of the form

$$\nabla_{\mathbf{x}} \cdot (\mathbf{F}(u)) = 0 \tag{1}$$

where $(\nabla_{\mathbf{x}} \cdot)$ represents the spatiotemporal divergence, $\mathbf{F} : \mathbb{R} \to \mathbb{R}^M$ represents the flux, $u(\mathbf{x}) \in \mathbb{R}$ represents the solution of the PDE at a position $\mathbf{x} \in \Omega$. The domain $\Omega \subseteq \mathbb{R}^M$ generally represents space and time in physical PDEs, but can also represent other quantities such as the value of stochastic process if the PDE evolves a density function over time.

We assume some known distribution of boundary conditions μ on some non-characteristic boundary $\Gamma \subset \Omega$. In practice, this known distribution is often the result of conditioning the prior distribution on a collection of noisy measurements $\{Y_i\}$ along the boundary observed at known points $\{\mathbf{x}_i \in \Gamma\}$. Such a non-characteristic boundary may be defined by sensors collecting measurements over time at fixed locations or by a snapshot in time of some physical surface or volume.

We further assume that Γ and $\mathbf{F}$ are such that there exists a unique physical solution satisfying the boundary conditions. It is worth noting that while such a condition is not trivial and depends on the usage of entropy conditions, without such an assumption even exact knowledge of the boundary would be insufficient to reconstruct the solution. Finally, we consider the distribution of solutions u to be that induced by the distribution of boundary conditions μ.

The goal of shock detection is to interpret the probability distribution of discontinuities above a threshold γ in u dependent on the position $\mathbf{x}$. Due to common computational issues with the exact localization of shockwaves, we instead approximate this objective by considering an open ϵ-ball centered at $\mathbf{x}$ to be $B_\epsilon(\mathbf{x}) = \{\mathbf{x}' : \|\mathbf{x} - \mathbf{x}'\| < \epsilon\}$ and define the magnitude of the shock to be $\lim_{\epsilon \to 0} d^{(\mathbf{x},\epsilon)}(u)$ where

$$d^{(\mathbf{x},\epsilon)}(u) = \left[\max_{\mathbf{x}' \in B_\epsilon(\mathbf{x})} u(\mathbf{x}')\right] - \left[\min_{\mathbf{x}' \in B_\epsilon(\mathbf{x})} u(\mathbf{x}')\right]. \tag{2}$$

For statistical tractability, we consider estimating

$$p(\mathbf{x}, \gamma, \epsilon) = P\left(d^{(\mathbf{x},\epsilon)}(u) > \gamma\right) \tag{3}$$

for some chosen $\gamma > 0$ and $\epsilon > 0$, where P denotes the probability of the event. In the next section, we will consider the behavior of this quantity as $\epsilon \to 0$. As the induced distribution of shocks behaves similarly to an arrival process in many cases, we introduce a normalization term and an interpretation as a rate function.

III. Shock Arrival Rate Function

In this section, we construct a rate function $R^*(\mathbf{x}, \gamma)$ for the arrival of shocks in the solutions of the PDE similar to that of a point process. While this function lacks many of the properties associated with common point processes, it is immediately interpretable as a quantification of how likely shocks are to exist in different regions of Ω.

We begin by introducing a normalized density function, motivated by the well-established notion of shock fronts, or twice continuously differentiable (C^2), $(M-1)$-dimensional discontinuity surfaces in the M-dimension spatiotemporal region [9], [18]. By the C^2 structure, we can locally approximate shock fronts as flat and consider the intersection between an $(M-1)$-dimensional hyperplane with $B_\epsilon(\mathbf{x})$. As only perturbations normal to the hyperplane impact the existence of an intersection, we normalize by the diameter of the ϵ-ball, rather than the volume. Thus, we introduce

$$R(\mathbf{x}, \gamma, \epsilon) = \frac{p(\mathbf{x}, \gamma, \epsilon)}{2\epsilon} \tag{4}$$

to be our normalized density function.

This normalized density contains some undesirable features: the placement of shocks is highly correlated and $p(\mathbf{x}, \gamma, \epsilon)$ additionally counts regions with steep gradients rather than only discontinuities. These issues clearly vanish in the limit, constructing our shock rate function

$$R^*(\mathbf{x}, \gamma) = \lim_{\epsilon \to 0} R(\mathbf{x}, \gamma, \epsilon), \tag{5}$$

which can immediately be seen to only include shocks, since $p(\mathbf{x}, \gamma, \epsilon)$ requires gradients of magnitude at least $\frac{\gamma}{2\epsilon}$.

There is still an essential question regarding this rate function before we proceed: Does the limit in Equation (5) exist, and is it non-zero? We introduce the following conjecture based on our construction of Equation (4) and the claim will be further justified through simulation in Section V.

Conjecture 1. *For each $\mathbf{x}$ and γ, if $\lim_{\epsilon \to 0} p(\mathbf{x}, \gamma, \epsilon) = 0$, then for sufficiently small ϵ*

$$p(\mathbf{x}, \gamma, \epsilon) \approx 2\epsilon R^*(\mathbf{x}, \gamma). \tag{6}$$

In what situations can either Conjecture 1 or its assumption be violated?

First, if there exists some point at which shocks form with non-zero probability resulting from some invariant in the boundary distribution, then the function diverges to infinity. This behavior is expected, as it corresponds to Dirac delta functions representing discrete events in a continuous space.

Second, the convergence of $p(\mathbf{x}, \gamma, \epsilon)$ may not be linear, causing either $R^*(\mathbf{x}, \gamma) = 0$ or $R(\mathbf{x}, \gamma, \epsilon) \to \infty$. As $\epsilon \to 0$, our motivating hyperplane approximation from Equation (4) dominates, resulting in the dimensionality reduction and associated claim.

Finally, oscillations in $p(\mathbf{x}, \gamma, \epsilon)$ which are proportional to the volume of $B_\epsilon(\mathbf{x})$ may cause $R^*(\mathbf{x}, \gamma)$ to not exist. The frequency of these oscillations must go to infinity as $\epsilon \to 0$, similar to a Cantor function. As the solutions on either side of a shock front are well-behaved, continuously differentiable solutions of the PDE [18], such pathological behavior seems unlikely to occur in most real-world systems.

There are some nuances to the shock rate function when compared to standard arrival processes. First, shocks are not generally isolated, rather forming curves in Ω on which the magnitude of the discontinuity may vary. Thus, the existence of a large discontinuity at one point all but guarantees that there will be a large discontinuity at an adjacent point. This observation motivates the second, more fundamental, observation: By the construction of the process, we lack any desirable independence structure commonly used to analyze arrival processes. In particular, one can construct bounded regions of Ω for which there is a guarantee of a shock wave for some particular boundary distribution μ.

Thus, while we we will refer to Eq. (5) as a rate function, it is important to remember that much of the classical intuition will not carry over. Despite this, the rate function still provides a key tool for characterizing the distribution of shocks and making informed decisions.

IV. Computational Methods

At this point, we introduce computational techniques to approximate $R^*(\mathbf{x}, \gamma)$. In particular, we provide a brief description of the convergence of a Monte Carlo based estimator

$$\hat{R}(\mathbf{x}, \gamma, \epsilon) = \frac{1}{2\epsilon N} \sum_{i=1}^{N} \begin{cases} 1 & d^{(\mathbf{x},\epsilon)}(u_i) > \gamma \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where each $u_i \sim \mu$ is independently drawn from the boundary distribution.

Proposition 1. *Given N samples, for any normalized density function $R(\mathbf{x}, \epsilon)$, the MSE is*

$$\begin{aligned} &\mathbb{E}\left[\left(\hat{R}(\mathbf{x}, \gamma, \epsilon) - R(\mathbf{x}, \gamma, \epsilon)\right)^2\right] \\ &\qquad = \left(\frac{1}{4\epsilon^2 N}\right) p(\mathbf{x}, \gamma, \epsilon)(1 - p(\mathbf{x}, \gamma, \epsilon)) \end{aligned} \tag{8}$$

Proof. This follows immediately by observing that $2\epsilon\hat{R}(\mathbf{x}, \gamma, \epsilon)$ is an estimator of the parameter of a Bernoulli distribution with parameter $p(\mathbf{x}, \gamma, \epsilon)$. □

While Proposition 1 represents the convergence rate, it is further useful to know how many additional samples are required to maintain the same error for different values of ϵ.

Under Conjecture 1, we can consider the convergence under a sampling scheme in which the number of samples is a function of the radius ϵ, i.e. $N(\epsilon)$. The following proposition describes the convergence as a function of ϵ.

Proposition 2. *Under Conjecture 1, let $N(\epsilon)$ be the number of Monte Carlo samples as a function of the radius of the integration region. Then*

$$\mathbb{E}\left[\left(\hat{R}(\mathbf{x}, \gamma, \epsilon) - R(\mathbf{x}, \gamma, \epsilon)\right)^2\right] = O\left(\epsilon^{-1} N(\epsilon)^{-1}\right) \tag{9}$$

as $\epsilon \to 0$.

Proof. Substitute Equation (6) into Equation (8) as

$$\frac{1}{N(\epsilon)} \cdot \frac{R^*(\mathbf{x}, \gamma)(1 - 2\epsilon R^*(\mathbf{x}, \gamma))}{2\epsilon} \tag{10}$$

Note that $2\epsilon R^*(\mathbf{x}, \gamma) \ll 1$. □

Proposition 2 suggests that under Conjecture 1, such a sequence of estimators converges to R^* if $N(\epsilon)$ grows faster than ϵ^{-1} as $\epsilon \to 0$. Under these conditions, we observe convergence in mean square to $R^*(\mathbf{x}, \gamma)$. This characterization of sample count requirements for convergence based on the radius of the approximation region additionally elucidates the interdependence of the numerical and statistical parameters.

V. Numerical Experiments

In this section, we provide numerical experiments[1] representing an application of the techniques proposed in this work. In particular, we forecast the formation of shock waves based on initial condition measurements of flow rate under a simple fluid dynamics model known as Burgers' equation. The simulations demonstrate that the arrival rate function provides an interpretable forecast of shock formation and propagation over time.

The advection form of Burgers' equation is

$$u_t + uu_x = 0, \tag{11}$$

where u represents the flow rate of the fluid and subscripts are used to denote partial derivatives with respect to time and space.

The importance of Burgers' equation is much greater than that of a simplified fluid dynamics model. In particular, it represents a canonical example of a hyperbolic PDE which forms shock waves through wave steepening and is an exceedingly common representative example in research discussing shock waves [2]–[5], [19]. From Equation (11), we can see that shock waves form due to features flowing spatially at a rate proportional to the value: higher flow rate regions will travel faster and collide with lower flow rate regions.

In our simulations, we use a Brownian motion prior on the boundary conditions, a common prior when little is known about a generative process. While such a boundary condition initializes the system such that discontinuities form immediately, we will see that the PDE quickly filters the shocks into a few dominant regions.

The initial conditions are defined by the state-space model

$$\begin{aligned} u(x_i + \delta, 0) &= u(x_i, 0) + w_i \\ Y_i &= u(x_i, 0) + v_i \end{aligned} \tag{12}$$

where $w_i \sim \mathcal{N}(0, \sigma_w^2)$ and $v_i \sim \mathcal{N}(0, \sigma_v^2)$ represent additive i.i.d. Gaussian noise and $x_{i+1} = x_i + \delta$ for some fixed $\delta > 0$. We apply backward Markovian state sampling based on the Rauch-Tung-Striebel (RTS) smoothing algorithm to construct sample trajectories conditioned on the observation [20].

The solutions of Burgers equation were computed using a Godunov discretization scheme to limit advection-based errors in shock locations [21]. The spatial grid resolution was chosen as 0.05 to ensure good localization of the discontinuities and the temporal grid resolution was chosen as 0.005 to ensure

[1] Available at https://github.com/Helmuthn/BayesianShockEstimation.jl

stability of the numerical integrator. $\gamma = 1$ was empirically chosen to capture many of the discontinuities while rejecting regions with significant gradients.

The results of the simulations appear in Figure 1. Panel (a) shows the true boundary conditions in orange as well as the observations in blue. Panel (c) shows 20 example boundary trajectories sampled according to the RTS smoothing distribution. The two key features in this plot are that the sample trajectories are still non-smooth, resulting in rapid formation of shocks, and they still cover a large area which has the potential to significantly impact the location of a shock. Panel (b) shows the numerically computed solution for the boundary conditions, and Panel (d) shows the numerically computed shock arrival rates conditioned on the boundary observations. Despite the rough boundary conditions and significant variation in snapshot samples, the arrival of shocks is fairly stable.

In Panels (a), (b), and (d) of Figure 2, we show the results of the estimator with three different values of ϵ. The particular values represent 1, 2, and 3 spatial steps. Step sizes below the spatial grid resolution have little additional information about the shock positions. This figure supports Conjecture 1 by demonstrating that the amplitude of $p(\mathbf{x}, \gamma, \epsilon)$ scales linearly with ϵ. In Panel (c), we plot one future time slice in the region in which there is empirically exactly one discontinuity and illustrated the approximate arrival rate as a function of space and ϵ. These curves numerically integrate to approximately 1, illustrating that the arrival rate may become a valid density function in sufficiently constrained regions in which there is a guarantee of exactly one discontinuity existing.

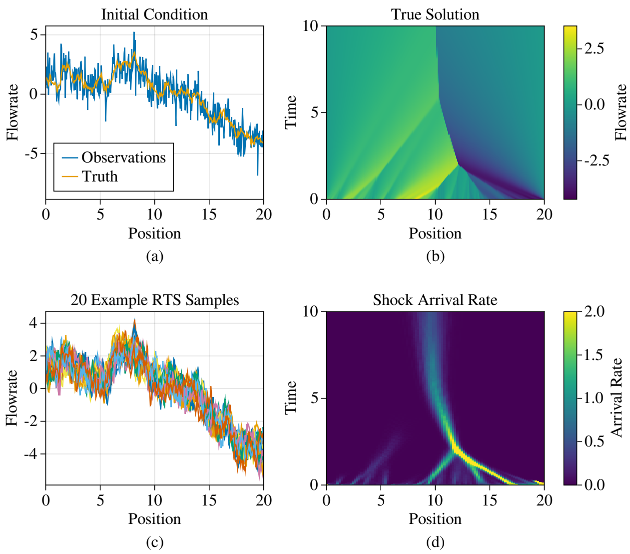

Fig. 1. Example approximation of arrival rate based on observations. (a) Noisy observations of the boundary in blue and true boundary conditions in yellow; (b) Solution with true boundary conditions; (c) 20 Rauch-Tung-Striebel samples; (d) Conditional shock arrival rate given observations.

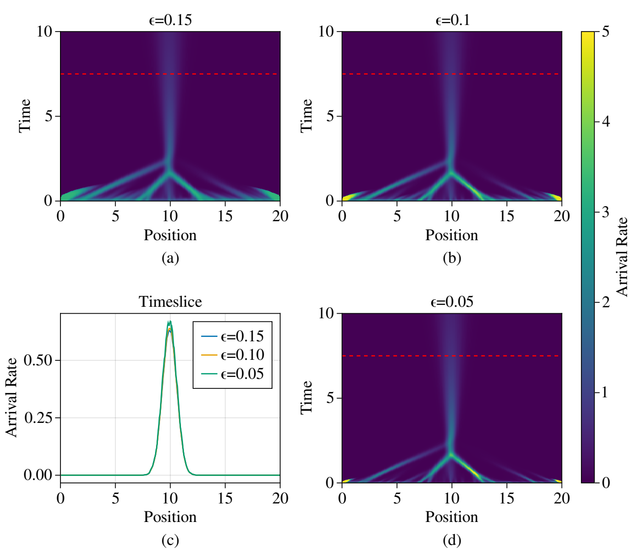

Fig. 2. Convergence behavior of the estimator in Equation (7). Panels (a), (b), and (d) show the approximation of arrival rate for different values of ϵ. Panel (c) shows a snapshot in time at 7.5 seconds, demonstrating the consistent amplitude.

VI. Conclusion

We proposed an interpretable method for the Bayesian prediction of shock formation and propagation in PDEs from noisy observations of boundary conditions. Motivated by an analogy to point processes, we constructed a Monte Carlo approximation scheme to compute an arrival rate function for the conditional distribution of shocks. We further provided sample count requirements for convergence based on the radius of the approximation region, characterizing the interdependence of numerical and statistical parameters. Finally, we tested these techniques on Burgers' equation, forecasting shocks based on a noisy snapshot of initial conditions and demonstrating the sparse structure of the rate function.

References

[1] Lawrence C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, RD, second edition, 2010.

[2] Zhen-Su She, Erik Aurell, and Uriel Frisch, "The inviscid Burgers equation with initial data of Brownian type," *Communications in Mathematical Physics*, vol. 148, no. 3, pp. 623–641, Sep 1992.

[3] Ya. G. Sinai, "Statistics of shocks in solutions of inviscid Burgers equation," *Communications in Mathematical Physics*, vol. 148, no. 3, pp. 601–621, Sep 1992.

[4] Yuri Bakhtin, "Burgers equation with random boundary conditions," *Proceedings of the American Mathematical Society*, vol. 135, no. 7, pp. 2257–2263, 2007.

[5] Yuri Bakhtin, Eric Cator, and Konstantin Khanin, "Space-time stationary solutions for the Burgers equation," *Proceedings of the American Mathematical Society*, vol. 27, no. 1, pp. 193–238, 2014.

[6] Stefan Ulbrich, "A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms," *SIAM Journal on Control and Optimization*, vol. 41, no. 3, pp. 740–797, 2002.

[7] Dongbin Xiu and George Em Karniadakis, "Supersensitivity due to uncertain boundary conditions," *International Journal for Numerical Methods in Engineering*, vol. 61, no. 12, pp. 2114–2138, 2004.

[8] S N Kružkov, "First order quasilinear equations in several independent variables," *Mathematics of the USSR-Sbornik*, vol. 10, no. 2, pp. 217, feb 1970.

[9] Guy Métivier, "Stability of multidimensional shocks," in *Advances in the Theory of Shock Waves*, pp. 25–103. Birkhäuser Boston, Boston, MA, 2001.

[10] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli, "Scientific machine learning through physics–informed neural networks: Where we are and what's next," *Journal of Scientific Computing*, vol. 92, no. 3, pp. 88, Jul 2022.

[11] I.E. Lagaris, A. Likas, and D.I. Fotiadis, "Artificial neural networks for solving ordinary and partial differential equations," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 987–1000, 1998.

[12] M. Raissi, P. Perdikaris, and G.E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[13] Alexander J. Howse, Hans De Sterck, Robert D. Falgout, Scott MacLachlan, and Jacob Schroder, "Parallel-in-time multigrid with adaptive spatial coarsening for the linear advection and inviscid Burgers equations," *SIAM Journal on Scientific Computing*, vol. 41, no. 1, pp. A538–A565, 2019.

[14] Jong Chul Ye, C.A. Bouman, K.J. Webb, and R.P. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 909–922, 2001.

[15] Liisa-Ida Sorsa, Mika Takala, Christelle Eyraud, and Sampsa Pursiainen, "A time-domain multigrid solver with higher-order born approximation for full-wave radar tomography of a complex-shaped target," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 579–590, 2020.

[16] I. Spiliopoulos, D.T. Hristopulos, M.P. Petrakis, and A. Chorti, "A multigrid method for the estimation of geometric anisotropy in environmental data from sensor networks," *Computers & Geosciences*, vol. 37, no. 3, pp. 320–330, 2011, Geoinformatics for Environmental Surveillance.

[17] Seungseok Oh, A.B. Milstein, C.A. Bouman, and K.J. Webb, "A general framework for nonlinear multigrid inversion," *IEEE Transactions on Image Processing*, vol. 14, no. 1, pp. 125–140, 2005.

[18] Andrew Majda, *The existence of multi-dimensional shock fronts*, Memoirs of the American Mathematical Society, v. 281. American Mathematical Society, Providence, R.I, 1983.

[19] Maziar Raissi and George Em Karniadakis, "Hidden physics models: Machine learning of nonlinear partial differential equations," *Journal of Computational Physics*, vol. 357, pp. 125–141, 2018.

[20] Olivier Cappé, Eric Moulines, and Tobias Rydén, *Inference in Hidden Markov Models*, chapter 6, p. 164, Sprinter Science and Business Media Inc., New York, NY, 2005.

[21] Randall J. LeVeque, *Numerical Methods for Conservation Laws*, pp. 136–145, Birkhäuser Basel, Basel, 1992.