



Research article

A minimization approach to conservation laws with random initial conditions and non-smooth, non-strictly convex flux

Carey Caginalp*

University of Pittsburgh, Mathematics Department, 301 Thackeray Hall, Pittsburgh PA 15260

* **Correspondence:** Email: carey_caginalp@alumni.brown.edu.

Abstract: We obtain solutions to conservation laws under any random initial conditions that are described by Gaussian stochastic processes (in some cases discretized). We analyze the generalization of Burgers' equation for a smooth flux function $H(p) = |p|^j$ for $j \geq 2$ under random initial data. We then consider a piecewise linear, non-smooth and non-convex flux function paired with general discretized Gaussian stochastic process initial data. By partitioning the real line into a finite number of points, we obtain an exact expression for the solution of this problem. From this we can also find exact and approximate formulae for the density of shocks in the solution profile at a given time t and spatial coordinate x . We discuss the simplification of these results in specific cases, including Brownian motion and Brownian bridge, for which the inverse covariance matrix and corresponding eigenvalue spectrum have some special properties. We calculate the transition probabilities between various cases and examine the variance of the solution $w(x, t)$ in both x and t . We also describe how results may be obtained for a non-discretized version of a Gaussian stochastic process by taking the continuum limit as the partition becomes more fine.

Keywords: conservation laws; random initial conditions; Lax-Oleinik; shocks; variational problems in differential equations; minimization; stochastic processes

Mathematics Subject Classification: 35F20, 35F31

1. Introduction

The conservation law in the form

$$\begin{aligned} w_t + (\mathcal{H}(w))_x &= 0 \text{ in } \mathbb{R} \times (0, \infty) \\ w(x, 0) &= g'(x) \text{ on } \mathbb{R} \times \{t = 0\} \end{aligned} \tag{1.1}$$

and its related Hamilton-Jacobi problem

$$u_t + \mathcal{H}(u_x) = 0 \text{ in } \mathbb{R} \times (0, \infty)$$

$$u(x, 0) = g(x) \text{ on } \mathbb{R} \times \{t = 0\} \quad (1.2)$$

have a wide array of applications among fluid mechanics, shocks, and turbulence. An interesting feature is that even for a smooth flux function \mathcal{H} and smooth initial data, discontinuous solutions to the conservation law (1.1) may occur [4, 6, 10, 11, 12, 14, 15, 17, 18, 19, 21, 22, 23, 24, 25, 31, 32]. In particular, this is observed in the prototypical case of Burgers' equation, for which $\mathcal{H}(w) = w^2/2$. The work of Dafermos [8] was instrumental in constructing an approximation for this case and more generally, locally Lipschitz \mathcal{H} by using a series of piecewise linear flux functions to construct the solution for deterministic initial conditions. Using our methods for random initial conditions, one also can construct the continuum limit for the deterministic problem as a special case.

In this paper, we consider the introduction of randomness into the initial conditions for the conservation law above. The classical approach for solving the conservation law (1.1) entails using variational methods to derive the Lax-Oleinik formula [11, 17, 21, 22]

$$w(x, t) := \frac{\partial}{\partial x} \left[\min_{y \in \mathbb{R}} \left\{ t \mathcal{L} \left(\frac{x-y}{t} \right) + g(y) \right\} \right], \quad (1.3)$$

requiring assumptions on smoothness, superlinearity, and strict convexity of the flux function. In a previous paper [5], we showed that (1.3) will still hold when these conditions are relaxed, with the only other assumption being that of Lipschitz continuity on the initial condition $g'(x)$. These methods thus enable us to derive a number of results for the case of non-deterministic initial data. When dealing with randomness, an important notion is that of a stochastic process $X(t)$. The process is a set of random variables linked together by a law of probability with the set of possible paths given by some state space Ω . In general this is far too broad a class to attack analytically, so attempts to derive exact solutions are often restricted to one specific stochastic process, for example Brownian motion [12]. In this work, we derive results for an entire family of stochastic processes. In Section 2 we provide an introduction and references to some of the key properties of such a family known as *Gaussian stochastic processes*. The motivation for this class of process is not only that it has sufficient structural properties to obtain closed-form results but also has the breadth to analyze a number of distinct cases. We also utilize this approach in calculating transition probabilities between different states and examining the variance in space as time increases.

Consider the basic problem $w_t + |w|_x = 0$ subject to continuous initial conditions $w(x, 0) = g'(x)$. Application of our results gives $g(x-t)$ as the solution if the minimum of $g(y)$ is at the left endpoint of the interval, $g(x+t)$ if the minimum is on the right endpoint, and 0 for an interior minimum. Thus for g' that is given by Brownian motion, one minimizes integrated Brownian motion over a finite interval and obtains the shock statistics.

This example can provide a pathway to considering the behavior of solutions to more complicated flux functions under random initial data, for example when one has not one but a number of affine segments linked together. This sort of construction is denoted as a *polygonal flux function* and is thus particularly suited to analysis under the Hopf-Lax minimization approach [17, 21, 22] discussed below. Together with the results proven by Dafermos [8], this methodology can be the basis for a general class of flux and random initial data.

Notably, both Brownian motion and Brownian bridge initial conditions for $g'(x)$ will satisfy the requisite regularity properties, as well the additional property of being Gaussian stochastic processes. Furthermore, when integrated to obtain integrated Brownian motion and integrated Brownian bridge,

they remain Gaussian stochastic process, leading to useful properties on both g and g' that aid in the analysis and computation of minimizers as specified in (1.3). As is shown in [5], we have the key relationship for the solution w given by

$$w(x, t) = g'(y^*(x, t)) \quad (1.4)$$

(when $g'(y^*(x, t))$ exists) expressed as a function of the greatest minimizer,

$$y^* = \arg^+ \min_{y \in \mathbb{R}} \left\{ t \mathcal{L} \left(\frac{x-y}{t} \right) + g(y) \right\}. \quad (1.5)$$

For the case of Burgers' equation, one uses the relation $g'(y^*(x, t)) = \frac{x-y^*(x, t)}{t}$ to obtain a special case of (1.4) given by

$$w(x, t) = \frac{x - y^*(x, t)}{t} \quad (1.6)$$

When one introduces Brownian motion initial data, (1.4) leads to the following result for the variance of the solution expressed in terms of the variance of the minimizer in (1.5), for a flux function given by $\mathcal{H}(p) = \frac{1}{j} |p|^j$

$$\text{Var}(w(x, t)) = t^{-\frac{2}{j-1}} \text{Var} \left(\text{sgn}(x - y^*(x, t)) |x - y^*(x, t)|^{\frac{1}{j-1}} \right) \quad (1.7)$$

which for Burgers' equation reduces to

$$\text{Var}(w(x, t)) = \frac{1}{t^2} \text{Var}(x - y^*(x, t)). \quad (1.8)$$

Later on, we specialize to the case $x = 0$.

In Section 3 we discuss extensions of (1.8) to more general flux functions. For example, with the flux $\mathcal{H}(p) = p^{2n}$, we have a generalization of (1.8), providing an understanding of the propagation of variance over the sample space Ω . We also consider a related case of $\mathcal{H}(p) = |p|$, in light of its Legendre transform

$$\mathcal{L}(q) = \begin{cases} 0 & |q| \leq 1 \\ +\infty & |q| > 1 \end{cases} \quad (1.9)$$

for which the solution has a very simple probabilistic interpretation. In particular, one can show formally that the frequency of jumps in $w(x, t)$ decreases in time.

The minimum defined on the right-hand side of (1.5) is of course unique under the classical assumptions that \mathcal{H} is smooth, superlinear, and uniformly convex [11]. However, the analog of the result (1.4) remarkably continues to hold even when these assumptions are relaxed, i.e. when the flux function has a lesser degree of regularity, the convexity of \mathcal{H} is not uniform or even strict, and when it is no longer superlinear. Even if the initial condition g' is also discontinuous, one can prove a modified form of (1.4) by accounting for the vertices of g .

A particularly interesting and yet broad range of cases results from studying a Gaussian stochastic process, $X(r)$, a random process completely determined by its second order statistics. In other words, such a process at arbitrary times $\{r_i\}_{i=1}^n$ is completely determined by the means $\mathbb{E}\{X(r_i)\}$ and the corresponding covariance matrix Σ defined by $(\Sigma)_{ij} = \text{Cov}(X(r_i), X(r_j))$. We denote the inverse covariance matrix by $A := \Sigma^{-1}$.

In Section 4, we consider a Gaussian stochastic process $X(r)$ and partition the real line into sets ζ_i with measure $\nu(\zeta_i) > \varepsilon > 0$, and take a finite number of points, among which we compute the minimum. We pair this initial condition with a piecewise linear flux function. The Legendre transform L of this flux function only takes a finite number of slopes, each on an interval of finite measure. By taking the intersection of our partition with this interval, we obtain a finite covering for it, and thus a finite number of points at which to analyze the quantity minimized in (1.5). We want to calculate the solution $\mathbb{E}\{w(x, t)\}$, which we have shown in [5] to depend only on the derivative of L or g at the point of the minimizer. By considering first the local minima of this quantity for each of the finite number of segments of L , and then taking the minimum over those segments, we obtain a remarkable closed-form expression for the solution. Among the results we can obtain from this is the expectation of the solution at a point (x, t) .

$$\mathbb{E}\{w(x, t)\} = - \sum_{i=1}^N p_i c_{N+1-i} + \sum_{i=N+1}^{2N+1} p_i \mathbb{E}\{g'(y^*(x, t))\} \quad (1.10)$$

where

$$p_i = \sum_{j=2}^{n_i} (2\pi)^{-\frac{n_i}{2}} |A|^{-\frac{1}{2}} \int_{-\infty}^{\infty} dx_{i,j} \prod_{(m,l) \neq (i,j)} \int_{x_{i,j}}^{\infty} dx_{m,l} e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T A(\bar{x}-\bar{\mu})}$$

for $1 \leq i \leq N$.

For $N+1 \leq i \leq 2N+1$ one has $p_i = \mathbb{P}\{R_{i-N}\}$.

(1.11)

are expressed in terms of integrals of nested error functions, and the sets $\{R_i\}$ refer to the event that a minimum occurs at a vertex of L (discussed further in (5.6)). The quantity A_i in the first term of (1.11) here denotes the inverse covariance matrix corresponding to the i th affine segment of L , and A for the same quantity over the whole domain. The vector μ is given by the discrete points of $\mu = \mathbb{E}\left\{g(y) + tL\left(\frac{x-y}{t}\right)\right\}$, where the second term is deterministic. The quantity $\tilde{\mu}$ refers to the analogous quantity in the i th segment. The integers n_i correspond to technical details of partitioning the intervals for which L has different slopes, as described in Section 5.

An important aspect of this evolution is that for Burgers' equation with random initial conditions, one has an increasing variance in the probabilistic sense for the solution $w(x, t)$ as a function of t for a given x , but formal calculations suggest that it has decreasing variance in the sense of $TV(w(x, t))$ as a function of time. In this manner, it is almost paradoxical that conservation laws which lead to shocks from smooth solutions can also act in such a way as to smooth out random initial conditions.

A key issue in kinetic theory is the distribution and formation of discontinuities in the solution, known as *shocks*. In particular, for a Gaussian stochastic process paired with a piecewise linear flux function, we can expand on the results of Section 5, including (1.11). In doing so, we derive expressions (both exact and approximate) for the distribution of these shocks based on the minimizer switching from a point on one segment of L to another. From this result, we also compute the total density of shocks.

The expressions obtained in Sections 5 and 6 hold for *any* Gaussian stochastic process. Some cases of particular interest include that of Brownian motion or Brownian bridge initial conditions, as well as Ornstein-Uhlenbeck ([26], p. 439). In Section 6, we apply our results to special cases and observe the simplifications for these specific processes. Since many of our results are exact and in terms of error

functions, this also provides a pathway to further numerical computation. By taking the limit of a finer partition of the real line, we also explain how one may be able to pass to a limit and obtain results for the continuum problem in addition to the discretized one.

2. Background on Gaussian stochastic processes

2.1. Stochastic processes

In considering randomness for the initial conditions along the entire real line \mathbb{R} , it is important to have a well-defined construction. A *stochastic process* $X(s)$ is a collection of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The space Ω can be thought of as the set of all possible outcomes for the random variable, \mathcal{F} as a set of reasonable sets of outcomes (i.e. *measurable* sets) and \mathbb{P} is a measure with total mass 1 assigning the probabilities for which an outcome or set of outcomes in \mathcal{F} can occur. The expectation $\mathbb{E}\{X\}$ of a random variable X is defined as $\mathbb{E}\{X\} = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)$. We also require that the measure \mathbb{P} be countably additive, i.e. $\mathbb{P}(\cup_{i \in \mathbb{N}} F_i) = \sum_{i=1}^{\infty} \mathbb{P}(F_i)$ for $\{F_i\}$ disjoint.

Example 2.1. Consider a collection of random variables $\{X_{\alpha}\}_{\alpha \in I}$ for some index set I such that $X_{\alpha} \sim \mathcal{N}(0, 1)$. That is, take identically distributed random variables that are normally distributed with mean 0 and variance 1. Define

$$S(n) = \sum_{i=1}^n X_{\alpha}. \quad (2.1)$$

The process $S(n)$ is known as a random walk and is a stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\Omega = \mathbb{N}$, \mathcal{F} is the set of Borel sets on \mathbb{R} , and \mathbb{P} is the usual probability measure.

2.2. Gaussian stochastic processes

Let $\{R_i\}_{i=1}^m$ be a collection of independent, identically distributed random variables with standard normal distribution, i.e. $R_i \sim \mathcal{N}(0, 1)$ as in Example 2.1. If a set of random variables $\{X_i\}_{i=1}^n$ satisfy

$$\begin{aligned} X_1 &= a_{11}R_1 + \dots + a_{1m}R_m + \mu_1 \\ &\dots \\ X_n &= a_{n1}R_1 + \dots + a_{nm}R_m + \mu_n \end{aligned} \quad (2.2)$$

for constants $\{a_{ij}\}$, $\{\mu_i\}$, we say that (X_1, \dots, X_n) form a *multivariate normal distribution*. We also define the *covariance* of two random variables below.

Definition 2.2. Let X_i and X_j be random variables. We define the covariance of the random variables by

$$\begin{aligned} \text{Cov}[X_i, X_j] &= \mathbb{E}[X_i - \mathbb{E}[X_i]] \mathbb{E}[X_j - \mathbb{E}[X_j]] \\ &= \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \mathbb{E}[X_j] \end{aligned} \quad (2.3)$$

This is the basis of the definition for a Gaussian stochastic process. In particular, one has the following [27]:

Theorem 2.3. If (X_1, \dots, X_n) have a multivariate normal distribution with density given by

$$f_X(X_1, \dots, X_n) = \frac{e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1}(\vec{x}-\vec{\mu})}}{\sqrt{(2\pi)^n |\Sigma|}} \quad (2.4)$$

where $\vec{\mu}$ is the mean, and Σ the covariance matrix, then the joint moment generating function is given by

$$\phi(\lambda_1, \dots, \lambda_n) = \exp \left\{ \sum_{i=1}^n \lambda_i \mu_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \text{Cov}[X_i X_j] \right\} \quad (2.5)$$

This is easily proven using the definition and algebraic manipulation and properties of characteristic functions. We can now define what we mean by a Gaussian stochastic process:

Definition 2.4. Let $\{t_i\}_{i=1}^n$ be arbitrary and increasing without loss of generality. We call a stochastic process $\{X(t)\}$ a Gaussian stochastic process if $(X(t_1), \dots, X(t_n))$ has a multivariate normal distribution.

Definition 2.4 implies that a stochastic process that is Gaussian is completely characterized by the properties of its mean and covariance matrix. A natural question to ask is what sort of processes might satisfy this requirement, and how one proves this property. We next give several examples relating to cases of particular interest in the broader literature. More details can be found in [27].

Definition 2.5. Let $W(t)$ be a stochastic process. We call $W(t)$ Brownian motion if

- (i) $W(0) = 0$
- (ii) $W(t)$ is almost surely continuous
- (iii) $W(t)$ has independent increments with a normal distribution, i.e.

$$W(t) - W(s) \sim \mathcal{N}(0, t - s). \quad (2.6)$$

Definition 2.6. We call a Brownian motion conditioned on $W(T) = 0$, denoted $\{W_b(t)\}_{0 \leq t \leq T}$ a Brownian bridge. An equivalent definition is to set

$$W_b(t) = W(t) - \frac{t}{T} W(T). \quad (2.7)$$

In particular, we note that $W_b(0) = W_b(T) = 0$.

Definition 2.7. We call a process $W_o(t)$ stationary Ornstein-Uhlenbeck if it is defined by

$$W_o(t) = e^{-\alpha t} W(e^{2\alpha t}). \quad (2.8)$$

The stationary Ornstein-Uhlenbeck process has the important property that its associated covariance matrix is invariant under translation in time. A basic result is the following.

Lemma 2.8. (a) Let $\{W(t)\}_{t \geq 0}$ be a Brownian motion. Then $W(t)$ is a Gaussian stochastic process.

(b) The integrated Brownian motion $S(t)$ given by $S(t) := \int_0^t W(r) dr$. Then $S(t)$ is a Gaussian stochastic process. Further, defining $Z(t)$ as integrated Brownian motion together with a drift, $Z(t)$ is a Gaussian stochastic process.

(c) Let $T > 0$ be given and $W(t)$ be a Brownian motion on $[0, T]$. The stochastic process $W_b(t) = W(t) - \frac{t}{T} W(T)$, called Brownian bridge, is a Gaussian stochastic process.

(d) Define $Z(t) = \int_0^t W_b(r) dr$. Then $Z(t)$ is also a Gaussian stochastic process.

Part (b) is proven in the Appendix. A basic result that follows from the standard theory is given below.

Theorem 2.9. (Covariance Matrices for Brownian Motion and Brownian Bridge). Let T and $\{t_i\}_{i=1}^N \in [0, T]$ be given.

(a) Let $\{S(t)\}_{t>0}$ be an integrated Brownian motion with piecewise linear drift. Then the covariance at times $0 \leq s \leq t$ is given by

$$\text{Cov}[S(t), S(s)] = s^2 \left(\frac{t}{2} - \frac{s}{6} \right) \quad (2.9)$$

(b) For an integrated Brownian bridge $Z(t) = \int_0^t W(r) - \frac{r}{T} W(T) dr$, the covariance matrix at times $0 \leq s \leq t \leq T$ is given by

$$\text{Cov}[Z(t), Z(s)] = s^2 \left(\frac{t}{2} - \frac{s}{6} \right) - \frac{1}{T} \frac{t^2 s^2}{4} \quad (2.10)$$

For s and t outside of $[0, T]$, the variance of $Z(t)$ is constant but nonzero.

An outline of the proof of Theorem 2.9 is provided in the Appendix.

Remark 2.10. Later on, we consider Brownian motion on the interval $(0, 1)$. In order to use our results on a different interval, one can instead consider Brownian motion on a translated and stretched interval, i.e. $(-M, M)$ where $B(-M) = 0$ is the starting point instead of 0 as is typically used.

3. Generalizations of Burgers' equation with random initial data

We consider a generalization of Burgers' equation by considering the flux function for $j \geq 2$ as follows:

$$\mathcal{H}(p) = \frac{p^j}{j}. \quad (3.1)$$

In this way, we obtain the following generalization of (1.8).

Lemma 3.1. With \mathcal{H} defined by (3.1), consider the conservation law

$$\begin{aligned} w_t + (\mathcal{H}(w))_x &= 0 \text{ in } \mathbb{R} \times (0, \infty) \\ w(x, 0) &= g'(x) \text{ on } \mathbb{R} \times \{t = 0\}. \end{aligned} \quad (3.2)$$

The solution is then

$$w(x, t) = \text{sgn} \left(\frac{x - y^*(x, t)}{t} \right) \left(\frac{x - y^*(x, t)}{t} \right)^{1/(j-1)}. \quad (3.3)$$

Proof. The Legendre transform of this flux function (3.1) is

$$\begin{aligned} \mathcal{L}(q) &= \sup_p \{pq - p^j\} = |q|^{j/(j-1)} - \frac{1}{j} |q|^{j/(j-1)} \\ &= \frac{j-1}{j} |q|^{j/(j-1)} \end{aligned} \quad (3.4)$$

and its derivative is given by

$$\mathcal{L}'(q) = \text{sgn}(q) q^{1/(j-1)}. \quad (3.5)$$

The solution $w(x, t)$ to the conservation law (1.1) can then be expressed as

$$w(x, t) = \mathcal{L}'\left(\frac{x - y^*(x, t)}{t}\right) = \operatorname{sgn}\left(\frac{x - y^*(x, t)}{t}\right) \left|\frac{x - y^*(x, t)}{t}\right|^{1/(j-1)}$$

□

Remark 3.2. For Burgers' equation, $j = 2$ and the relation (3.3) reduces to

$$w(x, t) = \frac{x - y^*(x, t)}{t},$$

i.e., the result (1.6).

Lemma 3.3. Consider the conservation law as in Lemma 3.1 with $\mathcal{H}(p) = |p|$ with initial data g' . Then the solution is given by

$$\begin{aligned} w(x, t) &= g'(y^*(x, t)) \text{ at points where } g' \text{ is continuous,} \\ w(x, t) &= 0 \text{ elsewhere,} \end{aligned} \quad (3.6)$$

where y^* is the value of the minimizer as before.

Proof. The Legendre transform of \mathcal{H} is simply

$$L(q) = \begin{cases} 0 & |q| \leq 1 \\ +\infty & |q| > 1 \end{cases}. \quad (3.7)$$

The result then follows immediately from the solution formula (1.3). □

This result is particularly of interest as it shows, for the special case of the flux function $\mathcal{H}(p) = |p|$, that the solution of the conservation law reduces to finding the minimum of the integral of the initial data, i.e. the minimum of g . For example, if g' is given by a Brownian motion, the solution at given x and t is given by solving the problem of finding the minimum of an integrated Brownian motion on a finite interval (whose endpoints are dependent on x and t). The integrated Brownian motion is a Gaussian stochastic process (Lemma 2.8 (b)), so the minimizer y^* can easily be approximated by using the mean and covariance matrix (Theorem 2.9 (a)).

For example, if the initial condition g' consists of ± 1 with probabilities p and $1 - p$, then one has $w(x, t) \in \{0, g'(x - t), g'(x + t)\}$. The value taken by w depends on whether the minimum is attained on the interior, left endpoint, or right endpoint of the interval $[x - t, x + t]$.

Remark 3.4. Due to the Legendre transform L taking on infinite value outside of $|q| > 1$, for a given (x, t) , the minimizer $y^*(x, t)$ takes its value in the domain

$$|x - y| < t. \quad (3.8)$$

This is a manifestation of the finite domain of influence strictly enforced, i.e. within a cone $y = x \pm t$.

Theorem 3.5. Consider the conservation law (1.1) with flux function given by (3.1), i.e. $\mathcal{H}(p) = \frac{p^j}{j}$, $j \geq 2$. For Brownian motion initial conditions one has, with y^* depending on (x, t) ,

$$\operatorname{Var}(w(x, t)) = t^{-\frac{2}{j-1}} \operatorname{Var}\left(\operatorname{sgn}(x - y^*) |x - y^*|^{\frac{1}{j-1}}\right), \quad (3.9)$$

noting that y^* is short for $y^*(x, t)$.

Proof. For a flux function of the form (3.1), a minimum is attained when

$$\mathcal{L}'\left(\frac{x-y^*}{t}\right) = \operatorname{sgn}\left(\frac{x-y^*}{t}\right)\left(\frac{x-y^*}{t}\right)^{\frac{1}{j-1}} = g'(y^*) = w(x, t). \quad (3.10)$$

One then writes

$$\begin{aligned} \mathbb{E}\{w(x, t)\} &= t^{-\frac{1}{j-1}} \mathbb{E}\left\{\operatorname{sgn}\left(\frac{x-y^*}{t}\right)|x-y^*|^{\frac{1}{j-1}}\right\}, \\ \mathbb{E}\{w(x, t)^2\} &= t^{-\frac{2}{j-1}} \mathbb{E}\{|x-y^*|^{\frac{2}{j-1}}\}. \end{aligned} \quad (3.11)$$

Therefore

$$\operatorname{Var}(w(x, t)) = t^{-\frac{2}{j-1}} \{\mathbb{E}\{|x-y^*|^{\frac{2}{j-1}}\} \quad (3.12)$$

$$- \mathbb{E}\left\{\operatorname{sgn}\left(\frac{x-y^*}{t}\right)|x-y^*|^{\frac{1}{j-1}}\right\}^2 \quad (3.13)$$

□

Remark 3.6. To obtain the result for Burgers' equation, one sets $j = 2$, and observes that Theorem 3.5 then yields the result (1.8):

$$\operatorname{Var}(w(x, t)) = \frac{1}{t^2} \operatorname{Var}(x - y^*(x, t)). \quad (3.14)$$

Further, for the case of Brownian motion and for an arbitrary time t , $\operatorname{Var}(g'(y^*(x, t)))$ is increasing in x since $y^*(x, t)$ is increasing in x (see [5]), and thus the variance of $w(x, t)$ also increases.

Remark 3.7. From (3.9) for Burgers' equation at $x = 0$, using Theorem 3.5 we have

$$\operatorname{Var}(w(0, t)) \sim \frac{C}{t^2} \cdot \operatorname{Var}[y^*(0, t)] \quad (3.15)$$

Using the solution formula, we also have

$$w(x, t) = g'(y^*(x, t)) \text{ so } \operatorname{Var}[w(x, t)] = \operatorname{Var}[g'(y^*(x, t))] \quad (3.16)$$

Combining (3.15) and (3.16) yields

$$t^2 \operatorname{Var}[g'(y^*(0, t))] = \operatorname{Var}[y^*(0, t)]. \quad (3.17)$$

Since g' is Brownian motion, one has $\operatorname{Var}[g'(z)] = z$ for any $z > 0$. Thus, one can formally write

$$\operatorname{Var}[g'(y^*(0, t))] \sim \mathbb{E}y^*(0, t) \quad (3.18)$$

Combining (3.17) and (3.18) then yields

$$\mathbb{E}y^*(0, t) \sim \frac{C}{t^2} \operatorname{Var}[y^*(0, t)] \quad (3.19)$$

Further, if we are willing to make the conjecture that

$$\text{Var} [y^* (0, t)] \sim \mathbb{E} [y^* (0, t)]^p, \quad (3.20)$$

for some power $p \in \mathbb{N}$, then we can deduce

$$\text{Var} (y^*) = t^2 \text{Var} (g' (y^*)) \sim t^2 \mathbb{E} [y^*]^p \sim Ct^2 \cdot t^p. \quad (3.21)$$

Then the result (3.15) can be further extended to obtain

$$\text{Var} (w (0, t)) \sim Ct^{3p-2}, \quad (3.22)$$

which seems to formally suggest that the variance of the solution on $x = 0$ increases as time increases. One can also perform computations using a finite difference scheme through a numerical study, e.g. from [30].

3.1. Spatial density of shocks

We now present a formal argument that the density of shocks in the solution $w(x, t)$ in x decreases as t increases for the case $H(p) = |p|$ with any random initial data. As noted in 3.4, the range of minimization problem is restricted to the interval $|x - y| \leq t$, so that we have

$$w(x, t) = \partial_x \min_{y \in [x-t, x+t]} g(y) = g'(y^*(x, t)). \quad (3.23)$$

The minimum of $g(y)$ is either on the interior or attained at one of the endpoints, $x \pm t$. If it occurs on the interior, we must have $g'(y^*) = 0$. However, in the event of a minimum at $x + t$, one may have $g'(y^*) = g'(x + t) < 0$. The case $g'(x + t) = 0$ will occur on a set of measure zero, and thus we have $w(x, t) < 0$ for this case. Similarly, if the minimum is attained at $y^* = x - t$, we will have $w(x, t) = g'(x - t) > 0$. We want to examine probabilities and consider how minima can change with a small change in the spatial coordinate.

Let time t be fixed and consider a small change in x from x to $x + \Delta x$. If the minimizer does not change, then the value of $w(x, t)$ also does not change by the argument above. If it does change, then so does g' , and consequently, $w(x, t)$. We calculate the probability that there is a transition in w , i.e. the probability of a lower minimum on $(x + \Delta x - t, x + \Delta x + t)$ compared to that on $(x - t, x + t)$. This is given by

$$\mathbb{P} \{ \min \{g(y) : |x - y| < t\} > \min \{g(y) : |x + \Delta x - y| < t\} \}. \quad (3.24)$$

In order for (3.24) to hold, there must be a minimum outside the region of overlap, i.e. one of two events: a minimum on the segment $(x + t, x + \Delta x + t)$ below the original minimum, or a minimum on $(x - t, x + \Delta x - t)$, as illustrated in Figure 1. In the latter case, the minimum attained on this segment fails to be in the new interval when x is increased by Δx , so the new minimum is a consequence of the original no longer being contained in the domain of consideration.

In the former case (with the new minimum attained on the right), the minimum moves lower, so $w(x, t) = \partial_x \min \{g(y)\}$ jumps from zero to a negative value. We express this probability as

$$\mathbb{P} \{w(x_-, t) > w(x_+, t)\}$$

$$= \mathbb{P} \{ \min \{g(y) : |x - y| < t\} > \min \{g(y) : x + t < y < x + \Delta x + t\} \} \quad (3.25)$$

In the latter case, the minimum on the left gives rise to a larger minimum in the central region, so the jump is to a minimum value.

The inequalities above are valid for any x and Δx . In particular, we could write a set of inequalities for a partition $\{x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\}$ of \mathbb{R} with corresponding $\Delta x_i := x_i - x_{i-1}$. We now perform the same computation for any x but with time $t + \Delta t$ to obtain the following:

$$\begin{aligned} & \mathbb{P} \{w(x, t + \Delta t) > w(x, t + \Delta t)\} \\ &= \mathbb{P} \left\{ \begin{array}{l} \min \{g(y) : |x - y| < t + \Delta t\} > \\ \min \{g(y) : x + t + \Delta t < y < x + \Delta x + t + \Delta t\} \end{array} \right\} \end{aligned} \quad (3.26)$$

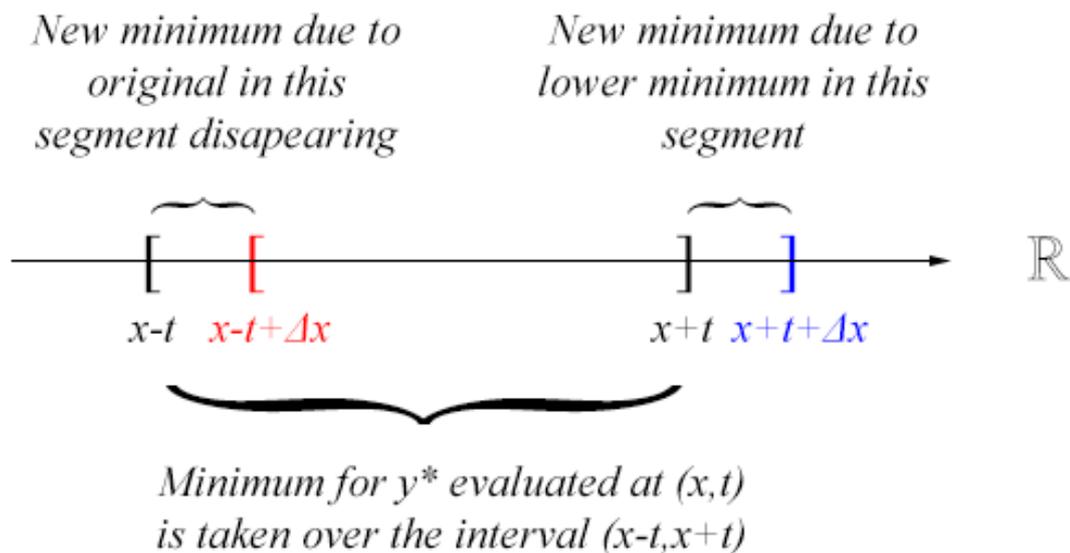


Figure 1. When moving from (x, t) to $(x + \Delta x, t)$, the interval of minimization is shifted Δx to the right. To have different minimizers at points x and $x + \Delta x$, at least one of two events must occur: (i) the minimizer was attained in the leftmost segment $(x - t, x - t + \Delta x)$ which is omitted for the minimization problem at $x + \Delta x$ and hence lost; (ii) a new, lower minimum is introduced by considering the segment $(x + t, x + t + \Delta x)$ on the right and hence gained.

We can now evaluate the probability in (3.26) to obtain the probability of a transition for $(x - \Delta t, t + \Delta t)$, i.e. obtain

$$\begin{aligned} & \mathbb{P} \{w(x - \Delta t, t + \Delta t) > w(x_+ - \Delta t, t + \Delta t)\} \\ &= \mathbb{P} \left\{ \begin{array}{l} \min \{g(y) : x - t - 2\Delta t < y < x + t\} > \\ \min \{g(y) : x + t < y < x + \Delta x + t\} \end{array} \right\}. \end{aligned} \quad (3.27)$$

We now compare (3.27) with (3.25) noting that the second terms are identical. One has the inequality

$$\min \{g(y) : |x - y| < t\} \geq \min \{g(y) : x - t - 2\Delta t < y < x + t\}. \quad (3.28)$$

In other words, in the notation of uniformly spaced points $\{x_i\}_{i \in \mathbb{Z}}$, we have

$$\mathbb{P}\{w(x_i, t) > w(x_{i+1}, t)\} \geq \mathbb{P}\{w(x_{i+1}, t + \Delta t) > w(x_{i+1} - \Delta t, t + \Delta t)\} \quad (3.29)$$

and

$$\mathbb{P}\{w(x_{i+1}, t) > w(x_i, t)\} \geq \mathbb{P}\{w(x_i, t + \Delta t) > w(x_i + \Delta t, t + \Delta t)\}. \quad (3.30)$$

By combining (3.29) and (3.30) and then summing over all i , one sees that the density of jumps is decreasing in time, i.e.

$$\sum_i \mathbb{P}\{\text{jump between } x_i \text{ and } x_{i+1}\} \quad (3.31)$$

$$\geq \sum_i \left\{ \begin{array}{l} \mathbb{P}\{w(x_{i+1}, t + \Delta t) > w(x_{i+1} - \Delta t, t + \Delta t)\} \\ + \mathbb{P}\{w(x_{i+1}, t + \Delta t) > w(x_{i+1} + \Delta t, t + \Delta t)\} \end{array} \right\}. \quad (3.32)$$

Further, if the process is stationary, one has a stronger result in that the local variation at a given x_i is also nonincreasing. In the limit, one may also write

$$\mathbb{P}\{w(x, t) > w(x_+, t)\} \geq \mathbb{P}\{w(x_+, t + \Delta t) > w([x - \Delta t]_+, t + \Delta t)\} \quad (3.33)$$

and

$$\mathbb{P}\{w(x, t) < w(x_-, t)\} \geq \mathbb{P}\{w(x_-, t + \Delta t) > w([x + \Delta t]_-, t + \Delta t)\}. \quad (3.34)$$

Thus, we can bound the probabilities that the right and left-hand limits of the solution at a particular point x in space change as time is increased slightly.

4. Application to a prototype case

As discussed earlier, an important application of our methodology involves the case where the flux function is given simply as $H(p) = |p|$. We will first consider this case for (1.1) to illustrate the main ideas and a convergence proof before proceeding to the more general case in the next section. As we shall see, though the formulation of this problem is simple, it raises a number of interesting points and the solution under our methods presents a clear geometric interpretation. Even for an initial condition given by a stochastic process such as Brownian motion, the essence of computing the solution profile is reduced to calculating in which of three regions Brownian motion has a minimum. To calculate the solution $w(x, t)$, one needs to determine whether the minimum of the Brownian motion is at $x - t$, $x + t$, or within $(x - t, x + t)$. The more general case is based on the same principles but involves more calculation.

To ease notation, we will present the following prescription as to how one determines the value of the solution profile $w(x, t)$ for a particular realization of one Brownian path and for one specific point $(x, t) \in \mathbb{R} \times (0, \infty)$. Implicit will be the existence of an underlying, appropriate partition of the real line, the technical details of which are omitted here and are to be outlined in the next section.

First note that the Legendre transform of the flux function H given above will be 0 inside a compact interval and $+\infty$ elsewhere. More specifically, we have

$${}^tL\left(\frac{y-x}{t}\right) = \begin{cases} 0 & y \in [x-t, x+t] \\ +\infty & y \in (-\infty, x-t) \cup (x+t, \infty) \end{cases}. \quad (4.1)$$

To this end, we consider the quantity

$$tL\left(\frac{y-x}{t}\right) + g(y). \quad (4.2)$$

One then has the remarkable simplification that minimizing the expression (4.2) over the entire real line is equivalent to computing the minimum of $g(y)$ in the interval $[x-t, x+t]$. Although finding the minimum of $g(y)$ for a stochastic process, say integrated Brownian motion, may not be so simple, for our purposes it will suffice to determine whether one of three distinct cases can occur. Indeed, either the minimum of $g(y)$ will occur at one of the endpoints or somewhere in the interior.

In our setup, we discretize Brownian motion in the following way. Without loss of generality (see Remark 2.10), consider the interval $[0, 1]$ and let $M \in \mathbb{N}$ be given. Divide the interval into 2^M pieces, resulting in a partition $\{r_n^{(M)}\}$, and choose random variables so that the discretized Brownian motion is constant on each subinterval, with its values prescribed by a random walk. Our goal is to show that the solution with initial conditions given by the discretized version of the integrated Brownian motion indeed converges to integrated Brownian motion as we let $M \rightarrow \infty$. We define our setup of the discretization more formally in Definition 5.2 before introducing our main results.

We will denote our discretized Brownian motion by W_M , and the integrated process as $I_M = \int W_M$. Note that our procedure superficially resembles the classical construction of Brownian motion as a limit of interpolated (and thus continuous) random variables. However, using the interpolated random walk would not be as useful in our case, as the minima of its integral will not general be at the specified n points (vertices of W_M and L).

We define

$$g'_M(r) = W_M(r) := W(r_i) \quad \text{for } r \in [r_i, r_{i+1}). \quad (4.3)$$

We then integrate and define

$$I_M(t) = \int_0^t W_M(s) ds, \quad (4.4)$$

where (4.4) defines our discretized integrated Brownian motion. The initial value problem we want to consider then becomes (using superscripts for M to ease notation)

$$w_t^M + |w^M|_x = 0, \quad w^M(x, 0) = g'_M(x). \quad (4.5)$$

We construct a solution to this discretized problem. A key result, derived in Section 3.1, is that

$$w^M(x, t) = \begin{cases} 0 & \text{if interior minimum} \\ g'_N(x-t) & \text{if minimum is left endpoint} \\ g'_N(x+t) & \text{if minimum is right endpoint} \end{cases}. \quad (4.6)$$

The events that the minimum occurs in these different cases (i.e., either somewhere in the interior of $[x-t, x+t]$ or at one of the endpoints) can be described as Gaussian integrals. In other words, the value of the solution is completely determined by the value of the initial condition depending on whether the minimum is attained at the left endpoint, in the interior, or at the right endpoint. The details are presented in the next section for the more general case, but take the form of a series of nested Gaussian integrals.

A second set of questions involve whether the solution converges, and if so in what sense to solutions to the limiting problem, i.e.

$$w_t + |w|_x = 0, \quad w(x, 0) = g'(x) \quad (4.7)$$

where g' is given by a Brownian motion. Indeed, one can show that we have convergence in probability of w^M to w in the following sense.

Given $A = B + C$, where $\mathbb{P}(|C| > \alpha)$ will be small, we will want to bound the probabilities of $\{A > 0\}$ and $\{A < 0\}$ in terms of sets only involving B . One may write

$$\mathbb{P}(B \leq -\alpha) - \mathbb{P}(C \geq \alpha) \leq \mathbb{P}(A < 0) \leq \mathbb{P}(B \leq \alpha) + \mathbb{P}(C \leq -\alpha). \quad (4.8)$$

We use Billingsley Theorem 37.9, p. 534 [3]. Let $n \in \mathbb{M}$ be fixed and let D be the dyadic rationals

$$D := \{k2^{-M} : k \in \mathbb{M}\} \quad (4.9)$$

Then we have for Brownian motion, W , with $\delta := 2^{-M}$, the bound for any $\alpha > 0$

$$\mathbb{P} \left[\sup_{r \in [0,1]} |W(t + \delta) - W(t)| > \alpha \right] \leq K \frac{\delta^2}{\alpha^4}, \quad (4.10)$$

where K is a positive real number. By continuity of Brownian motion, we can extend this inequality to all t (not just dyadic). We consider for simplicity the interval $[0, 1]$ so $k \in \{1, 2, \dots, 2^n\}$ and let $Q := 2^M$, (r_0, \dots, r_Q) be a set of equally spaced points with $r_0 := 0$ and $r_N := 1$ with spacing

$$\Delta r = \delta = 2^{-M}. \quad (4.11)$$

We have the, for any $r \in [0, 1]$, the bound

$$|I(r)| \leq \sup_{s \in [0,1]} |W(s) - W_N(s)|. \quad (4.12)$$

Now define the following:

$$\begin{aligned} \tilde{A}(s) &:= I(x-t) - I(s), \quad \tilde{B}(s) := I_M(x-t) - I_M(s) \\ \tilde{C}(s) &:= \{I(x-t) - I_M(x-t)\} + \{I(s) - I_M(s)\}. \end{aligned} \quad (4.13)$$

and the real numbers

$$\begin{aligned} A &:= \sup_{s \in [x-t, x+t]} \tilde{A}(s), \quad B := \sup_{s \in [x-t, x+t]} \tilde{B}(s), \\ C &:= \sup_{s \in [x-t, x+t]} \tilde{A}(s) - \sup_{s \in [x-t, x+t]} \tilde{B}(s) \end{aligned} \quad (4.14)$$

Thus, we have $A = B + C$, and C has the property

$$|C| \leq \sup_{(s \in [x-t, x+t])} |\tilde{C}(s)| > \alpha \quad (4.15)$$

Using the above, one can then write

$$\mathbb{P}\{\omega : |C| > \alpha\} \leq \mathbb{P}\left\{\omega : \sup_{s \in [0,1]} |I(s) - I_M(s)| > \alpha\right\} \leq K \frac{M^{-2}}{\alpha^4} \quad (4.16)$$

and recalling that $\delta = \frac{1}{Q}$ by definition, we have

$$\mathbb{P}(B < -\alpha) - K \frac{M^{-2}}{\alpha^4} \leq \mathbb{P}(A < 0) \leq \mathbb{P}(B < \alpha) + K \frac{M^{-2}}{\alpha^4} \quad (4.17)$$

Then, by choosing α and N appropriately, one has

$$\mathbb{P}\left(B < -M^{-\frac{2}{5}}\right) - KM^{-\frac{2}{5}} < \mathbb{P}(A < 0) < \mathbb{P}\left(B < M^{-\frac{2}{5}}\right) + KM^{-\frac{2}{5}} \quad (4.18)$$

for a pure number K independent of N . We use these inequalities in conjunction with the convergence theorems below.

4.1. Properties of the Brownian motion discretization

In particular, we want to prove that I_N defined by (4.4) is a Cauchy sequence in probability. Given an $\varepsilon > 0$ we choose J such that $\Delta r_{N_j} \leq \varepsilon$. Let $t \in [r_i^{(j)}, r_{i+1}^{(j)}]$. This means that for all $j, k > J$ we have, for any particular ω in the Polish space, the inequality

$$|W_{N_j}(t) - W_{N_k}(t)| \leq \sup_{t_1, t_2 \in [r_i^{(j)}, r_{i+1}^{(j)}]} |W(t_1) - W(t_2)|. \quad (4.19)$$

This means that for any ω and any real number α , one has

$$\begin{aligned} & P\left\{\omega : |W_{N_j}(t) - W_{N_k}(t)| > \alpha\right\} \\ & \leq P\left\{\omega : \sup_{t_1, t_2 \in [r_i^{(j)}, r_{i+1}^{(j)}]} |W(t_1) - W(t_2)| > \alpha\right\}. \end{aligned} \quad (4.20)$$

A similar bound is obtained for the integrals of the discretized Brownian motion. Using the definitions above we write, for $t \in [0, 1]$, and any particular ω so that W_{N_j} and W_{N_k} are continuous, bounded functions,

$$\begin{aligned} |I_{N_j}(t) - I_{N_k}(t)| & \leq \int_0^t |W_{N_j}(s) - W_{N_k}(s)| ds \\ & \leq \sup_{s \in [0, 1]} |W_{N_j}(s) - W_{N_k}(s)| \\ & \leq \sup_{s \in [0, 1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| \end{aligned} \quad (4.21)$$

so that for any ω and $\alpha \in \mathbb{R}$ we have

$$|I_{N_j}(t) - I_{N_k}(t)| \geq \alpha \text{ implies } \sup_{s \in [0, 1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| \geq \alpha. \quad (4.22)$$

Hence we have

$$\left\{\omega : |I_{N_j}(t) - I_{N_k}(t)| \geq \alpha\right\} \subset \left\{\omega : \sup_{s \in [0, 1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| \geq \alpha\right\}$$

$$P\left\{|I_{N_j}(t) - I_{N_k}(t)| \geq \alpha\right\} \leq P\left\{\sup_{s \in [0,1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| \geq \alpha\right\}. \quad (4.23)$$

Next, we claim that if s and r are in their respective closed, bounded intervals, and one has for some $C \in \mathbb{R}$ and any fixed ω the inequality

$$P\{|W(s+r) - W(s)| \geq \alpha\} \leq C \quad (4.24)$$

then one also has

$$P\left\{\sup_{s \in [0,1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| \geq \alpha\right\} \leq C. \quad (4.25)$$

To prove this note that for each ω the Brownian motion $W(s; \omega)$ is continuous, so its difference $|W(s+r) - W(s)|$ is also continuous in (r, s) and will have a maximum and minimum on $[0, \Delta r^{(j)}] \times [0, 1]$. Let (r^*, s^*) be a point of maximum. Then there is a sequence, (r_n, s_n) converging to (r^*, s^*) . We have thus

$$\sup_{s \in [0,1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| = |W(s^* + r^*) - W(s^*)|. \quad (4.26)$$

Let $R_n := |W(s_n + r_n) - W(s_n)|$ and $R^* := |W(s^* + r^*) - W(s^*)|$. For each ω we have by continuity in s for $W(s; \omega)$, the pointwise convergence, $R_n \rightarrow R^*$. I.e., for any $\varepsilon > 0$ we have convergence with probability 1, also expressed as (B p. 75)

$$P[|R_n - R^*| \geq \varepsilon \text{ i.o.}] = 0. \quad (4.27)$$

This implies convergence in probability, which implies convergence in distribution, i.e., for any α such that $P[R^* = \alpha] = 0$, which in our case is all α by continuity, one has

$$\lim_{n \rightarrow \infty} P[R_n > \alpha] = P[R > \alpha], \quad (4.28)$$

i.e., one has

$$\begin{aligned} \lim_{n \rightarrow \infty} P[|W(s_n + r_n) - W(s_n)| > \alpha] &= P[|W(s^* + r^*) - W(s^*)| > \alpha] \\ &= P\left[\sup_s \sup_r |W(s+r) - W(s)| > \alpha\right]. \end{aligned} \quad (4.29)$$

Hence, the assumption that $P[|W(s_n + r_n) - W(s_n)| > \alpha] \leq C$ implies the conclusion (4.25), proving the claim.

We use these in conjunction with a basic bound on Brownian motion to prove the following Lemma.

Lemma 4.1. *For any $\alpha > 0$, one has*

$$P\left\{|I_{N_j}(s) - I_{N_k}(s)| > \alpha\right\} \leq 3 \frac{|\Delta r^{(j)}|}{\alpha^4} \text{ for any } s \in [0, 1]. \quad (4.30)$$

Proof. A random variable X with finite k th moment satisfies Markov's inequality for any $k \in \mathbb{N}$

$$P\{|X| > \alpha\} \leq \frac{E[|X|^k]}{\alpha^k}. \quad (4.31)$$

We apply this to $X = W(t) - W(s)$ with $k := 4$. Note that X is a Gaussian with mean 0 and variance $\sigma^2 = |t - s|$. The fourth moment of this Gaussian is then $3\sigma^4 = 3|t - s|^2$.

Thus, we can write, for $r \in [0, \delta]$ the inequality

$$P\{|W(s+r) - W(s)| > \alpha\} \leq \frac{3|t - s|^2}{\alpha^4}. \quad (4.32)$$

From this inequality, one obtains

$$P\left\{\sup_{s \in [0,1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| > \alpha\right\} \leq \frac{3[\Delta r^{(j)}]^2}{\alpha^4}. \quad (4.33)$$

Combining this with the inequality (4.23), for $j, k > J$ one has for any $t \in [0, 1]$

$$\begin{aligned} P\{|I_{N_j}(t) - I_{N_k}(t)| > \alpha\} &\leq P\left\{\sup_{s \in [0,1]} \sup_{r \in [0, \Delta r^{(j)}]} |W(s+r) - W(s)| > \alpha\right\} \\ &\leq \frac{3[\Delta r^{(j)}]^2}{\alpha^4}. \end{aligned} \quad (4.34)$$

□

Our next lemma entails passing to a subsequence to achieve the convergence in the desired sense. We state this without proof as it follows easily from analysis.

Lemma 4.2. *There is a measurable function $I^*(\omega, s)$ such that $I_{N_j} \rightarrow I^*$ in probability and a subsequence converges a.u. in ω and uniformly in s .*

With the definition

$$B_M := \sup_{s \in [x-t, x+t]} \{I_M(x-t) - I_M(s)\} \quad (4.35)$$

one has from (4.18) the bounds

$$\begin{aligned} P\{B_M < -M^{-2/5}\} - KM^{-2/5} &< P\{A \leq 0\} \\ &< P\{B_M < M^{-2/5}\} + KM^{-2/5}. \end{aligned} \quad (4.36)$$

By Lemma 4.2 we know that there is a subsequence of $\{I_{N_j}\}$ converging a.u., although the convergence in measure should be adequate. We consider only the subsequence that converges almost uniformly in ω and also in measure in ω uniformly in the s variable. We will just refer to this subsequence as $\{I_j\}$ for brevity rather than $\{I_{N_j}\}$.

By containment of sets in conjunction with (4.36) for $j \in \mathbb{N} : j^{-2/5} < \varepsilon$, one has the inequalities

$$P(B_j < -\varepsilon) - Kj^{-2/5} < P(B_j < -j^{-2/5}) - Kj^{-2/5}$$

$$\begin{aligned}
&< P\{A \leq 0\} \\
&< P\{B_j < j^{-2/5}\} + Kj^{-2/5} < P(B_j < \varepsilon) + Kj^{-2/5}
\end{aligned} \tag{4.37}$$

Define, analogous to B_M above, the quantity

$$B^* := \sup_{s \in [x-t, x+t]} \{I^*(x-t) - I^*(s)\}.$$

Since Lemma 4.1 says that $\{I_{N_j}(\omega, s)\}$ is fundamental in measure uniformly in s , we know that there is an I^* to which I_{N_j} converges in measure. Then we have using I_j in place of I_{N_j} ,

$$I_j(x-t) - I(s) \rightarrow I^*(x-t) - I^*(s) \text{ in prob, unif in } s.$$

Using Lemma 4.2 we then have that

$$\begin{aligned}
\sup_s \{I_j(x-t) - I(s)\} &\rightarrow \sup_s \{I^*(x-t) - I^*(s)\} \text{ in prob, i.e. ,} \\
B_j &\rightarrow B^* \text{ in prob.}
\end{aligned}$$

Convergence in probability implies convergence in distribution, so that we have

$$P\{B_j < \alpha\} \rightarrow P\{B^* < \alpha\}.$$

Now using this in (4.37) we take limits and obtain,

$$P\{B^* < -\varepsilon\} \leq P\{A \leq 0\} \leq P\{B^* < \varepsilon\}. \tag{4.38}$$

If we consider only the subsequence that converges almost uniformly, then we have that for any $\delta > 0$ we have a set F such that $P(F) < \delta$ and on $\Omega \setminus F$ one has uniform convergence of continuous functions B_j so that B^* is continuous. Hence we can take the limit as $\varepsilon \rightarrow 0$ and obtain from (4.38) that

$$P\{B^* \leq 0\} \leq P\{A \leq 0\} \leq P\{B^* \leq 0\}. \tag{4.39}$$

Thus we have the following theorem.

Theorem 4.3. *In the limit $M \rightarrow \infty$, we have that $\mathbb{P}\{w_M(x, t) = g'_M(x-t)\} \rightarrow \mathbb{P}\{w(x, t) = g'(x, t)\}$ for $1 \leq i \leq 2N+1$. In other words, we have convergence in probability of the solution to the discrete problem to that of the continuous problem (1.1).*

Recalling that $A < 0$ is equivalent to $I(s)$ having its minimum at the left endpoint, i.e. (4.14), we see that the probability that the minimum is attained at the left endpoint is a limit of probabilities involving the discretizations I_M .

While we have developed these ideas in the context of $H(p) = |p|$, the same methodology can be implemented for general polygonal flux, as the latter simply introduces deterministic drift, and the bounds on Brownian motion remain valid.

5. Analysis of the piecewise linear flux function with Gaussian stochastic process initial data

We now consider the flux function in (1.1). To be precise, let the flux function H be piecewise linear with slopes given by $\{m_i\}_{i=1}^{N+1}$ in increasing order and break points $\{c_i\}_{i=1}^N$. The Legendre transform L will consequently have slopes $\{c_i\}_{i=1}^N$ and break points $\{m_i\}_{i=1}^{N+1}$, with slope $-c_{N+1-i}$ in the interval $I_i := [m_i, m_{i+1}]$, and infinite value outside the interval $[m_1, m_{N+1}]$. When evaluated at the argument $\frac{x-y}{t}$, the function exhibits limiting behavior under small or large time leading to interesting results in these limits for the solution of the minimization problem. For small time t , the function L is infinite outside a small interval, making it more likely that a minimum is obtained at a vertex of L . For large t , L can be approximated as a single linear segment. In this case the minimum is then likely to be at a vertex of g rather than at a vertex of L . In numerical studies, this can be illustrated by comparing the variance in the spatial variable x at given fixed times t_1, t_2 , etc. As is apparent in Figure 2 computed using [30], the total variation tends to decrease as a function of t . This augments the formal calculations shown in Section 3. This is despite the fact that the probabilistic variance at a point x increases in time.

In addition to the intuitive interpretation of the asymptotic limiting behavior of the solutions, we prove a rigorous result for a broad class of initial conditions. We consider a discretized approximation to a Gaussian stochastic process as an initial condition paired with the piecewise linear flux function. The virtue of this approach is that the result holds for any stochastic process in this class, and is exact without relying on the particular structure of say Brownian motion or Brownian bridge. We recall a result from [5].

Theorem 5.1. *Let L be polygonal convex, as described above, and g' be piecewise constant. Then*

$$\begin{aligned} w(x, t) &= g'(y^*(x, t)) \text{ when the minimizer } y^* \text{ is at a vertex of } L \\ &= L'\left(\frac{x - y^*(x, t)}{t}\right) \text{ when the minimizer } y^* \text{ is at a vertex of } g \end{aligned} \quad (5.1)$$

a.e. in x (for fixed $t > 0$) is a solution to (1.1). Should a minimum occur at a point where both g and L have a vertex, one has a discontinuity in the solution with $w(x-, t)$ taking the value of the left limit $g'(x, t)$ and $w(x, t)$ having the value $g'(x+, t)$. Note that for a fixed $t > 0$ and x a.e., one of the cases in (5.1) occurs.

This theorem will be needed as the final step in proving our main result of this section. Before stating the theorem, we clarify the choice of discrete points at which we will sample the stochastic process. This partition will depend on the value of x and t . Given x and t , the Legendre transform L of the flux function will only be finite on a finite interval, and therein L' , the quantity of interest, will take N different values. We partition each of these intervals into pieces at which we sample the values of L' and g' . In addition, we will need to consider the value at each of the vertices of L . One is then left with a *finite* number of points, and computing the minimum amongst these is a far more manageable task than calculating the minimum along a continuum. Indeed, our main result presents an explicit formula for computing this minimum, and thus an exact expression for the expectation value of the solution $w(x, t)$. One can then consider refining such partitions and taking a limit to the continuum. In a computational context, it may be necessary to first fix a partition, and restrict oneself to that same grid, but we leave the technical details to future works.

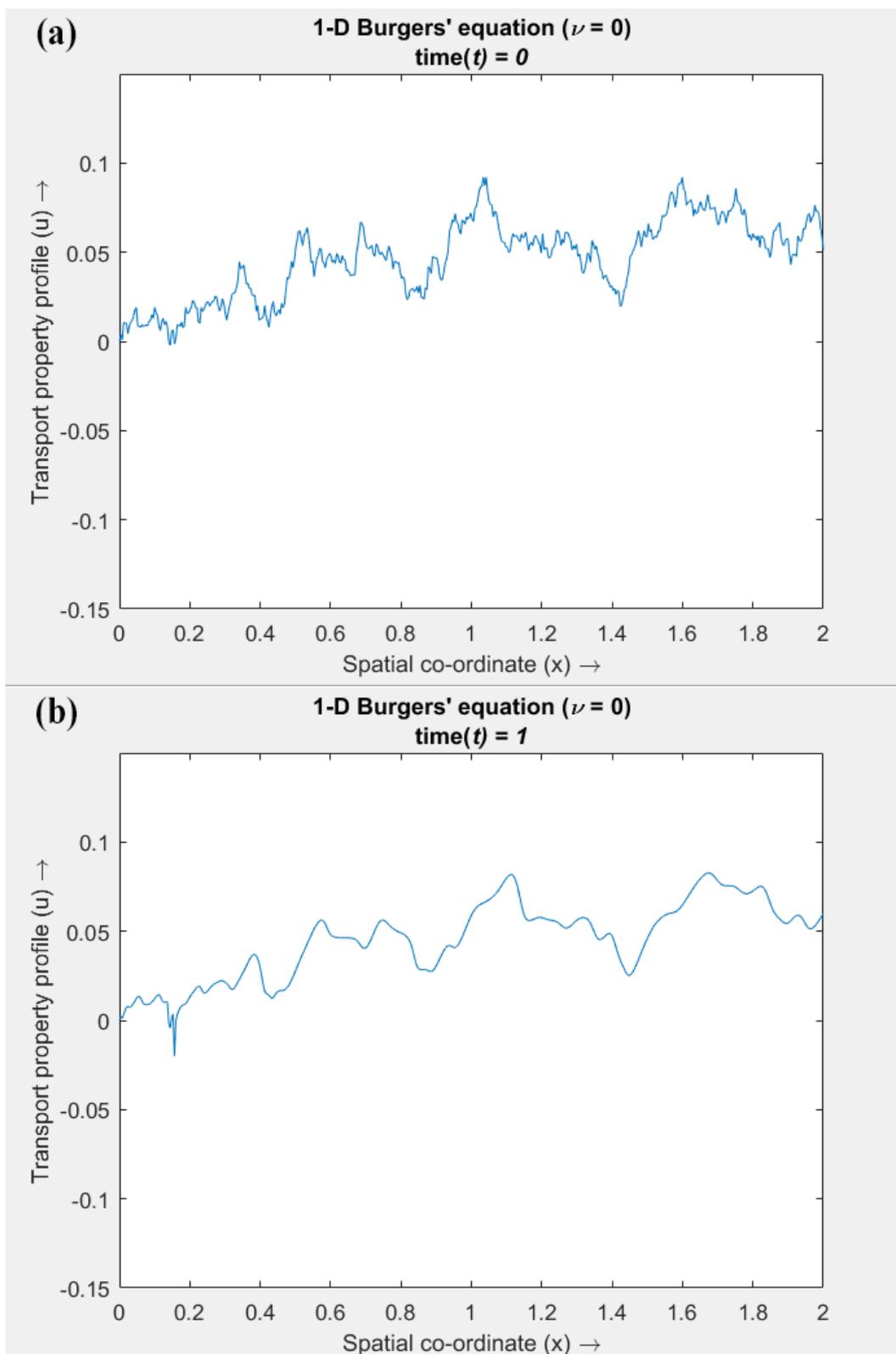


Figure 2. By performing numerical simulations using a finite difference scheme on Burgers' equation for Brownian motion initial conditions, one can see that the longer time behavior exhibits less variance in x and appears more smooth, in accordance with the analytic results.

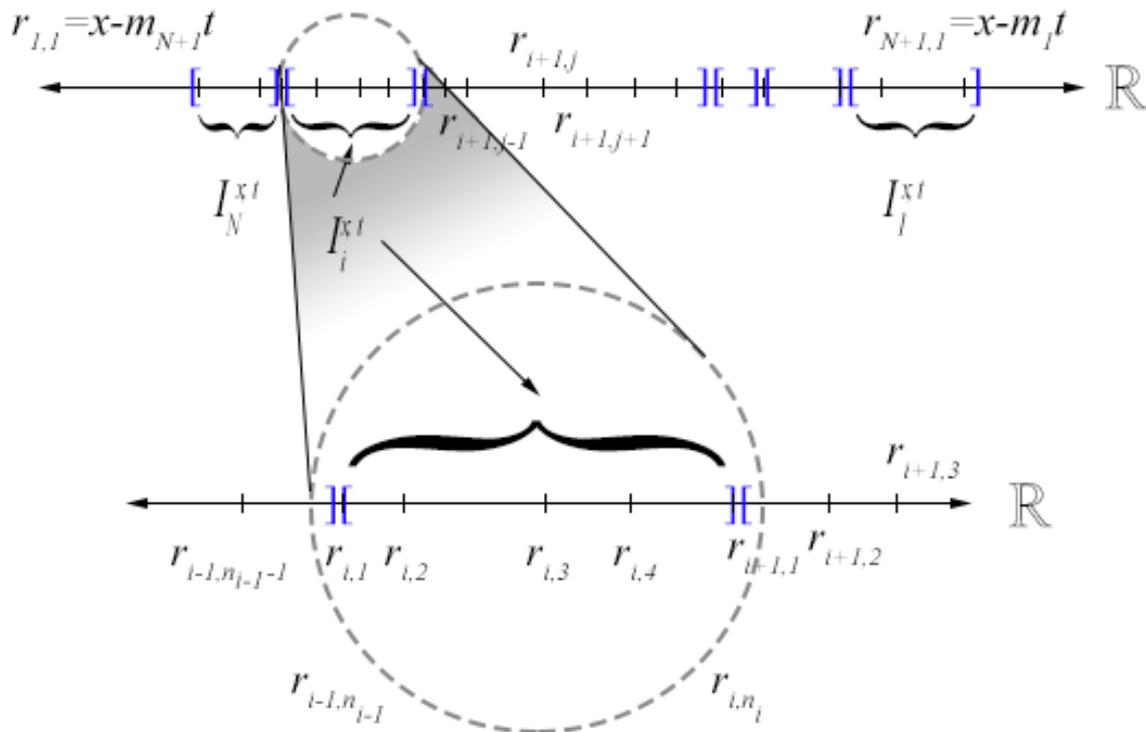


Figure 3. Illustration of the partition defined in Definition 5.2. Given a fixed x and t , we split the interval $I^{x,t} = [x - m_{N+1}t, x - m_1t]$ into subintervals $I_i^{x,t} = [x - m_{i+1}t, x - m_it]$ and partition each of these N intervals into n_i points. The problem of finding the minimum over the entire interval $I^{x,t}$ is then simplified to finding the minimum over these $\sum_{i=1}^N n_i = n$ points and $N + 1$ vertices of L , which are the endpoints of $I_i^{x,t}$.

Definition 5.2. On the interval $[x - m_{N+1}t, x - m_1t]$ we set grid points for a fixed x and t as follows. Note that L is divided into N segments on which it takes finite values, and has $N + 1$ associated vertices. Starting from the leftmost segment, we label these vertices as $r_{1,1}, \dots, r_{1,n_1}, r_{2,1}, \dots, r_{i,1}, \dots, r_{N+1,1}$. We number the n_i points that lie between each vertex sequentially using the second index, pairing them with the first index representing the vertex most immediately to its left, i.e. $r_{1,2}, \dots, r_{1,n_1}$ for points between the first and second vertex. We denote the total number of points (vertices and points between) by n , and define the intervals $I_i^{x,t} := (r_{i,1}, r_{i+1,1})$.

We illustrate the definition of these partitions in Figure 3.

Definition 5.3. Let x and t be given and choose partitions of $I_i^{x,t}$ as described in Definition 5.2. Let the number of such points be denoted by n_i , and define $\sum_i n_i =: n$. Define $A_{i,j}$ as the inverse of the covariance matrix of $X(s)$ with respect to these points $\{s_{i,j}\}$ in the subpartition $\{r_{j,2}, \dots, r_{j,n_j}\}$ and suppress the second index of A , s , etc. Denote the eigenvalues of A_i as $\tilde{\lambda}_i$, and define $\tilde{\mu}_i$ as the means of

each random variable, i.e.

$$\mu_{i,j} = \mathbb{E}\left\{g(r_{i,j}) + tL\left(\frac{x - r_{i,j}}{t}\right)\right\} \text{ for } j \neq 1, 1 \leq i \leq N. \quad (5.2)$$

The matrix A_i can be diagonalized by

$$A_i = U_i D_i U_i^T. \quad (5.3)$$

Similarly, define $A, \{\lambda_i\}, \{\mu_i\}$ etc. as the covariance matrix, eigenvalues, means of the random variables for the entire process within the range $(x - t, x + t)$.

Define also Q_i as follows:

$$Q_i = Q(i, x, t) := \min_{y \in I_i^{x,t}} \left\{ tL\left(\frac{x - y}{t}\right) + g(y) \right\}, \quad (5.4)$$

i.e. the minimum on the i th segment of the flux function L .

Theorem 5.4. Let the flux function H in the conservation law (1.1) be piecewise linear and convex, with N segments and $N + 1$ vertices. Let the initial condition $g'_N(s)$ be given by a discretized Gaussian stochastic process, as outlined in Definition 5.2. The probability that the minimum occurs on the i th segment, i.e.

$$p_i := \mathbb{P}\{Q(y^*(x, t)) = Q(y(i, x, t))\}, \text{ i.e., } w(x, t) = c_{N+1-i} \quad (5.5)$$

is given by

$$p_i = \sum_{j=2}^{n_i} (2\pi)^{-\frac{n}{2}} |A|^{\frac{1}{2}} \int_{-\infty}^{\infty} dx_{i,j} \prod_{(m,l) \neq (i,j)} \int_{x_{i,j}}^{\infty} dx_{m,l} e^{-\frac{1}{2}(\bar{x} - \bar{\mu})^T A (\bar{x} - \bar{\mu})}$$

for $1 \leq i \leq N$. (5.6)

The probability of such a minimum occurring at a vertex point of L is given by an expression similar to (5.6) with the integral over the j th vertex as the outermost integral (treating the vertex as a segment with only one point).

Finally, define the event that the minimum occurs at the j th vertex of L by R_j and set (for notational convenience) $p_{N+i} = \mathbb{P}\{R_i\}$. Then the expected value of the solution is given by

$$\begin{aligned} \mathbb{E}\{w(x, t)\} &= \mathbb{E}\left\{L'\left(\frac{x - y^*(x, t)}{t}\right)\right\} \\ &\quad + \sum_{i=1}^{N+1} \mathbb{P}\{R_i\} \mathbb{E}\{g'(y^*(x, t))\} \\ &= - \sum_{i=1}^N p_i c_{N+1-i} + \sum_{i=N+1}^{2N+1} p_i \mathbb{E}\{g'(y^*(x, t))\} \end{aligned} \quad (5.7)$$

a.e., that is, an average of the slopes of L weighted by the probability of the minimum occurring on the i th segment of L , plus terms at vertices of L .

Remark 5.5. In Theorem 5.4, we present closed-form probability of the solution taking on one of a number of values for a given x and t , as well as the expectation of the solution at this given value. In fact, we can also write an expression for the distribution of the local minimum on segment i as a function of its value s by instead taking the last integral in (5.6) up to s instead of ∞ .

We may also write expressions for the cdf of the value of $g + L$ on the vertex, and for the cdf of the minimum value obtained on just one segment. Consider first, for given $i \in [1, N]$, the domain $I_i^{x,t}$ (with parameters x, t) where L is a purely affine function. Then the cumulative probability distribution (cdf) for the minimum on this i th segment, i.e. the quantity Q_i is given by

$$F_{Q_i}(s) := (2\pi)^{-\frac{n_i}{2}} |A_i|^{\frac{1}{2}} \int_{-\infty}^s \dots \int_{-\infty}^s e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T A_i (\bar{x}-\bar{\mu})} dx_{i,n_i} \dots dx_{i,1}. \quad (5.8)$$

where $\bar{s} = (s, s, \dots, s)$ and $\bar{\mu}$ the vector given by $\mathbb{E} \left\{ g(\bar{y}) + tL \left(\frac{\bar{x}-\bar{y}}{t} \right) \right\}$ evaluated at the discrete set of points.

Furthermore, the cdf of a potential minimum that occurs at the j th of the $N + 1$ vertices of L is given by

$$F_{v_j}(s) = (2\pi)^{-\frac{1}{2}} |\mu_j|^{-\frac{1}{2}} \int_{-\infty}^s e^{-\frac{1}{2}\mu_j(x-\mu_j)^2} dx. \quad (5.9)$$

Remark 5.6. The case of the Gaussian stochastic process includes the deterministic case, i.e. when $p_i = 1$ for some i dependent on x and t and $p_j = 0$ for $j \neq i$.

Proof of Theorem 5.4. We consider an initial condition given by the discretized Gaussian stochastic process $g'_N(x)$. The key feature is that, using our methods, this problem of finding the minimum over the entire interval $(x - t, x + t)$ is reduced to finding a minimum over n points, n_i of which are on each segment, and $N + 1$ points that are on the vertices of L . Given an x and t , these n points are fixed. Since we have a joint probability density for the values of this function at any of these points, it is a matter of integrating over the appropriate region(s) to obtain the desired probabilities. From there we can calculate not only the expectation value $\mathbb{E}w(x, t)$ but the entire distribution of the solution w . To this end, let $\{Y_j\}_{j=1}^n$ be a set of random variables with a multivariate probability density given by

$$f(x; \mu, A) = (2\pi)^{-\frac{n}{2}} |A_i|^{\frac{1}{2}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T A_i (\bar{x}-\bar{\mu})}, \quad (5.10)$$

where $A_i := \sum_i^{-1}$ and \sum_i is the covariance matrix of the random variables in the subpartition ζ^i , and $\bar{x}, \bar{\mu} \in \mathbb{R}^{n_i}$ are vectors. Here the components of μ and x are related by $\mu_i = tL \left(\frac{x_i - y_i}{t} \right) + \mathbb{E}g(y_i)$. Thus, f contains all the information we will need to compute the minimum along the i th segment of L .

For simplicity, we write out the distribution over the points $n_{i,2}, \dots, n_{i,n_i}$ to have values less than or equal to s_i . This is expressed as

$$\begin{aligned} \mathbb{P} \{ Y_2 \leq s_1, \dots, Y_{n_i} \leq s_{n_i} \} &= \mathbb{P} \left\{ Y_i \in \bigotimes_{i=1}^n (-\infty, s_i] \right\} \\ &= \int_{-\infty}^{s_1} \dots \int_{-\infty}^{s_n} \frac{e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T A_i (\bar{x}-\bar{\mu})}}{(2\pi)^{\frac{n_i-1}{2}} |\sum_i|^{\frac{1}{2}}} dx_{n_i} \dots dx_2. \end{aligned} \quad (5.11)$$

Note we have suppressed the first index in $Y_{i,j}$ (which would be i) in these quantities over segment i for ease of notation. To find the probability that a minimum on one segment is below a value s , we can use

the result (5.11) and must all sum over all the (mutually exclusive) events that each random variable Y_i is a minimum as follows

$$\mathbb{P}\{Y_i \in (-\infty, s)^{n_i}\} = \sum_{j=2}^{n_i} (2\pi)^{-\frac{n_i-1}{2}} |A_i|^{\frac{1}{2}} \left\{ \int_{-\infty}^s \dots \int_{-\infty}^s e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T A_i(\bar{x}-\bar{\mu})} dx_{n_i} \dots dx_1 \right\}, \quad (5.12)$$

where we have $(n_i - 1)$ -fold integrals from $-\infty$ to s on the right-hand side of (5.12). A similar construction then yields the result (5.6).

We now want to use equation (5.1) from Theorem 5.1, which states the solution $w(x, t)$ takes the value of L' at any point where the minimum is achieved at a vertex of g , and the value g when achieved at a vertex of L . By accounting for both the possibility of a minimum at one of the N segments or at one of the vertices of L , we can write not only the expectation value of the solution at a given point, but the entire distribution. The expectation value is consequently

$$\begin{aligned} \mathbb{E}\{w(x, t)\} &= \mathbb{E}\left\{L' \left(\frac{x - y^*(x, t)}{t}\right)\right\} \\ &\quad + \sum_{j=1}^{N+1} \mathbb{P}\{R_j\} \mathbb{E}\{g'(y^*(x, t))\} \\ &= - \sum_{i=1}^N p_i c_{N+1-i} + \sum_{i=N+1}^{2N+1} p_i \mathbb{E}\{g'(y^*(x, t))\}, \end{aligned} \quad (5.13)$$

i.e., it is given by the average over these possibilities. The event of a minimum at a vertex of L may be more than a set of measure zero at the discrete level, but one might expect convergence to zero as we take the limit $\tilde{N} \rightarrow \infty$. The situation in regards to these two cases is illustrated in Figure 4. \square

For a specific realization of the stochastic process $g_{\tilde{N}}(x)$, the values of the solution w are within the range $\{c_i\}$. The solution w takes the value c_{N+1-i} when the minimum is achieved along the i th segment. If the minimum is achieved at a vertex of L , the expectation of the solution is zero by assumption, so the second set of terms drops out.

Remark 5.7. *As an aside, we note that one can transform the integrals to a different coordinate system and simplify the exponentials to only square terms. However, the domain of integration is transformed in such a way that the integration is nontrivial and results in a nested sequence of integrated error functions. By construction, the matrix A is symmetric and real, so the eigenvalues are real there exists a matrix U_i such that $U^T A U = D$, where D is the diagonal matrix consisting of the eigenvalues λ_{n_i} . Any repeated eigenvalues are assigned to $\{\lambda_i\}$ as needed, i.e. $\lambda_k = \lambda_{k+1}$ is permitted. Furthermore, we know $U^{-1} = U^T$, so we can write*

$$(\bar{x} - \bar{u})^T A (\bar{x} - \bar{\mu}) = (\bar{x} - \bar{\mu}) U D U^{-1} (\bar{x} - \bar{\mu}) = z^T D z = \sum_{i=1}^n \lambda_i z_i^2 \quad (5.14)$$

where

$$\bar{z} = U^{-1} (\bar{x} - \bar{\mu}) \quad (5.15)$$

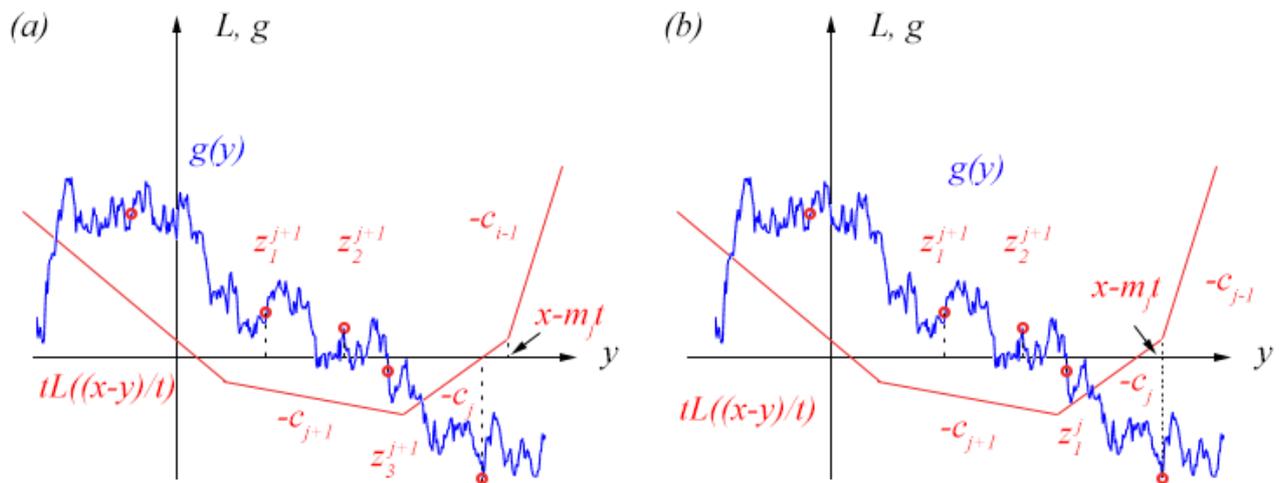


Figure 4. We are seek the minimum of $L\left(\frac{x-y}{t}\right) + g(y)$ amongst a discrete, finite set of points $\{z_i^j\}$ which are chosen based on a fixed x and t . Here we display integrated Brownian motion, but these arguments hold for *any* stochastic process: (a) In most cases, none of the points $\{z_i^j\}$ coincide with the vertices of the function L , so the minimum on each segment is well-defined and one computes $\mathbb{E}\{w\}$ as described below; (b) Occasionally the points $\{z_i^j\}$ may intersect $\{x - m_k t\}$ for some i, j, k but this is a set of measure zero and can be ignored.

for a unitary matrix U so that $|U|^2 = 1$. Hence, given any vector μ and a real, symmetric matrix A , we can write

$$-\frac{1}{2}(\bar{x} - \bar{\mu}) = -\frac{1}{2} \sum_{i=1}^{n_i} \lambda_i z_i^2. \tag{5.16}$$

Thus, one has

$$f(z_j, \{\lambda_j\}) = (2\pi)^{-n/2} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2} \sum_{i=1}^{n_i} \lambda_i z_i^2} \tag{5.17}$$

or equivalently

$$\mathbb{P}\{Z_1 \leq z_1, \dots, Z_{n_i} \leq z_{n_i}\} = \int \dots \int_{\Omega} \frac{e^{-\frac{1}{2} \sum_{i=1}^{n_i} \lambda_i z_i^2}}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} d\tilde{z}_{n_i} \dots d\tilde{z}_1 \tag{5.18}$$

over the transformed domain Ω (rotated and scaled from the original region $[-\infty, s]^{n_i}$).

6. Exact and approximate results for the density of shocks

We now consider the density of shocks in the solution $w(x, t)$, i.e. a spatial coordinate x where

$$w(x_-, t) \neq w(x_+, t). \tag{6.1}$$

This was discussed in the special case $H(p) = |p|$ in Section 3. Recall from Section 4 the notation $I_i^{x,t} = (r_{i,1}, r_{i+1,1})$ to track the intervals along which L has slope d_i . From the arguments above, it is clear

that at point x where a shock occurs, the minimizer $y^*(x, t)$ jumps from one segment of L to another. By continuity, we can assume that for sufficiently small $\varepsilon > 0$, $y^*(x - \varepsilon, t)$ is located within the same segment in the limit. In other words, one writes

$$\{y^*(x - \varepsilon, t) \mid \varepsilon < \varepsilon_0\} \subset I_i^{x,t} \setminus \partial I_i^{x,t} \quad (6.2)$$

Let $\Delta x > 0$ and consider the following. Suppose that $y_1 := y^*(x - \Delta x, t)$ is located where L has slope d_1 and $y_2 := y^*(x, t)$ where L has slope d_2 . When we shift the spatial coordinate by amount Δx , the value of the quantity to be minimized changes by $R_1 = -d_1 \Delta x$ at y_1 and by $R_2 = -d_2 \Delta x$ at y_2 . Thus, the net change is a decrease $(d_2 - d_1) \Delta x$ at y_2 relative to y_1 . This means a shock from this limit can only occur when $d_2 > d_1$. This argument is illustrated in Figure 5. If we let Δx be negative, i.e. we consider the points x and $x + \Delta x$ instead of $x - \Delta x$ and x , we can have a shock for the case $d_1 > d_2$. We now want to find an expression for the probability density of such shocks.

To be more precise, we let $a := \Delta x(d_2 - d_1)$. We want to compute the probability that the minimum shifts from one segment to another. First consider the simpler problem with only two segments, and the probability of the minimum shifting from a particular point on the first segment to a particular point on the second segment. To illustrate this in a simple way, consider the points $r_{1,2}$ and $r_{2,2}$ among the $n_1 + n_2 + 1$ points in the system. We also use the notation that the random variable $Y_{i,j}$ corresponds to the point $r_{i,j}$, suppressing the subscript r . We compute the probability that the variable $Y_{1,2}$ associated with the point $r_{1,2}$ is at least s , $Y_{2,2}$ is in the range between s and $s + a$, and $Y_{1,i}$, $Y_{2,i}$, and $Y_{3,1}$ are all at least $s + a$ for all $i \neq 2$ as follows

$$\begin{aligned} & \mathbb{P}\{Y_{1,2} = s_{1,2}, Y_{2,2} \in (s_{1,2}, s_{1,2} + a), Y_{j,i} \geq s_{1,2} + a \text{ for all } i \neq 2, 1 \leq j \leq 3\} \\ &= \int_{s_{1,2}}^{s_{1,2}+a} ds_{2,2} \prod_{(m,n) \neq (1,2) \text{ or } (2,2)} \int_{s_{1,2}+a}^{\infty} ds_{m,n} f(s_{1,1}, s_{1,2}, \dots, s_{3,1}) \end{aligned} \quad (6.3)$$

where f is the joint distribution for the variables $s_{1,1}, \dots, s_{3,1}$ at the $n_1 + n_2 + 1$ points, and we have fixed $s_{1,2}$. Now, we want to divide by a and take the limit. When we differentiate with respect to x , the integral with respect to $s_{2,2}$ (in the $(n_1 + 1)$ -th argument of f) drops out due to the Fundamental Theorem of Calculus, assuming sufficient continuity properties. The probability of transitioning from a minimum from one segment to the other can be written explicitly as follows

$$\begin{aligned} \mathbb{P}_{(1,2),(2,2)} &:= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{P}\{\text{min at } (1, 2) \text{ at } (x, t) \text{ and min at } (2, 2) \text{ at } (x + \Delta x, t)\} \\ &= (d_2 - d_1) \int_{-\infty}^{\infty} ds_{1,2} \prod_{(u,v) \neq (1,2),(2,2)} ds_{u,v} \int_{s_{1,2}}^{\infty} f(s_{1,1}, s_{1,2}, \dots, s_{2,1}, s_{2,2}, \dots, s_{3,1}, \dots). \end{aligned} \quad (6.4)$$

i.e. the $s_{2,2}$ entry is evaluated at $s_{1,2}$ in f . This can be generalized to computing the probability of transition from one segment to another, which will entail $n_i n_j$ terms. These terms correspond to a jump from a minimum at s attained at one of the n_i possible points from segment i to one of the n_j possible points on segment j . Now we extend this computation from one point to another, to all points in a segment i to another segment j . Let

$$\tilde{P}_{i,j}(s; a) := \mathbb{P}\left\{ \min_{2 \leq k \leq n_i} Y_{i,k} = s, \min_{2 \leq k \leq n_j} Y_{j,k} \in (s, s + a) \right\},$$

$$\begin{aligned}
 \mathbb{P}_{i,j} &:= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{P} \left\{ w(x, t) = d_i \text{ and } w(x + \Delta x, t) = d_j \right\} \\
 &= \sum_{k=1}^{n_i} \sum_{k'=1}^{n_j} (d_{k'} - d_k) \int_{-\infty}^{\infty} ds \prod_{(m,n) \neq (i,k), (j,k')} \\
 &\int_s^{\infty} ds_{m,n} f(s_{1,1}, \dots, s_{N+1,1}) |_{s_{i,k}=s_{j,k'}=s}.
 \end{aligned} \tag{6.6}$$

for the density of any transitions of a minimum between segments i and j . The joint probability distribution function f is as outlined in (5.8).

One can diagonalize this matrix. Indeed, denoting the eigenvalues of A by $\{\lambda_m\}_{m=1}^n$, A can thus be diagonalized by $U^T A U = D$, where the diagonal matrix D is given by $D_{jk} = \delta_{jk} \lambda_j$. We may then rewrite the integrand $f(\cdot)$ in (6.6) as $\prod_{l=1}^{n-2} e^{-\lambda_l z_l^2}$ and have, similar to the proof of Theorem 5.4:

$$P_{i,j}(s) = (2\pi)^{-\frac{n}{2}} \sum_{k=1}^{n_i} \sum_{k'=1}^{n_j} (d_{k'} - d_k) |A|^{\frac{1}{2}} \int_{\Omega_{(i,k),(j,k')}} \prod_{l=1}^{n-2} e^{-y_l^2} dy_l \tag{6.7}$$

where the transformed domains of integration depends upon k and k' . The density of transitions, in a range $(s, s + ds)$ can be expressed using these results. The integration over the domain Ω may appear complicated due to the high dimension of the space and the nested nature of the integrals, but one can make some simplifications. For example, as is further illustrated in the next section, the eigenvalue spectrum of A falls off rapidly. Thus, only a (smallest) few of the eigenvalues are relevant and many of the integrals can be well-approximated as δ -functions about $y_l = 0$.

6.1. Transitions between vertex and non-vertex points

Above, we have considered the question of transition from an overall minimum from a segment i to a different segment j . There is also the issue of a minimum obtained at one of the N segments of L replaced by a minimum at a vertex, and vice versa. We will quantify the probability of such an occurrence.

First, we comment that without discretization, we would have a transition from $w(x, t) = d_i$ to $w(x, t) = g'(j, 1)$ if $Y_{i,k} = s$ and $Y_{j,1} \in (s, s - \Delta x(d_i + g'(j, 1)))$. Now, with the discrete model, we will evaluate exactly the probability that if $Y_{i,k} = s$, then $Y_{j,1} \in (s, s - \Delta x(d_i + g'(j, 1)))$ and vice versa. We assume for the first case that $d_i + g'(j, 1) < 0$.

Recall that the grid we place is for a fixed x and t . In this section, we will consider changing x (or similarly, t) by such a small amount Δx that all of the original grid points within $(x - t, x + t)$ remain, and no new grid points are introduced into the interval $(x + \Delta x - t, x + \Delta x + t)$. Thus, the points $r_{k,l}$ remain unchanged. We also assume that the minimum attained at (x, t) occurs at the vertex $r_{j,1}$ and make the approximation that L still nearly has a vertex at $r_{j,1}$ despite the Δx shift.

The first step is to identify the condition for local minima at vertices $r_{j,1}$ to overtake $Y_{i,k}$, the absolute minimum over \mathbb{R} located at some k th point on the i th segment. We use $Y_{i,k}^{new}$ to denote the new local minimum when L is shifted by Δx . We calculate

$$Y_{i,k}^{new} := Y_{i,k} - \Delta x d_i > Y_{j,1} + g'(r_{j,1}) \Delta x =: Y_{j,1}^{new} \tag{6.8}$$

In other words, (6.8) is the requirement for a transition from a minimum at the vertex of g at $r_{i,k}$ to the vertex of L originally at $r_{j,1}$ that has now shifted. Of course, we also need the condition

$$Y_{i,k} < Y_{j,1} \quad (6.9)$$

The condition (6.8) is equivalent to

$$Y_{i,k} - \Delta x(d_i + g'(r_{j,1})) > Y_{j,1} \quad (6.10)$$

Once we have established conditions (6.9–6.10), we can now calculate that probability using the discrete setup in which we deal with only the points $r_{i,1}, r_{i,2}, \dots$, i.e. perform the calculation using Gaussian integrals as follows. Let $\mathbb{P}_{i,j}^{\Delta x}$ be defined as the probability of a transition from segment i of L to the vertex point $r_{j,1}$ of L , so that

$$\begin{aligned} \mathbb{P}_{i,j}^{\Delta x} &:= \sum_{k=1}^{n_i} \int_{-\infty}^{\infty} ds \mathbb{P} \{ Y_{i,k} = s, Y_{j,1} \in (s, s - (d_i + g'(r_{j,1})) \Delta x) \} \\ &= \sum_{k=1}^{n_i} \int_{-\infty}^{\infty} ds \int_s^{s - (d_i + g'(r_{j,1})) \Delta x} dx_{j,1} \\ &\quad \prod_{(m,n) \neq (i,k), (j,1)} \int_s^{\infty} dx_{m,n} f(x_{1,1}, \dots, x_{N+1,1}) |_{x_{i,k}=s} \\ \mathbb{P}_{i,j} &:= \lim_{\Delta x \rightarrow 0} \frac{\mathbb{P}_{i,j}^{\Delta x}}{\Delta x} = - (d_i + g'(r_{j,1})) \sum_{k=1}^{n_i} \int_{-\infty}^{\infty} ds \\ &\quad \prod_{(m,n) \neq (i,k), (j,1)} \int_s^{\infty} dx_{m,n} f(\dots) |_{x_{i,k}=s, x_{j,1}=s} \\ &= \mathbb{P} \{ w(x_-, t) = d_i \text{ and } w(x_+, t) = g'(r(j, 1)) \}. \end{aligned} \quad (6.11)$$

The calculation for the opposite transition, which can only occur when $d_i + g'(r(j, 1))$ is positive, is

$$\begin{aligned} \mathbb{P}_{i,j} &:= \mathbb{P} \{ w(x_-, t) = g'(r(j, 1)) \text{ and } w(x_+, t) = d_i \} \\ &= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \sum_{k=1}^{n_i} \int_{-\infty}^{\infty} ds \mathbb{P} \{ Y_{j,1} = s, Y_{i,k} \in (s, s + (d_i + g'(j, 1)) \Delta x) \} \\ &= (d_i + g'(r_{j,1})) \int_{-\infty}^{\infty} ds \prod_{(m,n) \neq (i,k), (j,1)} \int_s^{\infty} dx_{m,n} f(\dots) |_{x_{i,k}=x_{j,1}=s}. \end{aligned} \quad (6.12)$$

7. Applications to selected examples and open problems

The results of the preceding sections apply in broad generality for any kind of randomness that falls under the umbrella of Gaussian stochastic processes. In addition to the general results proven for an arbitrary Gaussian stochastic process, it is also interesting to analyze the behavior of the solution to the conservation law under a few specific cases. In particular, we will illustrate the results for the cases of Brownian motion and Brownian bridge. As these processes correspond to the initial condition g' , we

must also note that $g(x)$ will be a Gaussian stochastic process for both of these cases, termed integrated Brownian motion and integrated Brownian bridge, respectively. For each of these cases, the covariance matrix takes a special form as was described in Section 2.

7.1. Brownian motion and Brownian bridge as initial conditions

In cases such as Brownian motion, Brownian bridge, or Ornstein-Uhlenbeck, one observes specific structure in the matrices and integrals entailed in potential applications of Theorem 5.4, which may make it somewhat more tractable to numeric computation than an arbitrary Gaussian stochastic process. For example, in some cases the relevant inverses of covariance matrices can be neglected aside from a narrow band of elements close to the diagonal. For example, consider an integrated Brownian motion $S(t)$ and let $\{t_i\}_{i=1}^n$ be given. Observe that the covariance matrix Σ for integrated Brownian motion from (2.9) can be written as [27]

$$\Sigma_{ij} = \frac{1}{N^3} \min(i, j)^2 \left(\frac{\max(i, j)}{2} - \frac{\min(i, j)}{6} \right), \quad (7.1)$$

so it is symmetric with respect to permuting i and j . One of the main quantities of interest in our results is of course the inverse A of this matrix. Upon taking N large, one observes an interesting property in that a substantial fraction of the diagonal elements of the matrix A are concentrated at one value. Numerical computation indicates that this number is close to 14.354... independent of N . Furthermore, the smallest of the eigenvalues start off at one value, with many concentrated near it, and then increase sharply, as illustrated in Figure 6. In addition, the numerically significant parts of the columns of the matrix A are repetitive.

We denote the smallest eigenvalue by λ_1 . Due to the nature of the integrands taking the form

$$\int e^{-\lambda_n z_n^2} dz_n, \quad (7.2)$$

there will be a rapid falloff as λ_n increases. As a first-order approximation, one may assume the first N' eigenvalues have the value λ_1 exactly, and ignore the rest as they will not make a meaningful contribution to the solution. From the minimization perspective, we are simply neglecting the terms corresponding to points at which the probability of attaining the minimum is very small. This will lead to a simplification of the form where one can set $\lambda_i = \lambda_1$ for a number n' of the eigenvalues, and ignore the subsequent ones.

For the shock densities, one observes a similar phenomenon, and expects a transition probability of the form

$$P_{i,i'} \cong (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} \sum_{k=1}^{n_i} \sum_{k'=1}^{n_{i'}} (d_{k'} - d_k) \int_{-\infty}^s \dots \int_{-\infty}^s e^{-\frac{1}{2}(\bar{s}-\bar{\mu})^T A_i(\bar{s}-\bar{\mu})} \cdot e^{-\left(\frac{\lambda_i}{2}[U^T(\bar{s}-\bar{\mu})]_k\right)^2 - \left(\frac{\lambda_{i'}}{2}[U^T(\bar{s}-\bar{\mu})]_{k'}\right)^2} ds_{N+1,1} \dots ds_{1,1} ds. \quad (7.3)$$

where $\bar{\mu}$ and A now correspond to the mean and inverse covariance matrix on segments i and i' together.

In Section 2, the covariance matrix for Brownian bridge in the range where it is defined is very similar to that of Brownian motion, despite some key differences in the processes. Outside of $[0, T]$

(under linear transformation), the variance of the integrated Brownian motion is constant. In applying this initial condition, one must set $g'(x) = g(x) = 0$ outside of a finite range $[0, T]$. In fact, one observes the same qualitative behavior for the eigenvalues of its inverse covariance matrix. This will lead to a result similar to (5.6) for the solution to the conservation law with Brownian bridge initial data. Another interesting application is that of a stationary Ornstein-Uhlenbeck process, i.e. with a covariance matrix invariant under translation, which may facilitate computations for further results. The advantage of this process is that it has constant variance, making it advantageous for some physical models where the randomness is fairly homogenous, rather than Brownian motion where the process is fixed at 0 and variance grows with time.

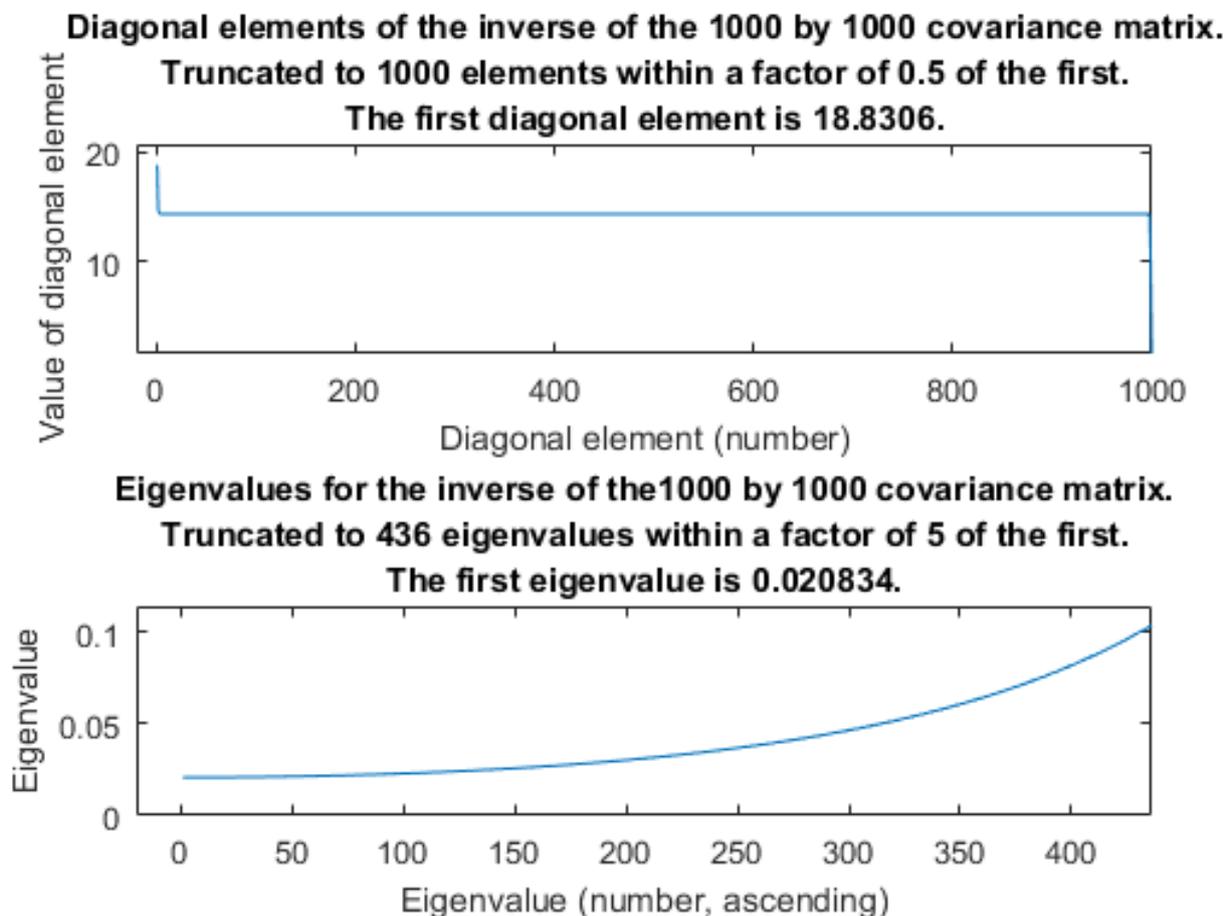


Figure 6. When N is large, the diagonal elements of the inverse covariance matrix A are concentrated very close to a particular value which can be calculated as approximately 14.3... The eigenvalues are also concentrated near one value, and sharply increase. Larger eigenvalues can be neglected due to the asymptotic behavior of the error function.

7.2. Approximation to smooth flux

A classical result by Dafermos [8] utilizing shock techniques describes how solutions to conservation laws with smooth flux can be approximated by those of polygonal flux. Our theorems can be combined with Dafermos' work in order to use polygonal flux as a building block for an arbitrary smooth flux function with randomness of the form of any Gaussian stochastic process.

One may also take the limits of $N \rightarrow \infty$ and make the partitions $\{r_{i,j}\}$ increasingly fine in our solution (5.6). In this case, one should obtain the result in [12] for which Burgers' equation is considered with Brownian initial conditions and exact solutions are computed. Our approach thus provides an extension of exact, closed-form results without relying on the exact properties of a particular flux function.

8. Appendix: results for Gaussian stochastic processes

Here we present sketches of the proof of several of the results we cited earlier in Section 2.

Proof of Proposition 2.8. (b) Define $\{t_i\}_{i=1}^n$ as before, we need to show that we can write

$$\begin{aligned} S(t_1) &= a_{11}R_1 + \dots + a_{1m}R_m + \mu_1 \\ &\dots \\ S(t_n) &= a_{n1}R_1 + \dots + a_{nm}R_m + \mu_n \end{aligned} \quad (8.1)$$

We can write the integral as an approximate sum with interpolation points (r_1, \dots, r_n) given as $r_i := \frac{j}{Nt_n}$, so that

$$S(t) \cong S_N(t) := \sum_{i=1}^{Nt} \frac{1}{N} W\left(\frac{j}{Nt_n}\right), \quad \frac{Nt}{Nt_n} \leq t < \frac{Nt+1}{Nt_n} \quad (8.2)$$

By definition, then

$$\begin{aligned} NS_N(t_1) &= W\left(\frac{1}{N}\right) + W\left(\frac{2}{N}\right) + \dots + W\left(\frac{Nt_1}{N}\right) \\ &\dots \\ NS_N(t_n) &= W\left(\frac{1}{N}\right) + W\left(\frac{2}{N}\right) + \dots + W\left(\frac{Nt_n}{N}\right), \end{aligned} \quad (8.3)$$

i.e. the only difference in the identities in (8.3) is where one terminates the sum. Since $W(t)$ is Gaussian stochastic by the part (a), we may write

$$W\left(\frac{i}{N}\right) = b_{1i}R_1 + \dots + b_{1m}R_m + \tilde{\mu}_i \quad (8.4)$$

By substituting (8.4) into (8.3) and defining new constants

$$a_{i1} := \frac{(b_{i1} + b_{i2} + \dots + b_{i,Nt_1})}{N},$$

we obtain the result that integrated Brownian motion is also a Gaussian stochastic process by using the Central Limit Theorem for functions ([26], p. 399).

We do not prove (c) and (d) here as the ideas are similar. In particular, since a Brownian bridge is a linear combination of $W(t)$ and $W(T)$, it can easily be expressed in terms of independent, identically distributed normal random variables. \square

Proof of Theorem 2.9. (a) We first consider the covariance matrix for times s and t . Since piecewise drift does not alter the covariance matrix, we assume without loss of generality that this term is zero. Letting $W(t)$ be a Brownian motion, one observes

$$\begin{aligned} \text{Cov} \left[\int_0^s W(r) dr, \int_0^t W(r) dr \right] &= \mathbb{E} \left\{ \int_0^t W(r) dr \int_0^t W(r') dr' \right\} \\ &= \int_0^s \int_0^t \mathbb{E} \{ W(r) W(r') \} dr' dr \\ &= \int_0^s \int_0^t \min \{ r, r' \} dr dr' \\ &= s^2 \left(\frac{t}{2} - \frac{s}{6} \right). \end{aligned} \quad (8.5)$$

The cross term, $\mathbb{E} \left\{ \int_0^t W(r) dr \right\} \mathbb{E} \left\{ \int_0^t W(r') dr' \right\}$, vanishes since $\mathbb{E} \{ W(t) \} = 0$ for Brownian motion.

Using a discretization argument as in [27], one may also construct (8.5) from first principles.

(b) Let $Z(t)$ be an integrated Brownian bridge constructed as in Definition 2.6. As in part (a) we consider the sum approximation to the Brownian bridge assuming no drift, and consider first only two points $0 \leq s \leq t \leq T$. We make note of (2.9) and also note

$$\text{Cov} \left[W(T), \int_0^t W(r) dr \right] = \text{Cov} \left[W(t), \int_0^t W(r) dr \right] \quad (8.7)$$

since $W(T) - W(t)$ is independent of events at t and before. Continuing from (8.7), we have

$$\begin{aligned} \text{Cov} \left[W(T), \int_0^t W(r) dr \right] &= \mathbb{E} \left\{ W(t) \int_0^t W(r) dr \right\} \\ &= \int_0^t \mathbb{E} \{ W(t) \} W(r) dr \\ &= \int_0^t \mathbb{E} \{ (W(t) - W(r) + W(r)) W(r) \} dr \\ &= \int_0^t \mathbb{E} \{ W(r)^2 \} dr = \frac{t^2}{2}, \end{aligned} \quad (8.8)$$

noting that the second term in the definition of covariance vanishes since $\mathbb{E} \{ W(t) \} = 0$. One then writes, expanding the definition of $Z(t)$ and collecting terms from (8.7) and (8.8):

$$\text{Cov} [Z(s), Z(t)] = s^2 \left(\frac{t}{2} - \frac{s}{6} \right) - \frac{1}{T} \frac{t^2 s^2}{4}. \quad (8.9)$$

for times $0 \leq \{t_i\}_{i=1}^N \leq T$.

Now consider $0 \leq s \leq T < t$. Then we have

$$\text{Cov} [Z(s), Z(t)] = \text{Cov} \left[S(s) - \frac{s}{T} W(T), S(T) - W(T) \right]$$

One obtains similar results by (stochastically) reflecting the process to consider it on the interval $[-T, T]$ for symmetry. One can also obtain these results by discretizing the process and writing the approximation of the integral, then limiting to the continuum using the Central Limit Theorem for functions. We also note that although the process $Z(s)$ is constant outside the interval $[0, T]$ (or $[-T, T]$ depending on the definition), $Z(s)$ does not vanish for $s > T$ with probability 1. \square

Acknowledgments

The author thanks the referee for review of the manuscript.

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. D. Applebaum, *Levy Processes and Stochastic Calculus*, 2 Eds, Cambridge: Cambridge University Press, 2009.
2. J. Bertoin, *Levy Processes*, Cambridge: Cambridge University Press, 1996.
3. P. Billingsley, *Probability and Measure*, New York: Wiley, 2012.
4. Y. Briener, E. Grenier, *Sticky particles and scalar conservation laws*, SIAM J. Numer. Anal., **35** (1998), 2317–2328.
5. C. Caginalp, *Minimization Solutions to Conservation Laws with Non-Smooth and Non-Strictly Convex Flux*, AIMS Mathematics, **3** (2018), 96–130.
6. M. Chabanol, J. Duchon, *Markovian solutions of inviscid Burgers equation*, J. Stat. Phys., **114** (2004), 525–534.
7. M. Crandall, L. Evans, P. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **282** (1984), 487–502.
8. C. Dafermos, *Polygonal approximations of solutions of the initial value problem for a conservation law*, J. Math. Anal. Appl., **38** (1972), 33–41.
9. C. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics*, 3 Eds, New York: Springer, 2010.
10. E. Weinan, G. Rykov, G. Sinai, *Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics*, Commun. Math. Phys., **177** (1996), 349–380.
11. C. Evans, *Partial Differential Equations*, 2 Eds, New York: Springer, 2010.
12. L. Frachebourg, P. Martin, *Exact statistical properties of the Burgers equation*, J. Fluid Mech., **417** (2000), 323–349.
13. D. G. Gilbarg & N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, New York: Springer, 1977.
14. P. Groeneboom, *Brownian motion with a parabolic drift and Airy functions*, Probab. Theory Rel., **81** (1989), 79–109.
15. M. Hairer, J. Maas, H. Weber, *Approximating rough stochastic PDEs*, Comm. Pure Appl. Math., **67** (2013), 776–870.
16. H. Holden, N. Risebro, *Front Tracking for Hyperbolic Conservation Laws*, New York: Springer, 2015.

17. E. Hopf, *The partial differential equation $u_t + uu_x = \mu u_{xx}$* , Comm. Pure Appl. Math., **3** (1950), 201–230.
18. D. Kaspar, F. Rezakhanlou, *Scalar conservation laws with monotone pure-jump Markov initial conditions*, Probab. Theory Relat. Fields, **165** (2016), 867–899.
19. A. Kaufman, H. Lim, J. Glimm, *Conservative front tracking: the algorithm, the rationale and the API*, Bulletin of the Institute of Mathematics, **11** (2016), 115–130.
20. D. Khoshnevisan, Z. Shi, *Chung’s Law for Integrated Brownian Motion*, T. Am. Math. Soc., **350** (1998), 4253–4264.
21. P. Lax, *Hyperbolic systems of conservation laws II*, Comm. Pure Appl. Math., **10** (1957), 537–566.
22. P. Lax, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, Society for Industrial and Applied Mathematics, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, 1973.
23. G. Menon, *Complete integrability of shock clustering and Burgers turbulence*, Arch. Ration. Mech. An., **203** (2012), 853–882.
24. G. Menon, R. Pego, *Universality classes in Burgers turbulence*, Comm. Math. Phys., **273** (2007), 177–202.
25. G. Menon, R. Srinivasan, *Kinetic theory and Lax equations for shock clustering and Burgers turbulence*, J. Stat. Phys., **140** (2010), 1195–1223.
26. M. Pinsky, S. Karlin, *An Introduction to Stochastic Modeling*, Burlington: Elsevier, 2011.
27. S. Ross, *Introduction to Probability Models*, 10 Eds, Burlington: Elsevier, 2010.
28. H. Royden, P. Fitzpatrick, *Real Analysis*, 4 Eds, Boston: Prentice Hall, 2010.
29. Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach*, New York: Springer, 2010.
30. S. Shankar, *Burgers Equation in 1D and 2D*, 2012. Available from:
<https://www.mathworks.com/matlabcentral/fileexchange/38087-burgers-equation-in-1d-and-2d?focused=5246985&tab=function>
31. M. Slemrod, (2013) *Admissibility of the weak solutions for the compressible Euler equations, $n \geq 2$* , Philos. T. R. Soc. A, **371** (2013), pp20120351.
32. D. She, R. Kaufman, H. Lim, et al. *Handbook of Numerical Analysis*, Elsevier **17**, 383–402.
33. A. Vol’pert, *Spaces BV and quasilinear equations*, Mat. Sb. (N.S.), **73** (1967), 255–302.



AIMS Press

©2018 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)