APPROXIMATION SCHEMES FOR VISCOSITY SOLUTIONS OF FULLY NONLINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

BY BENJAMIN SEEGER

CEREMADE, Université Paris-Dauphine and Collège de France seeger@ceremade.dauphine.fr

The aim of this paper is to develop a general method for constructing approximation schemes for viscosity solutions of fully nonlinear pathwise stochastic partial differential equations, and for proving their convergence. Our results apply to approximations such as explicit finite difference schemes and Trotter–Kato type mixing formulas. The irregular time dependence disrupts the usual methods from the classical viscosity theory for creating schemes that are both monotone and convergent, an obstacle that cannot be overcome by incorporating higher order correction terms, as is done for numerical approximations of stochastic or rough ordinary differential equations. The novelty here is to regularize those driving paths with nontrivial quadratic variation in order to guarantee both monotonicity and convergence.

We present qualitative and quantitative results, the former covering a wide variety of schemes for second-order equations. An error estimate is established in the Hamilton–Jacobi case, its merit being that it depends on the path only through the modulus of continuity, and not on the derivatives or total variation. As a result, it is possible to choose a regularization of the path so as to obtain efficient rates of convergence. This is demonstrated in the specific setting of equations with multiplicative white noise in time, in which case the convergence holds with probability one. We also present an example using scaled random walks that exhibits convergence in distribution.

1. Introduction. We construct numerical schemes to approximate viscosity solutions of fully nonlinear pathwise stochastic partial differential equations, and prove that they converge under quite general assumptions. Among the approximations that we study are finite-difference schemes and Trotter–Kato type product formulas. The former raise the possibility of numerical implementation, which we justify with precise error estimates in the first-order setting.

More precisely, given a finite horizon $T > 0$, we consider pathwise viscosity solutions of the initial value problem

$$(1.1) \quad \begin{cases} du = F(D^2u, Du)\, dt + \sum_{i=1}^{m} H^i(Du) \circ dW^i & \text{in } \mathbb{R}^d \times (0, T] \text{ and} \\ u(\cdot, 0) = u_0 & \text{in } \mathbb{R}^d, \end{cases}$$

where $W = (W^1, W^2, \ldots, W^m) : [0, T] \to \mathbb{R}^m$ is a continuous path and the initial datum $u_0 : \mathbb{R}^d \to \mathbb{R}$ is bounded and uniformly continuous. The precise assumptions on $H = (H^1, H^2, \ldots, H^m) : \mathbb{R}^d \to \mathbb{R}^m$ and $F : S^d \times \mathbb{R}^d \to \mathbb{R}$, where S^d is the space of symmetric matrices, are specified later in the paper. We emphasize here that F is assumed to be degenerate elliptic, that is, $F(X, p) \leq F(Y, p)$ whenever $p \in \mathbb{R}^d$ and $X, Y \in S^d$ satisfy $X \leq Y$.

The technical assumptions and theorems are stated in full generality later in the paper. First, we describe the main results in a simplified context to provide a flavor for what is to

follow. Afterwards, we provide some background on the notion of pathwise viscosity solutions, the history of the study of the equation, and its applications. The Introduction concludes with a description of the organization of the rest of the paper.

We note that, in the sequel, the term "classical viscosity theory" refers to the Crandall–Ishii–Lions [10] theory of viscosity solutions, which applies to (1.1) when W is continuously differentiable, or to the theory of equations with L^1-time dependence put forth by Ishii [16] and Lions and Perthame [22] for Hamilton–Jacobi equations, and by Nunziante [28] in the second order case, which includes (1.1) when W has bounded variation.

1.1. *The main results.*   Assume for now that $d = m = 1$, F and H are both smooth, and F depends only on u_{xx}, so that (1.1) becomes

$$(1.2) \qquad du = F(u_{xx})\, dt + H(u_x) \circ dW \quad \text{in } \mathbb{R} \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R},$$

or, in the first order case, when $F \equiv 0$,

$$(1.3) \qquad du = H(u_x) \circ dW \quad \text{in } \mathbb{R} \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}.$$

Here and throughout the paper, the solutions of equations like (1.1), (1.2), and (1.3) are to be understood in the pathwise, or stochastic, viscosity sense (see Definitions 3.1 and 3.2 below).

The approximations are constructed through the use of a scheme operator, which, for $h > 0$, $0 \le s \le t \le T$, and $\zeta \in C([0, T]; \mathbb{R})$, is a map $S_h(t, s; \zeta) : BUC(\mathbb{R}) \to BUC(\mathbb{R})$, whose properties will be made more precise in Section 4. Here, $BUC(\mathbb{R}^d)$ is the space of bounded, uniformly continuous functions on $\mathbb{R}^d$.

Throughout the paper, the symbol $\mathcal{P}$ denotes a partition of $[0, T]$ and $|\mathcal{P}|$ its mesh size, that is,

$$\mathcal{P} := \{0 = t_0 < t_1 < \cdots < t_N = T\} \quad \text{and} \quad |\mathcal{P}| := \max_{n=0,1,\ldots,N-1} (t_{n+1} - t_n).$$

Given such a partition $\mathcal{P}$ and a path $\zeta \in C([0, T]; \mathbb{R})$, usually a piecewise linear approximation of W, we first define the function $v_h(\cdot; \zeta, \mathcal{P})$ by

$$(1.4) \qquad \begin{cases} v_h(\cdot, 0; \zeta, \mathcal{P}) := u_0, \\ v_h(\cdot, t; \zeta, \mathcal{P}) := S_h(t, t_n; \zeta) v_h(\cdot, t_n; \zeta, \mathcal{P}) \\ \quad \text{for } n = 0, 1, \ldots, N - 1 \text{ and } t \in (t_n, t_{n+1}]. \end{cases}$$

The strategy is to choose families of approximating paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$ satisfying

$$(1.5) \qquad \lim_{h \to 0^+} \|W_h - W\|_\infty = 0 = \lim_{h \to 0^+} |\mathcal{P}_h|,$$

in such a way that the function

$$(1.6) \qquad u_h(x, t) := v_h(x, t; W_h, \mathcal{P}_h)$$

is an efficient approximation of the solution of (1.1).

As an example of the types of schemes studied in this paper, we consider here the following adaptation of the Lax–Friedrichs finite difference approximation, a formulation for which can be found in the work of Crandall and Lions [11] in the classical viscosity setting.

For some $\epsilon_h > 0$, define

$$(1.7) \qquad \begin{aligned} S_h(t, s; \zeta) u(x) := {}& u(x) + H\left(\frac{u(x+h) - u(x-h)}{2h}\right)(\zeta(t) - \zeta(s)) \\ &+ \left(F\left(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2}\right)\right. \\ &\left. + \epsilon_h\left(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2}\right)\right)(t - s). \end{aligned}$$

The first result, which is qualitative in nature, applies to the simple setting above as follows:

THEOREM 1.1.   *Assume that, in addition to* (1.5), W_h *and* $\mathcal{P}_h$ *satisfy*

$$|\mathcal{P}_h| \leq \frac{h^2}{\|F'\|_\infty} \quad and \quad \epsilon_h := h\|\dot{W}_h\|_\infty \xrightarrow{h \to 0} 0.$$

Then, as $h \to 0$, *the function* u_h *defined by* (1.6) *using the scheme operator* (1.7) *converges locally uniformly to the solution* u *of* (1.2).

We obtain explicit error estimates for finite difference approximations of the stochastic Hamilton–Jacobi equation (1.3). The results below are stated for the following scheme, which is defined, for some $\theta \in (0, 1]$, by

(1.8)
$$S_h(t, s; \zeta)u(x) := u(x) + H\left(\frac{u(x+h) - u(x-h)}{2h}\right)(\zeta(t) - \zeta(s))$$
$$+ \frac{\theta}{2}(u(x+h) + u(x-h) - 2u(x)).$$

Note that this corresponds to choosing $\epsilon_h := \frac{\theta h^2}{2(t-s)}$ in (1.7).

The main tool for proving rates of convergence is the following pathwise estimate. For the remaining results in the Introduction, it is assumed that, for some $L > 0$, the initial datum u_0 is Lipschitz continuous with $\|u_0'\|_\infty \leq L$.

THEOREM 1.2.   *There exists* $C > 0$ *depending only on the Lipschitz constant* L *such that, if* $h > 0$, $\zeta \in C([0, T], \mathbb{R})$ *is piecewise linear over the partition* $\mathcal{P}$ *such that*

$$\max_{n=0,1,\dots,N-1} |\zeta(t_{n+1}) - \zeta(t_n)| \leq \frac{\theta}{\|H'\|_\infty}h,$$

and v *solves* (1.3) *with the path* ζ, *then, for all* $\epsilon > 0$,

$$\sup_{(x,t)\in\mathbb{R}\times[0,T]} |v_h(x, t; \zeta, \mathcal{P}) - v(x, t)|$$

$$\leq \frac{1}{\epsilon}\sum_{n=0}^{N-1}(t_{n+1} - t_n)^2 + C\sqrt{N}h + \max_{s,t\in[0,T]}\left\{C|\zeta(s) - \zeta(t)| - \frac{|s-t|^2}{2\epsilon}\right\}.$$

The rates of convergence are then established by choosing families of paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$ in order to optimize the estimate from Theorem 1.2.

We do so first for an arbitrary, fixed continuous path W with modulus of continuity $\omega : [0, \infty) \to [0, \infty)$. For $h > 0$, define ρ_h implicitly by

(1.9)
$$\lambda := \frac{(\rho_h)^{1/2}\omega((\rho_h)^{1/2})}{h} < \frac{\theta}{\|H'\|_\infty},$$

and let the partition $\mathcal{P}_h$ and path W_h satisfy

(1.10)
$$\begin{cases} \mathcal{P}_h := \{n\rho_h \wedge T\}_{n\in\mathbb{N}_0}, \qquad M_h := \lfloor(\rho_h)^{-1/2}\rfloor, \\ \quad \text{and, for } k \in \mathbb{N}_0 \text{ and } t \in \left[kM_h\rho_h, (k+1)M_h\rho_h\right), \\ W_h(t) := W(kM_h\rho_h) + \left(\frac{W((k+1)M_h\rho_h) - W(kM_h\rho_h)}{M_h\rho_h}\right)(t - kM_h\rho_h). \end{cases}$$

THEOREM 1.3. *There exists $C > 0$ depending only on L such that, if u_h is constructed using* (1.6) *and* (1.8) *with $\mathcal{P}_h$ and W_h as in* (1.9) *and* (1.10), *and u is the pathwise viscosity solution of* (1.3), *then*

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \left|u_h(x,t) - u(x,t)\right| \leq C(1+T)\omega\left((\rho_h)^{1/2}\right).$$

When W is a Brownian motion, we study the problem from different points of view, depending on whether the focus is on almost-sure convergence or convergence in distribution.

As a special case of Theorem 1.3, the approximating paths and partitions may be taken to satisfy (1.10) with ρ_h given by

$$(1.11) \qquad \lambda := \frac{(\rho_h)^{3/4}|\log \rho_h|^{1/2}}{h} < \frac{\theta}{\|H'\|_\infty}.$$

Alternatively, the constructions may be achieved through the use of certain stopping times:

$$(1.12) \quad \begin{cases} T_0 := 0, \qquad T_{k+1} := \inf\left\{t > T_k : \max_{r,s\in[T_k,t]} \left|W(r) - W(s)\right| > \frac{h^{1/3}}{|\log h|^{2/3}}\right\}, \\[2mm] W_h(t) := W(T_k) + \frac{W(T_{k+1}) - W(T_k)}{T_{k+1} - T_k}(t - T_k) \quad \text{for } t \in [T_k, T_{k+1}), \\[2mm] M_h := \left\lceil \frac{\|H'\|_\infty}{(h|\log h|)^{2/3}} \right\rceil, \\[2mm] \text{and} \quad \mathcal{P}_h := \left\{t_n := T_k + (n - kM_h)\frac{T_{k+1} - T_k}{M_h} : kM_h \leq n < (k+1)M_h, \right. \\[2mm] \left. \qquad k \in \mathbb{N}_0\right\}. \end{cases}$$

The various definitions for $\mathcal{P}_h$ and W_h above, while technical, are all made with the same idea in mind, namely, to ensure that the approximation W_h is "mild" enough with respect to the partition. In particular, for any consecutive points t_n and t_{n+1} of the partition $\mathcal{P}_h$, and for sufficiently small h, the ratio

$$\frac{|W_h(t_{n+1}) - W_h(t_n)|}{h}$$

should be less than some fixed constant. This is a special case of the kind of Courant–Lewy–Friedrichs (CFL) conditions required for the schemes in this paper, which are discussed in more detail in the following sections.

THEOREM 1.4. *Suppose that W is a Brownian motion, and assume either that $\mathcal{P}_h$ and W_h are as in* (1.10) *with ρ_h defined by* (1.11), *or $\mathcal{P}_h$ and W_h are as in* (1.12). *If u_h is constructed using* (1.6) *and* (1.8), *and u is the solution of* (1.3), *then there exists a deterministic constant $C > 0$ depending only on L and λ such that, with probability one,*

$$\limsup_{h\to 0} \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \frac{|u_h(x,t) - u(x,t)|}{h^{1/3}|\log h|^{1/3}} \leq C(1+T).$$

The final type of result involves convergence in distribution in the space $BUC(\mathbb{R}^d \times [0, T])$. Here, the paths W_h are taken to be appropriately scaled simple random walks. More precisely,

for some probability space $(\mathcal{A}, \mathcal{G}, \mathbb{P})$,

$$
(1.13) \quad
\begin{cases}
\lambda := \dfrac{(\rho_h)^{3/4}}{h} \le \dfrac{\theta}{\|H'\|_\infty}, \qquad M_h := \lfloor (\rho_h)^{-1/2} \rfloor, \\[2mm]
\mathcal{P}_h := \{t_n\}_{n=0}^N = \{n\rho_h \wedge T\}_{n \in \mathbb{N}_0}, \\[2mm]
\{\xi_n\}_{n=1}^\infty : \mathcal{A} \to \{-1, 1\} \quad \text{are independent,} \\[2mm]
\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = \dfrac{1}{2}, \qquad W(0) = 0, \quad \text{and} \\[2mm]
W_h(t) := W_h(kM_h\rho_h) + \dfrac{\xi_k}{\sqrt{M_h\rho_h}}(t - kM_h\rho_h) \\[2mm]
\quad \text{for } k \in \mathbb{N}_0, t \in \big[kM_h\rho_h, (k+1)M_h\rho_h\big).
\end{cases}
$$

THEOREM 1.5.   *If u_h is constructed using* (1.6) *and* (1.8) *with W_h and $\mathcal{P}_h$ as in* (1.13), *and u is the solution of* (1.3) *with W equal to a Brownian motion, then, as $h \to 0$, u_h converges to u in distribution.*

1.2. *Background for the study of* (1.1).   When W is continuously differentiable, or of bounded variation, the symbol dW^i in equation (1.1) stands for the time derivative $\frac{d}{dt}W^i(t) = \dot{W}^i(t)$ and "$\circ$" denotes multiplication. As already noted, the classical viscosity theory applies in this context.

The problem becomes more complicated when W is merely continuous, and therefore, possibly nowhere differentiable or of infinite variation. In many examples of interest, W is the sample path of a stochastic process, such as Brownian motion, and then the symbol "$\circ$" is regarded as the Stratonovich differential. More generally, W may be a geometric rough path, a specific instance being a Brownian motion enhanced with its Stratonovich iterated integrals.

The notion of pathwise viscosity solutions for equations like (1.1) was developed by Lions and Souganidis, first for Hamiltonians depending smoothly on the gradient Du [24], and later for nonsmooth Hamiltonians [25]. The comparison principle was proved in [26], and equations with Hamiltonians depending nonlinearly on u were considered in [27]. The theory has since been extended to treat Hamiltonians with spatial dependence, as by Friz, Gassiat, Lions, and Souganidis [12], or by the author [29]; these papers use techniques developed by Lions and Souganidis for more general settings that appear in forthcoming works [23]. An alternative existence result relying on Perron's method can be found in the work of the author [30]. Many more details and results are summarized in the notes of Souganidis [31].

The setting in which H depends linearly on the gradient has been explored from the point of view of rough path theory by many authors, including, but not limited to, Caruana, Friz, and Oberhauser [8] and Gubinelli, Tindel, and Torrecilla [14]. The semilinear problem was also studied by Buckdahn and Ma [5, 6] using the pathwise control interpretation.

It is of particular interest to have a way to analyze (1.1) when H is nonlinear and not necessarily C^1, because of the application, via the level set method, to the theory of the propagation of fronts with a stochastically perturbed normal velocity. For example, if, for $t > 0$, $\Gamma_t \subset \mathbb{R}^d$ is a smooth, $(d-1)$-dimensional surface moving with normal velocity

$$
(1.14) \qquad\qquad\qquad V = -\kappa + \alpha\, dW,
$$

where κ is the mean curvature of the surface, $\alpha \in \mathbb{R}$ is a constant, and dW is white noise in time, and if Γ_t is the 0-level set of some function $u(\cdot, t)$, that is, $\Gamma_t = \{x \in \mathbb{R}^d : u(x, t) = 0\}$, then, formally, u solves the equation

$$
(1.15) \qquad du = \left(\Delta u - \left\langle D^2u \frac{Du}{|Du|}, \frac{Du}{|Du|} \right\rangle \right) dt + \alpha |Du| \circ dW \quad \text{in } \mathbb{R}^d \times (0, T).
$$

This is a special case of (1.1) for which F is singular. The stochastic viscosity interpretation of (1.15) has been used by Souganidis and Yip [34] to exhibit stochastic selection principles for some examples of nonuniqueness in mean curvature flow, and by Lions and Souganidis [23] to establish a sharp interface limit for the Allen–Cahn equation perturbed with an additive, mild approximation of time-white noise. For the latter problem, it was proved that, for some $\alpha \in \mathbb{R}$, the limiting front has a normal velocity as in (1.14).

As far as we know, the results in this paper on approximation schemes for stochastic viscosity solutions are the first of their kind. We are also aware of a work by Hoel, Karlsen, Risebro, and Storrøsten [15] using numerical methods to study a related class of equations, namely, stochastic scalar conservation laws.

1.3. *Organization of the paper.* Section 2 begins with a discussion of the theory of monotone approximation schemes in the classical viscosity setting, as well as the difficulties faced for pathwise equations. In Section 3, we recall some definitions and results from the theory of pathwise (stochastic) viscosity solutions. Some of the material may be found in [24, 25], or [31], while other facts, whose proofs are given here, are developed by Lions and Souganidis in a forthcoming work [23].

In Section 4, we make the notion of the scheme operator S_h more precise, and use the method of half-relaxed limits to prove that, for an appropriate family of partitions $\{\mathcal{P}_h\}_{h>0}$ and paths $\{W_h\}_{h>0}$ as in (1.5), if u_h is defined by (1.4) and (1.6), then u_h converges locally uniformly to the solution of (1.1). Various examples are presented to which the general convergence result may be applied.

Section 5 lays the framework for the quantitative analysis of schemes for stochastic Hamilton–Jacobi equations by proving a generalization of the pathwise estimate in Theorem 1.2. This result is then used in Section 6 to obtain explicit rates of convergence, such as those stated in Theorems 1.3 and 1.4, as well as the result on convergence in distribution as in Theorem 1.5.

1.4. *Notation.* Throughout most of the proofs in this paper, the symbol C will stand for a generic constant that may change from line to line, and whose dependence will be specified or made clear from context. $C^{0,1}(\mathbb{R}^d)$ is the space of Lipschitz continuous functions, and $\|Du\|_\infty$ is the Lipschitz constant for a function $u \in C^{0,1}(\mathbb{R}^d)$. For $\alpha \in (0,1)$, $C^\alpha([0,T])$ denotes the space of α-Hölder continuous paths on $[0,T]$, and $[W]_{\alpha,T}$ is defined to be the Hölder seminorm of $W \in C^\alpha([0,T])$. Given a continuous path W and $s, t \in [0,T]$, the maximum oscillation of W between s and t is denoted by

$$\operatorname{osc}(W, s, t) := \max_{r_1, r_2 \in I_{s,t}} |W(r_1) - W(r_2)| = \max_{I_{s,t}} W - \min_{I_{s,t}} W,$$

where $I_{s,t} := [\min(s,t), \max(s,t)]$.

The spaces of upper- and lower-semicontinuous functions on $\mathbb{R}^d \times [0,T]$ are respectively $USC(\mathbb{R}^d \times [0,T])$ and $LSC(\mathbb{R}^d \times [0,T])$, and $(B)UC(U)$ is the space of (bounded) uniformly continuous functions on a domain U.

For $a \in \mathbb{R}$, $\lfloor a \rfloor$ and $\lceil a \rceil$ denote respectively the largest (smallest) integer k satisfying $k \leq a$ ($k \geq a$). The mesh-size of a partition $\mathcal{P} = \{0 = t_0 < t_1 < t_2 < \cdots < t_N = T\}$ of $[0,T]$ is defined by $|\mathcal{P}| := \max_{n=0,1,2,\ldots,N-1}(t_{n+1} - t_n)$. S^d is the space of symmetric d-by-d matrices, and, for $X, Y \in S^d$, the inequality $X \leq Y$ means $\xi \cdot X\xi \leq \xi \cdot Y\xi$ for all $\xi \in \mathbb{R}^d$. The set of positive integers is written as $\mathbb{N}$, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

2. Monotone schemes for viscosity solutions.

2.1. *The classical viscosity setting.* It is well known that viscosity solutions of the nonlinear degenerate parabolic equation

$$(2.1) \qquad u_t = F(D^2u, Du) \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d$$

satisfy a comparison principle. That is, if u and v are respectively a sub- and super-solution of (2.1), then, for all $t \in [0, T]$,

$$(2.2) \qquad \sup_{x \in \mathbb{R}^d} (u(x, t) - v(x, t)) \le \sup_{x \in \mathbb{R}^d} (u(x, 0) - v(x, 0)).$$

In particular, if $u(\cdot, 0) \le v(\cdot, 0)$, then $u(\cdot, t) \le v(\cdot, t)$ for all future times $t > 0$.

Moreover, (2.1) is stable under local uniform convergence. That is, if, for $n \ge 0$, $u_{0,n}, u_0 \in BUC(\mathbb{R}^d)$, $F_n, F \in C(\mathbb{R}^d)$, $u_n \in BUC(\mathbb{R}^d \times [0, T])$ solves

$$(2.3) \qquad u_{n,t} = F_n(D^2u_n, Du_n) \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad u_n(\cdot, 0) = u_{0,n} \quad \text{in } \mathbb{R}^d,$$

and, as $n \to \infty$,

$$(2.4) \qquad u_{0,n} \to u_0 \quad \text{and} \quad F_n \to F \quad \text{locally uniformly,}$$

then, as $n \to \infty$, u_n converges locally uniformly to u, the viscosity solution of (2.1).

These and other properties can be summarized in terms of the solution operators for (2.1), which are, for $t \ge 0$, the maps $S(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ for which the solution u of (2.1) is given by $u(x, t) = S(t)u_0(x)$. For all $s, t \ge 0$, $\phi, \psi \in BUC(\mathbb{R}^d)$, and $k \in \mathbb{R}$, these satisfy

$$(2.5) \qquad \begin{cases} \text{(a)} & S(0)\phi = \phi, \\ \text{(b)} & S(t + s) = S(t)S(s), \\ \text{(c)} & S(t)(\phi + k) = S(t)\phi + k, \quad \text{and} \\ \text{(d)} & \sup_{\mathbb{R}^d}(S(t)\phi - S(t)\psi) \le \sup_{\mathbb{R}^d}(\phi - \psi). \end{cases}$$

Property (2.5)(c) implies that (2.5)(d) is equivalent to the monotonicity of $S(t)$. That is, if $\phi \le \psi$, then $S(t)\phi \le S(t)\psi$ for all $t \ge 0$.

The stability property above can be rephrased as saying that, if (2.4) holds and if $S_n(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ is the family of solution operators corresponding to (2.3), then, as $n \to \infty$, $S_n(t)u_{0,n}(x) \to S(t)u_0(x)$ locally uniformly. The philosophy behind the creation of approximation schemes is to generalize this result, by constructing, for $h > 0$ and $\rho > 0$, suitable operators $S_h(\rho) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ that satisfy properties similar to those in (2.5). In particular, for all $\phi \in BUC(\mathbb{R}^d)$ and $k \in \mathbb{R}$, and for some increasing function $h \mapsto \rho_h$ satisfying $\lim_{h \to 0} \rho_h = 0$,

$$(2.6) \qquad \begin{cases} \text{(a)} & S_h(t)(\phi + k) = S_h(t)\phi + k \quad \text{for all } h, t > 0, \\ \text{(b)} & \sup_{\mathbb{R}^d}(S_h(\rho)\phi - S_h(\rho)\psi) \le \sup_{\mathbb{R}^d}(\phi - \psi) \\ & \quad \text{whenever } h > 0 \text{ and } 0 < \rho \le \rho_h, \quad \text{and} \\ \text{(c)} & \lim_{h \to 0} \sup_{0 < \rho \le \rho_h} \left| \dfrac{S_h(\rho)\phi - \phi}{\rho} - F(D^2\phi, D\phi) \right| = 0 \quad \text{for all } \phi \in C^2(\mathbb{R}^d). \end{cases}$$

Given a partition $\mathcal{P}_h$ satisfying $|\mathcal{P}_h| \le \rho_h$, the approximate solution $u_h : BUC(\mathbb{R}^d \times [0, T])$ is assembled by first setting $u_h(\cdot, 0) := u_0$ and then iteratively defining

$$(2.7) \qquad u_h(\cdot, t) := S_h(t - t_n)u_h(\cdot, t_n) \quad \text{for } n = 0, 1, 2, \ldots, N - 1 \text{ and } t \in (t_n, t_{n+1}].$$

One example of particular interest is the class of finite difference approximations, for which $S_h(\rho)u$ depends on the function u only through its values on the discrete lattice $h\mathbb{Z}^d$. A major consideration for such schemes is to establish a relationship between the resolutions of the discrete grids in time and space, that is, to choose the map $h \mapsto \rho_h$ in such a way that the properties in (2.6) can be attained. Such a relationship is known as a Courant–Friedrichs–Lewy (CFL) condition [9], and various examples will be studied throughout the paper.

As is well known, solutions of (1.1) are generally not C^2 on all of $\mathbb{R}^d \times [0, T]$, even if F, H, and u_0 are all smooth, and so (2.6)(c) alone is not enough to prove the convergence of u_h to u as $h \to 0$. It is here that the monotonicity of $S_h(\rho)$, which is implied by (2.6)(a) and (b), is vital, since it allows the scheme operator to be applied to the smooth test functions coming from the definition of viscosity solutions.

A finite difference scheme operator S_h, in its simplest form, when $d = 1$ (the last assumption here made only to simplify the presentation), is given, for some $F_h \in C^{0,1}(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$, by

$$
(2.8) \quad
\begin{aligned}
S_h(\rho)u(x) := u(x) + \rho F_h\Big(&\frac{u(x+h) + u(x-h) - 2u(x)}{h^2}, \frac{u(x+h) - u(x)}{h}, \\
&\frac{u(x) - u(x-h)}{h}\Big).
\end{aligned}
$$

The scheme (2.8) automatically satisfies (2.6)(a), while (2.6)(b) holds if the function

$$
(u, u_-, u_+) \mapsto u + \rho F_h\Big(\frac{u_+ + u_- - 2u}{h^2}, \frac{u_+ - u}{h}, \frac{u - u_-}{h}\Big)
$$

is nondecreasing in each argument when $0 < \rho \leq \rho_h$, which, in turn, calls for

$$
(2.9) \qquad \rho_h := \lambda h^2
$$

for some sufficiently small constant $\lambda > 0$. In the case of first-order equations, that is, for the equation

$$
(2.10) \qquad u_t = H(Du) \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{on } \mathbb{R}^d,
$$

the CFL condition becomes

$$
(2.11) \qquad \rho_h = \lambda h.
$$

The function F_h is related to F through a consistency requirement, which here means that, for all $X \in \mathbb{R}$ and $p \in \mathbb{R}$,

$$
(2.12) \qquad \lim_{h \to 0} F_h(X, p, p) = F(X, p) \quad \text{and} \quad \sup_{h > 0} \|DF_h\|_\infty < \infty.
$$

Property (2.6)(c) can then be readily verified by using Taylor approximations to estimate the finite differences of functions $\phi \in C^2(\mathbb{R}^d)$.

An instructive example in the first-order setting is the following analogue of the Lax–Friedrichs scheme for scalar conservation laws. Let $\epsilon_h > 0$ and define, for $x \in \mathbb{R}$,

$$
(2.13) \quad
\begin{aligned}
S_h(\rho)u(x) := u(x) + \rho\Big\{&H\Big(\frac{u(x+h) - u(x-h)}{2h}\Big) \\
&+ \epsilon_h\Big(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2}\Big)\Big\}.
\end{aligned}
$$

Here, H_h is given by

$$
H_h(p, q) = H\Big(\frac{p+q}{2}\Big) + \frac{\epsilon_h}{h}(p - q).
$$

The final term in (2.13) is a discrete analogue of the method of vanishing viscosity, and is used here to inject monotonicity into the scheme. Indeed, if, for some fixed $\theta > 0$ and $\lambda > 0$, the small parameter ϵ_h is defined by

$$\epsilon_h := \frac{\theta h}{2\lambda}, \tag{2.14}$$

then (2.6)(b) is satisfied as long as (2.11) holds with $\theta \le 1$ and $\lambda \le \frac{\theta}{\|H'\|_\infty}$.

In [11], Crandall and Lions found explicit error estimates for this and and other explicit finite difference schemes for homogenous Hamilton–Jacobi equations. More precisely, it was proved for the above example that there exists a constant $C > 0$ depending only on $\|DH\|_\infty$, $\|Du_0\|_\infty$, and λ such that, if u_h is defined as in (2.7) and (2.13), and if u solves (2.10), then

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \left| u_h(x,t) - u(x,t) \right| \le C(1+T)h^{1/2}.$$

This same rate was later established by Souganidis [32] for both explicit and implicit finite difference schemes for equations with Lipschitz spatial and time dependence, and the same method was applied to study other approximations such as max-min representations and Trotter–Kato product formulas [33].

Barles and Souganidis [4] considered schemes for second order equations, using a shorter, qualitative proof of convergence relying on the method of half-relaxed limits. Kuo and Trudinger [20, 21] also investigated such schemes in great detail and constructed several examples. The question of estimating the rates of convergence for such approximations of second order equations was analyzed from many points of view. Barles and Jakobsen [1–3] achieved algebraic convergence rates for stochastic control problems, taking advantage of the fact that F is convex in that setting. Jakobsen [17, 18] and Krylov [19] also established rates of convergence for nonconvex problems under some restrictions on F. If F is uniformly elliptic, then rates of convergence can be found under very general assumptions using techniques from the regularity theory for fully nonlinear, uniformly elliptic equations, as exhibited by Caffarelli and Souganidis [7], and later by Turanova [35] for inhomogenous equations.

2.2. *Difficulties in the pathwise setting.* The lack of regularity for W complicates the task of constructing scheme operators for (1.1) that are both monotone and consistent.

Consider, for example, modifying the Lax–Friedrichs scheme (2.13) for the stochastic Hamilton–Jacobi equation

$$du = H(u_x) \circ dW \quad \text{in } \mathbb{R} \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}. \tag{2.15}$$

If W is sufficiently regular, then it is reasonable to define a time-inhomogenous scheme operator by

$$
\begin{aligned}
S_h(t, s)u(x) := u(x) + H\left(\frac{u(x+h) - u(x-h)}{2h}\right)(W(t) - W(s)) \\
+ \epsilon_h\left(\frac{u(x+h) + u(x-h) - 2u(x)}{h^2}\right)(t - s).
\end{aligned}
\tag{2.16}
$$

Proceeding as in the previous subsection, a simple calculation reveals that $S_h(t, s)$ is monotone for $0 \le t - s \le \rho_h$, if ρ_h and ϵ_h are such that, for some $\theta \in (0, 1]$,

$$\epsilon_h := \frac{\theta h^2}{2(t - s)}$$

and

$$\lambda := \max_{|t-s|\le\rho_h} \frac{\text{osc}(W, s, t)}{h} \le \lambda_0 := \frac{\theta}{\|H'\|_\infty}. \tag{2.17}$$

On the other hand, spatially smooth solutions Φ of (2.15) have the expansion, for any $s, t \in [0, T]$ with $|s - t|$ sufficiently small,

$$
\begin{aligned}
(2.18) \quad \Phi(x, t) &= \Phi(x, s) + H\big(\Phi_x(x, s)\big)\big(W(t) - W(s)\big) \\
&\quad + H'\big(\Phi_x(x, s)\big)^2 \Phi_{xx}(x, s)\big(W(t) - W(s)\big)^2 + O\big(|W(t) - W(s)|^3\big),
\end{aligned}
$$

so that, if $0 \le t - s \le \rho_h$, for some $C > 0$ depending only on H,

$$
\begin{aligned}
(2.19) \quad \sup_{\mathbb{R}} \big| S_h(t, s)\Phi(\cdot, s) - \Phi(\cdot, t)\big| &\le C \sup_{r \in [s,t]} \big\| D^2 \Phi(\cdot, r)\big\|_\infty \big(|W(t) - W(s)|^2 + h^2\big) \\
&\le C \sup_{r \in [s,t]} \big\| D^2 \Phi(\cdot, r)\big\|_\infty (1 + \lambda_0^2) h^2.
\end{aligned}
$$

Therefore, in order for the scheme to have a chance of converging, ρ_h should satisfy

$$
(2.20) \qquad \lim_{h \to 0} \frac{h^2}{\rho_h} = 0.
$$

Both (2.17) and (2.20) can be achieved when W is continuously differentiable, or merely Lipschitz, by setting

$$
\rho_h := \frac{\lambda h}{\|\dot{W}\|_\infty}.
$$

More generally, if W has Young–Hölder regularity, that is, $W \in C^\alpha([0, T])$ with $\alpha > \frac{1}{2}$, and if

$$
(2.21) \qquad (\rho_h)^\alpha := \frac{\lambda h}{[W]_{\alpha, T}},
$$

then both (2.17) and (2.20) are satisfied, since

$$
\frac{h^2}{\rho_h} = \left(\frac{[W]_{\alpha, T} h^{2\alpha - 1}}{\lambda} \right)^{1/\alpha} \xrightarrow{h \to 0} 0.
$$

However, this approach fails as soon as the quadratic variation

$$
\mathcal{Q}([0, T], W) := \lim_{|\mathcal{P}| \to 0} \sum_{n=0}^{N-1} |W(t_{n+1}) - W(t_n)|^2
$$

is nonzero, as (2.17) and (2.20) together imply that $\mathcal{Q}([0, T], W) = 0$. This rules out, for instance, the case where W is the sample path of a Brownian motion, for which $\mathcal{Q}([0, T], W) = T$ with probability one.

Motivated by the theory of rough differential equations, it is natural to explore whether the scheme operator (2.16) can be altered in some way to refine the estimate in (2.19), potentially allowing (2.20) to be relaxed and ρ_h to converge more quickly to zero as $h \to 0^+$. More precisely, the next term in the expansion (2.18) suggests that, for $W \in C^\alpha([0, T], \mathbb{R})$ with $\alpha > \frac{1}{3}$ (or more generally, W with p-variation for $p < 3$), one should define

$$
\begin{aligned}
(2.22) \quad S_h&(t, s)u(x) \\
&:= u(x) + H\left(\frac{u(x + h) - u(x - h)}{2h} \right)\big(W(t) - W(s)\big) \\
&\quad + \frac{1}{2} H'\left(\frac{u(x + h) - u(x - h)}{2h} \right)^2 \left(\frac{u(x + h) + u(x - h) - 2u(x)}{h^2} \right) \\
&\quad \times \big(W(t) - W(s)\big)^2 \\
&\quad + \frac{\theta}{2}\big(u(x + h) + u(x - h) - 2u(x)\big).
\end{aligned}
$$

As can easily be checked, (2.22) is monotone as long as (2.17) holds,

$$\|Du\|_\infty \le L, \qquad \theta + \|H'\|_\infty \lambda^2 \le 1 \quad \text{and} \quad \lambda \le \frac{\theta}{\|H'\|_\infty (1 + 2L\|H''\|_\infty)}.$$

On the other hand, the error in (2.19) would then be of order $h^2 + |W(t) - W(s)|^3$, which again leads to (2.20). This seems to indicate that we should also incorporate higher order corrections in (2.22) to improve the order of the error in the h variable. However, this will disrupt the monotonicity of the scheme in general. This is due to the fact that such terms involve the discrete second derivative of u, and, thus, will counter the effect of the term

$$\frac{\theta}{2}(u(x+h) + u(x-h) - 2u(x)),$$

which is included precisely for the purpose of creating monotonicity.

For this reason, we develop a more effective strategy that works for any continuous path. Namely, rather than modifying the scheme itself, we regularize the path W. If $\{W_h\}_{h>0}$ is a family of smooth paths converging uniformly, as $h \to 0$, to W, then $\mathcal{Q}(W_h, [0, T]) = 0$ for each fixed $h > 0$, and therefore, W_h and ρ_h can be chosen so that (2.17) and (2.20) hold for W_h rather than W. Various methods for implementing this procedure, both qualitative and quantitative, are explored throughout the paper.

3. The definition of pathwise viscosity solutions.

3.1. *Assumptions on the nonlinearities.* The nonlinear function $F : S^d \times \mathbb{R}^d \to \mathbb{R}$ is assumed to be Lipschitz and degenerate elliptic; that is,

$$(3.1) \qquad \begin{cases} F \in C^{0,1}(S^d \times \mathbb{R}^d) & \text{and} \\ F(X, p) \le F(Y, p) & \text{whenever } p \in \mathbb{R}^d \text{ and } X \le Y. \end{cases}$$

The results of this paper may be extended to the case where F has additional dependence on u, x, or t, in which case F requires additional structure in order for the comparison principle to hold. To simplify the presentation, we take F as in (3.1). One consequence is that the solution operator for (1.1) is invariant under translations in both the independent and dependent variables.

In order for (1.1) to be well-posed for all continuous W and uniformly continuous u_0, the Hamiltonians need to be more regular than what is required in the classical viscosity theory. As explained in [25] and [31], it is necessary to assume that

$$(3.2) \qquad H^i = H_1^i - H_2^i \quad \text{for convex } H_1^i, H_2^i : \mathbb{R}^d \to \mathbb{R} \text{ with } H_1^i, H_2^i \ge 0.$$

The nonnegativity is imposed here only to simplify some arguments in what follows, and the setting can be reduced to the general case by transforming the equation appropriately.

Letting H depend additionally on u or x makes the question of well-posedness for (1.1) highly nontrivial. Indeed, there is no pathwise theory for equations of the form

$$du = F(D^2u, Du)\,dt + \sum_{i=1}^m H^i(Du, u, x) \circ dW^i,$$

except for some special cases, for instance, if the dependence of H on Du is linear.

Under certain assumptions, (1.1) is well-posed for H depending nonlinearly on both Du and x. In this case, the lack of uniform regularity estimates for the solutions becomes an obstacle in the construction of schemes for (1.1). These issues will be the subject of a future work.

The homogeneity of H in space allows us to forego difficult questions about regularity, because the spatial modulus of continuity for the solution of (1.1) is retained for all time. In particular, throughout much of the paper, the initial condition u_0 is fixed and satisfies

$$(3.3) \qquad u_0 \in C^{0,1}(\mathbb{R}^d) \quad \text{and} \quad \|Du_0\|_\infty \leq L,$$

and therefore, $H(p)$ may be redefined for $|p| > L$ without affecting the solution. Since (3.2) implies that H is locally Lipschitz, we may then assume that

$$(3.4) \qquad \text{for some } C = C_L > 0, \quad \|DH\|_\infty = C_L < \infty.$$

Note that (3.4) implies that H grows at most linearly as $|p| \to +\infty$.

In some parts of the paper, to allow for a more flexible solution theory, especially when we discuss schemes for second order equations, the hypothesis (3.2) is replaced with the stronger assumption

$$(3.5) \qquad H \in C^k(\mathbb{R}^d, \mathbb{R}^m) \quad \text{for some } k = 2, 3, \ldots.$$

If H satisfies (3.5), and therefore $H \in C^{1,1}(\mathbb{R}^d, \mathbb{R}^m)$, then (3.2) is satisfied on every bounded set of gradients.

3.2. Smooth solutions of the Hamilton–Jacobi part of (1.1).

The definition of pathwise viscosity solutions relies on the existence of local-in-time, smooth-in-space solutions of the Hamilton–Jacobi part of equation (1.1). More precisely, for $t_0 \in [0, T]$ and $\phi \in C^{1,1}(\mathbb{R}^d)$, the goal is to find an open interval $I \subset [0, T]$ containing t_0 and a solution $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ of

$$(3.6) \qquad d\Phi = \sum_{i=1}^m H^i(D\Phi) \circ dW^i \quad \text{in } \mathbb{R}^d \times I \quad \text{and} \quad \Phi(\cdot, t_0) = \phi \quad \text{in } \mathbb{R}^d.$$

Such solutions are defined through a density argument, that is, the solution operator for (3.6) for smooth paths extends continuously to continuous paths. This is justified with the computations below, and is consistent with the cases where W is a Brownian motion or, more generally, a geometric rough path.

When H and ϕ are smooth, the construction of such solutions can be accomplished for any smooth ϕ by inverting the characteristics associated to (3.6). Because H is independent of x, this amounts to inverting the map

$$(3.7) \qquad x \mapsto X(x, t) := x - \sum_{i=1}^m DH^i(D\phi(x))(W^i(t) - W^i(t_0)).$$

The continuity of W implies that there exists an interval $I \ni t_0$ such that

$$\sup_{t \in I} |W(t) - W(t_0)| < \frac{1}{\|D^2 H\|_\infty \|D^2 \phi\|_\infty},$$

whence (3.7) is invertible for all $t \in I$. The solution is then given by $\Phi(x, t) := Z(X^{-1}(x, t), t)$, where

$$(3.8) \qquad Z(x, t) := \phi(x) + \sum_{i=1}^m \left(H^i(D\phi(x)) - D\phi(x) \cdot DH^i(D\phi(x))\right)(W^i(t) - W^i(t_0)).$$

This can be confirmed with a simple calculation when W is smooth. For general continuous paths, the formula holds by a density argument, since all expressions depend only on the values of W, and not on its derivatives.

Notice also that the regularity of Φ improves with that of H and ϕ. Indeed, differentiating (3.8) leads to the relation $D\Phi(X(x, t), t) = D\phi(x, t)$, and therefore, in view of (3.7), the

solution Φ belongs to $C(I, C^k(\mathbb{R}^d))$ as long as $H \in C^k(\mathbb{R}^d, \mathbb{R}^m)$ and $\phi \in C^k(\mathbb{R}^d)$ for some $k = 2, 3, \ldots$. Furthermore, if $D^j \phi$ is bounded for some $j = 0, 1, 2, \ldots, k$, then, shrinking I if necessary,

$$\sup_{t \in I} \| D^j \Phi(\cdot, t) \|_\infty < \infty.$$

This strategy breaks down when H is only assumed to satisfy (3.2), since such Hamiltonians are not even continuously differentiable in general. In this case, only very particular smooth solutions of (3.6) can be constructed. Assume $\eta : \mathbb{R}^d \to \mathbb{R}$ is strictly convex, and, for $\delta > 0$, define

$$(3.9) \quad \Phi(x, t) := \sup_{p \in \mathbb{R}^d} \left\{ p \cdot x - \eta(p) - \delta \sum_{i=1}^m (H_1^i(p) + H_2^i(p)) + \sum_{i=1}^m H^i(p)(W^i(t) - W^i(t_0)) \right\}.$$

LEMMA 3.1.   *Let H satisfy (3.2) and (3.4). If the open interval $I \ni t_0$ is such that*

$$\sup_{t \in I} \max_{i=1,2,\ldots,m} |W^i(t) - W^i(t_0)| < \delta,$$

then the function Φ defined by (3.9) belongs to $C(I, C^{1,1}(\mathbb{R}^d))$, and is a solution of (3.6) with

$$(3.10) \qquad \Phi(\cdot, t_0) = \phi(x) := \sup_{p \in \mathbb{R}^d} \left\{ p \cdot x - \eta(p) - \delta \sum_{i=1}^m (H_1^i(p) + H_2^i(p)) \right\}.$$

PROOF.   For all $x \in \mathbb{R}^d$ and $t \in I$, the function

$$p \mapsto \eta(p) + \delta \sum_{i=1}^m (H_1^i(p) + H_2^i(p)) - \sum_{i=1}^m H^i(p)(W^i(t) - W^i(t_0)) - p \cdot x$$

is strictly convex, and therefore attains a unique global minimum. The smoothness of Φ in x then follows from the implicit function theorem.

Now, for $t \in \mathbb{R}$, let $S^i(t) : UC(\mathbb{R}^d) \to UC(\mathbb{R}^d)$ be the solution operator for the equation $u_t = H^i(Du)$. If $\psi \in UC(\mathbb{R}^d)$ is convex, then the Hopf formula gives

$$S^i(t)\psi(x) = \sup_{p \in \mathbb{R}^d} \{ p \cdot x - \psi^*(p) + t H^i(p) \},$$

and so (3.9) can be rewritten as

$$\Phi(x, t) = \prod_{i=1}^m S^i(W^i(t) - W^i(t_0))\phi(x)$$

with ϕ as in (3.10). If W is smooth, then the fact that Φ is a solution of (3.6) is justified by the regularity of Φ and a simple calculation. The result holds for continuous W by a density argument. $\square$

As in the classical viscosity theory, many quantitative arguments involve doubling variables, and it is therefore important to have a smooth solution of (3.6) that behaves like the penalizing "distance function"

$$(3.11) \qquad\qquad\qquad\qquad (x, y) \mapsto \frac{|x - y|^2}{2\delta}.$$

In the present setting, this is accomplished with a function $\Phi_\delta : \mathbb{R}^d \times [0, T]^2 \times C([0, T], \mathbb{R}^m) \to \mathbb{R}$ that is equal to a particular choice of (3.9) near the diagonal $\{(t, t) \in [0, T]^2\}$, and such that $\Phi_\delta(x - y, s, t; W)$ exhibits similar growth as (3.11) when $|x - y|$ is large.

Define the neighborhood $U_\delta(W)$ by

$$U_\delta(W) := \{(s, t) \in [0, T]^2 : \operatorname{osc}(W, s, t) < \delta\},$$

let the projection $\pi_\delta(W) : [0, T]^2 \to \overline{U_\delta(W)}$ be such that $\pi_\delta(W)(s, t)$ is the element $(\tilde{s}, \tilde{t}) \in \overline{U_\delta(W)}$ closest to (s, t) on the line $\tilde{s} + \tilde{t} = s + t$, and set

(3.12)
$$\Phi_\delta(x, s, t; W)$$
$$:= \begin{cases} \sup_{p \in \mathbb{R}^d} \left\{ p \cdot x - \dfrac{\delta}{2}|p|^2 - \delta \sum_{i=1}^m (H_1^i(p) + H_2^i(p)) \right. \\ \qquad \left. + \sum_{i=1}^m H^i(p)(W^i(s) - W^i(t)) \right\} & \text{if } (s, t) \in \overline{U_\delta(W)}, \\ \Phi_\delta(x, \pi_\delta(W)(s, t); W) & \text{if } (s, t) \notin \overline{U_\delta(W)}. \end{cases}$$

LEMMA 3.2.   *Assume H satisfies* (3.2) *and* (3.4), *and let Φ_δ be defined as in* (3.12). *For some $C = C_L > 0$ and for all $\delta > 0$ and $W \in C([0, T]; \mathbb{R}^m)$, the following hold:*

(a) *For all $x \in \mathbb{R}^d$ and $(s, t), (\tilde{s}, t) \in U_\delta(W)$,*

$$\left| \Phi_\delta(x, s, t; W) - \Phi_\delta(x, \tilde{s}, t; W) \right| \le C\left(1 + \frac{|x|}{\delta}\right)|W(s) - W(\tilde{s})|.$$

(b) *For all $(s, t) \in [0, T]^2$, $\Phi_\delta(\cdot, s, t; W)$ is convex and semiconcave with constant $\frac{1}{\delta}$. That is,*

$$0 \le D^2 \Phi_\delta(x, s, t; W) \le \frac{1}{\delta} I_d \quad \text{in the sense of distributions.}$$

(c) *For all $x \in \mathbb{R}^d$ and $s, t \in [0, T]$,*

$$\frac{1}{2(C+1)\delta}|x|^2 - C\delta \le \Phi_\delta(x, s, t; W) \le \frac{1}{2\delta}|x|^2.$$

(d) *For any fixed $y \in \mathbb{R}^d$ and $t \in [0, T]$, the functions*

$$(x, s) \mapsto \Phi_\delta(x - y, s, t; W) \quad \text{and} \quad (x, s) \mapsto -\Phi_\delta(y - x, t, s; W)$$

are $C(I, C^{1,1}(\mathbb{R}^d))$-solutions of (3.6), *where $I := \{s \in [0, T] : \operatorname{osc}(W, s, t) < \delta\}$.*

Note that the local regularity given by (a) also applies to the second time variable, in view of the identity $\Phi_\delta(x, s, t; W) = \Phi_\delta(x, t, s; -W)$.

PROOF OF LEMMA 3.2.   To prove (a), we first show that there exists $C = C_L > 0$ such that, for any $x \in \mathbb{R}^d$ and $(s, t) \in U_\delta(W)$, the unique maximum p^* achieved in the definition of Φ_δ satisfies $\delta|p^*| \le C\delta + |x|$. Indeed, if

$$J(p) := p \cdot x - \frac{\delta}{2}|p|^2 - \delta \sum_{i=1}^m (H_1^i(p) + H_2^i(p)) + \sum_{i=1}^m H^i(p)(W^i(s) - W^i(t)),$$

then, for any $q \in \mathbb{R}^d$, (3.4) and the inequality $J(p^*) \ge J(p^* + q)$ imply that

$$\delta p^* \cdot \frac{q}{|q|} - \frac{\delta}{2}|q| \le |x| + C\delta.$$

Setting $q = t \frac{p^*}{|p^*|}$ and sending $t \to 0^+$ yields the claim. The time-regularity estimate in (a) is then immediate.

As a pointwise supremum of affine functions, Φ_δ is clearly convex, while the semiconcavity follows from elementary convex analysis and the convexity of

$$p \mapsto \delta \sum_{i=1}^{m} (H_1^i(p) + H_2^i(p)) - \sum_{i=1}^{m} H^i(p)(W^i(s) - W^i(t)).$$

The estimate in (c) can be deduced from Young's inequality and the fact that, for some $C = C_L > 0$ and for all $p \in \mathbb{R}^d$ and $(s, t) \in U_\delta(W)$,

$$0 \le \delta \sum_{i=1}^{m} (H_1^i(p) + H_2^i(p)) - \sum_{i=1}^{m} H^i(p)(W^i(s) - W^i(t)) \le C\delta(1 + |p|).$$

Finally, (d) is a consequence of Lemma 3.1. $\square$

The following definition for solutions of (1.1) relies on the existence of solutions of (3.6) that are C^2, and, in particular, is only valid if H is at least twice continuously differentiable.

DEFINITION 3.1.   A function $u \in USC(\mathbb{R}^d \times [0, T])$ (resp. $u \in LSC(\mathbb{R}^d \times [0, T])$) is called a pathwise viscosity sub-solution (resp. super-solution) of (1.1) for H satisfying (3.5) if $u(\cdot, 0) \le u_0$ (resp. $u(\cdot, 0) \ge u_0$) and, whenever $\psi \in C^1([0, T])$, $(x_0, t_0) \in \mathbb{R}^d \times [0, T]$, $I \ni t_0$, $\Phi \in C(I, C^2(\mathbb{R}^d))$ is a solution of (3.6) in $\mathbb{R}^d \times I$, and

$$u(x, t) - \Phi(x, t) - \psi(t)$$

attains a local maximum (resp. minimum) at $(x_0, t_0) \in \mathbb{R}^d \times I$, then

$$\text{(3.13)} \qquad \begin{aligned} \psi'(t_0) &\le F(D^2\Phi(x_0, t_0), D\Phi(x_0, t_0)) \\ &(\text{resp. } \psi'(t_0) \ge F(D^2\Phi(x_0, t_0), D\Phi(x_0, t_0))). \end{aligned}$$

A solution of (1.1) is both a sub- and super-solution.

If $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ is a solution of (3.6), then it is not possible to make sense of (3.13), since $D^2\Phi$ may not be defined at every point. The following definition is made to comply with the case when H only satisfies (3.2).

DEFINITION 3.2.   A function $u \in USC(\mathbb{R}^d \times [0, T])$ (resp. $u \in LSC(\mathbb{R}^d \times [0, T])$) is called a pathwise viscosity sub-solution (resp. super-solution) of (1.1) if $u(\cdot, 0) \le u_0$ (resp. $u(\cdot, 0) \ge u_0$) and, whenever $I \subset [0, T]$ and $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ is a solution of (3.6) in $\mathbb{R}^d \times I$, the function $v : \mathbb{R}^d \times I \to \mathbb{R}$ defined by

$$v(\xi, t) := \sup_{x \in \mathbb{R}^d} \{u(x, t) - \Phi(x - \xi, t)\} \quad \left(\text{resp. } v(\xi, t) := \inf_{x \in \mathbb{R}^d} \{u(x, t) + \Phi(x - \xi, t)\}\right)$$

is a classical viscosity sub- (resp. super-) solution of the equation

$$v_t = F(D^2 v, Dv) \quad \text{in } \mathbb{R}^d \times I.$$

A solution of (1.1) is both a sub- and super-solution.

When H satisfies (3.5), Definition 3.2 is equivalent to Definition 3.1. In the first-order setting, that is, when $F \equiv 0$, Definition 3.1 may be used even if H is not smooth, because it is not necessary to evaluate $D^2\Phi$ at any point.

With either definition, (1.1) satisfies the following comparison principle, a proof for which can be found in [26] or [31]: if $u \in USC(\mathbb{R}^d \times [0, T])$ and $v \in LSC(\mathbb{R}^d \times [0, T])$ are respectively a sub- and super-solution of (1.1), then, for all $t \in (0, T]$,

$$\text{(3.14)} \qquad \sup_{x \in \mathbb{R}^d} (u(x, t) - v(x, t)) \le \sup_{x \in \mathbb{R}^n} (u(x, 0) - v(x, 0)).$$

A variant of the proof of the comparison principle gives the following path-stability estimate [31].

LEMMA 3.3. *Assume that H satisfies* (3.2). *There exists $C = C_L > 0$ such that, if $u_0 \in C^{0,1}(\mathbb{R}^d)$ with $\|Du_0\|_\infty \leq L$, $W^1, W^2 \in C([0, T], \mathbb{R}^m)$, and $u^1, u^2 \in C([0, T], C^{0,1}(\mathbb{R}^d))$ are the solutions of* (1.1) *with respectively the paths W^1 and W^2, then*

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u^1(x,t) - u^2(x,t)| \leq C \max_{t\in[0,T]} |W^1(t) - W^2(t)|.$$

It can be shown that, when $W \in C^1([0, T])$, the above notions of pathwise viscosity solutions are consistent with the standard definitions from the classical viscosity theory. Furthermore, solutions of (1.1) are stable under uniform convergence. Therefore, the estimate in Lemma 3.3, which is proved first for smooth paths, also establishes the existence of pathwise viscosity solutions. Lemma 3.3 can then be seen to hold for arbitrary continuous paths via a density argument.

Although the class of test functions defined by (3.9) is rather restrictive, it is enough to prove both the comparison principle and the stability estimate. Indeed, only the "distance function" Φ_δ in (3.12) is used in both proofs. When $H \in C^2(\mathbb{R}^d)$, any initial condition $\Phi(\cdot, t_0) \in C^2(\mathbb{R}^d)$ with bounded second derivatives yields a solution as in (3.6). In particular, by adding quadratic functions to $\Phi(\cdot, t_0)$, it may be assumed that the test functions in Definition 3.1 satisfy

$$\lim_{|x|\to+\infty} \frac{\Phi(x,t)}{|x|} = +\infty \quad \text{or} \quad \lim_{|x|\to+\infty} \frac{\Phi(x,t)}{|x|} = -\infty \quad \text{uniformly for } t \in I.$$

Thus, as in the classical viscosity theory, the maxima and minima in Definition 3.1 may be assumed to be strict without loss of generality.

Finally, we remark that, if H satisfies (3.5), then it is enough to use functions $\Phi \in C(I, C^k(\mathbb{R}^d))$ in Definition 3.1. The argument is almost identical to one from the classical viscosity theory, and it uses the fact that the solution operator for (3.6) is contractive.

4. The general convergence result and applications. The constructions in this paper rely on the properties of a family of scheme operators, indexed by $h > 0$, $s, t \in [0, T]$ with $s \leq t$, and a path $\zeta \in C([0, T], \mathbb{R}^m)$:

$$S_h(t, s; \zeta) : (B)UC(\mathbb{R}^d) \to (B)UC(\mathbb{R}^d).$$

We assume throughout that S_h commutes with translations in both the independent and dependent variables, in order to reflect the corresponding translation invariance of (1.1). That is,

$$(4.1) \qquad S_h(t, s; \zeta)(u + k) = S_h(t, s; \zeta)u + k \quad \text{for all } k \in \mathbb{R} \text{ and } u \in (B)UC(\mathbb{R}^d),$$

and

$$(4.2) \qquad S_h(t, s; \zeta) \circ \tau_v = \tau_v \circ S_h(t, s; \zeta) \quad \text{for all } v \in \mathbb{R}^d, \text{ where } \tau_v u := u(\cdot + v).$$

For a Hamiltonian H satisfying (3.5) and a fixed continuous path $W \in C([0, T], \mathbb{R}^m)$, we consider a family of paths $\{W_h\}_{h>0} \subset C([0, T], \mathbb{R}^m)$ and a partition width $\rho_h > 0$ satisfying

$$(4.3) \qquad h \mapsto \rho_h \text{ is increasing}, \quad \lim_{h\to 0} \|W_h - W\|_\infty = 0 = \lim_{h\to 0} \rho_h;$$

$$(4.4) \qquad \begin{array}{l} \text{if } u_1 \leq u_2 \text{ and } s, t \in [0, T] \text{ satisfy } 0 \leq t - s \leq \rho_h, \\ \text{then } S_h(t, s; W_h)u_1 \leq S_h(t, s; W_h)u_2; \end{array}$$

and

$$
(4.5) \quad
\begin{cases}
\text{if } I \subset \mathbb{R}, \Phi_h \in C\big(I, C^k(\mathbb{R}^d)\big) \text{ is a solution of} \\
\quad d\Phi_h = \sum_{i=1}^{m} H^i(D\Phi_h) \circ dW_h^i \text{ in } \mathbb{R}^d \times I, \\
s_h, t_h \in I, 0 \le t_h - s_h \le \rho_h, \phi \in C^k(\mathbb{R}^d), R > 0, \text{ and} \\
\quad \lim_{h \to 0} \big\| \Phi_h(\cdot, s_h) - \phi \big\|_{C^k(\mathbb{R}^d)} = 0, \\
\text{then } \lim_{h \to 0} \dfrac{S_h(t_h, s_h; W_h)\Phi_h(\cdot, s_h)(x) - \Phi_h(x, s_h)}{t_h - s_h} = F(D^2\phi(x), D\phi(x)) \\
\text{uniformly for } x \in \mathbb{R}^d \text{ and } \max_{j=2,3,\dots,k} \big\| D^j\phi \big\|_\infty \le R.
\end{cases}
$$

The integer k in (4.5) corresponds to the level of regularity of H in (3.5). In Sections 5 and 6, we obtain error estimates for schemes for first-order equations with Hamiltonians satisfying the weaker condition (3.2), in which case the assumptions on the scheme operator will be modified.

The scheme operator is used to build approximate solutions as follows. For a fixed path $\zeta \in C([0, T]; \mathbb{R}^m)$, partition $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ of $[0, T]$, and initial datum $u_0 \in BUC(\mathbb{R}^d)$, define

$$
(4.6) \quad
\begin{cases}
v_h(\cdot, 0; \zeta, \mathcal{P}) := u_0, \\
v_h(\cdot, t; \zeta, \mathcal{P}) := S_h(t, t_n; \zeta) v_h(\cdot, t_n; \zeta, \mathcal{P}) \\
\quad \text{for } n = 0, 1, \dots, N - 1 \text{ and } t \in (t_n, t_{n+1}].
\end{cases}
$$

THEOREM 4.1. *Assume $u_0 \in BUC(\mathbb{R}^d)$, (3.1), (3.5), and S_h, W_h, and ρ_h satisfy (4.1)–(4.5). Let $\{\mathcal{P}_h\}_{h>0}$ be a family of partitions of $[0, T]$ such that $|\mathcal{P}_h| \le \rho_h$ for all $h > 0$, and define $u_h := v_h(\cdot; W_h, \mathcal{P}_h)$. Then, as $h \to 0$, u_h converges locally uniformly to the pathwise viscosity solution u of (1.1).*

The proof of Theorem 4.1, which, as in [4], makes use of the method of half-relaxed limits, will be postponed until the end of this section. In the following sub-sections, we demonstrate its utility in a variety of contexts.

4.1. *Finite difference schemes.* Define, for $x \in \mathbb{R}^d$ and $y \in \mathbb{Z}^d \backslash \{0\}$, the discrete derivatives

$$
(4.7) \quad
\begin{aligned}
& D_{h,y}^+ u(x) := \frac{u(x + hy) - u(x)}{h|y|}, \qquad D_{h,y}^- u(x) := \frac{u(x) - u(x - hy)}{h|y|} \quad \text{and} \\
& D_{h,y}^2 u(x) := D_{h,y}^+ D_{h,y}^- u(x) = \frac{u(x + hy) + u(x - hy) - 2u(x)}{h^2|y|^2}.
\end{aligned}
$$

Observe that there exists a universal constant $C > 0$ such that, if $u \in C^{1,1}(\mathbb{R}^d)$, $h > 0$, and $y \in \mathbb{Z}^d \backslash \{0\}$, then

$$
(4.8) \quad \left\| D_{h,y}^\pm u - Du \cdot \frac{y}{|y|} \right\|_\infty \le C \| D^2 u \|_\infty h,
$$

and, if $u \in C^2(\mathbb{R}^d)$,

$$
(4.9) \quad \left\| D_{h,y}^2 u - D^2 u \frac{y}{|y|} \cdot \frac{y}{|y|} \right\|_\infty \le C \sup_{|x_1 - x_2| \le h} |D^2 u(x_1) - D^2 u(x_2)|.
$$

For some fixed $N \in \mathbb{N}$, define

$$\mathbb{Z}_N^d := \left\{ y \in \mathbb{Z}^d : \max_{i=1,2,\ldots,d} |y_i| \leq N \right\}, \qquad D_{h,N}^{\pm} := \{D_{h,y}^{\pm}\}_{y \in \mathbb{Z}_N^d \setminus \{0\}},$$

$$D_{h,N} := (D_{h,N}^+ D_{h,N}^-) \quad \text{and}$$

$$D_{h,N}^2 := \{D_{h,y}^2\}_{y \in \mathbb{Z}_N^d \setminus \{0\}}.$$

Then, for some given functions

$$H_h \in C^{0,1}\big(\mathbb{R}^{(2N+1)^d-1} \times \mathbb{R}^{(2N+1)^d-1} \times \mathbb{R}\big) \quad \text{and}$$

$$F_h \in C^{0,1}\big(\mathbb{R}^{2N+1)^d-1} \times \mathbb{R}^{(2N+1)^d-1} \times \mathbb{R}^{(2N+1)^d-1}\big),$$

the scheme operators for finite difference approximations take the form

$$(4.10) \quad S_h(t,s;\zeta)u(x) := u(x) + F_h\big(D_h^2 u(x), D_h u(x)\big)(t-s) + H_h\big(D_h u(x), \zeta(t) - \zeta(s)\big).$$

Properties (4.1) and (4.2) are immediate, while the question of whether (4.10) satisfies (4.4) or (4.5) is reduced to routine calculations involving F_h and H_h.

4.1.1. *Hamilton–Jacobi equations.*

We first study the first-order setting, for which $F = F_h = 0$, and assume, in addition to (3.5), that

$$(4.11) \quad \begin{cases} D_{p,q} H_h(\cdot, \cdot, \Delta\zeta) \leq C(|\Delta\zeta| + h) \\ \quad \text{for some } C = C_L > 0 \text{ and all } h > 0 \text{ and } \Delta\zeta \in \mathbb{R}^m, \\ \text{and } H_h(p, p, \Delta\zeta) = \sum_{i=1}^m H^i(p)(\Delta\zeta)^i \\ \quad \text{for all } h > 0, \ p \in \mathbb{R}^{(2N+1)^d-1}, \text{ and } \Delta\zeta \in \mathbb{R}^m. \end{cases}$$

In order for monotonicity to hold, the Lipschitz bounds in (4.11) are made more precise. Let elements of $\mathbb{R}^{(2N+1)^d-1}$ be labeled by $\{p_y\}_{y \in \mathbb{Z}_N^d \setminus \{0\}}$, and assume that, for some $C = C_L > 0$, $\theta \in [0,1]$, and $\lambda_0 > 0$,

$$(4.12) \quad \begin{cases} \displaystyle\sum_{y \in \mathbb{Z}_N^d \setminus \{0\}} \frac{1}{|y|}\left(\frac{\partial H_h}{\partial q_y} - \frac{\partial H_h}{\partial p_y}\right) \leq \frac{1-\theta}{\lambda_0}|\Delta\zeta| + \theta h \quad \text{and} \\ \displaystyle\frac{\partial H_h}{\partial q_y} - \frac{\partial H_h}{\partial p_{-y}} \geq C\left(h - \frac{|\Delta\zeta|}{\lambda_0}\right) \quad \text{for all } y \in \mathbb{Z}_N^d \setminus \{0\}. \end{cases}$$

LEMMA 4.1. *Suppose that H satisfies (3.5) and H_h satisfies (4.11). Then there exists $C = C_L > 0$ such that, whenever $\zeta \in C([0,T], \mathbb{R}^m)$, $\mathrm{osc}(\zeta, s, t) \leq \lambda_0 h$ for some $s, t \in I$, and $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ is a solution of*

$$d\Phi = \sum_{i=1}^m H^i(D\Phi) \circ d\zeta^i \quad \text{in } \mathbb{R}^d \times I,$$

then

$$\big\| S_h(t,s;\zeta)\Phi(\cdot,s) - \Phi(\cdot,t) \big\|_\infty \leq C \|D^2\Phi\|_\infty h^2.$$

If, in addition, H_h satisfies (4.12), then, whenever $u_1, u_2 \in (B)UC(\mathbb{R}^d)$ with $u_1 \leq u_2$ and $\mathrm{osc}(\zeta, s, t) \leq \lambda_0 h$,

$$S_h(t,s;\zeta)u_1 \leq S_h(t,s;\zeta)u_2.$$

Motivated by the above result, the schemes for first-order equations in Sections 5 and 6, for which we obtain explicit error estimates, will be assumed to satisfy the conclusions of Lemma 4.1. In fact, the smoothness assumption (3.5) is not needed in the proof of Lemma 4.1, and the quantitative convergence results in those sections can be proved under the more general hypotheses (3.2) and (3.4).

PROOF OF LEMMA 4.1. Let $\Phi \in C(I, C^{1,1}(\mathbb{R}^d))$ be as in the statement of the lemma. Then there exists $C > 0$ depending only on $\max_{|p| \leq L} |DH(p)|$ such that, for all $s, t \in I$,

$$\left\| \Phi(\cdot, t) - \Phi(\cdot, s) - \sum_{i=1}^m H^i(D\Phi(\cdot, s))(\zeta^i(t) - \zeta^i(s)) \right\|_\infty \leq C \|D^2\Phi\|_\infty |\zeta(t) - \zeta(s)|^2.$$

Therefore,

$$|S_h(t, s; \zeta)\Phi(\cdot, s)(x) - \Phi(x, t)| \leq C \|D^2\Phi\|_\infty (h^2 + |\zeta(t) - \zeta(s)|h + |\zeta(t) - \zeta(s)|^2)$$

$$\leq C(1 + \lambda_0 + \lambda_0^2)h^2.$$

Meanwhile, if $\mathcal{S}_h : \mathbb{R}^{(2N+1)^d} \to \mathbb{R}$ is the map implicitly defined by

$$\mathcal{S}_h(\{u(x + y)\}_{y \in \mathbb{Z}_N^d}) = S_h(t, s; \zeta)u(x),$$

then (4.12) implies that $\mathcal{S}_h$ is increasing in each of its arguments as long as $\mathrm{osc}(\zeta, s, t) \leq \lambda_0 h$. $\qquad \square$

We now mention two specific examples. The first is the analogue of the Lax–Friedrichs scheme for scalar conservation laws discussed in the Introduction. Here, H_h is defined, for some $\theta \in (0, 1]$, by

$$H_h(p, q, \Delta\zeta) := H\left(\frac{p+q}{2}\right)\Delta\zeta + \frac{\theta h}{2d}\sum_{k=1}^d (q_k - p_k),$$

where the vector $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$ stands for the discrete derivatives

$$p = (D_{h,e_1}^+, D_{h,e_2}^+, \ldots, D_{h,e_d}^+), \qquad q = (D_{h,e_1}^-, D_{h,e_2}^-, \ldots, D_{h,e_d}^-),$$

$$e_k := (0, 0, \ldots, 0, \underbrace{1}_{k}, 0, \ldots, 0) \quad \text{for } k = 1, 2, \ldots, d.$$

A calculation verifies that (4.11) and (4.12) are satisfied with $\lambda_0 := \frac{\theta}{d\|DH\|_\infty}$.

If $d = 1$, the different regions of monotonicity of H may be exploited to create upwind schemes. As a simple example, assume that $H \geq H(0) = 0$ and H is increasing for $p > 0$ and decreasing for $p < 0$, and define

$$H_h(p, q, \Delta\zeta) := [H(p_+) + H(-q_-)](\Delta\zeta)_+ - [H(q_+) + H(-p_-)](\Delta\zeta)_-.$$

Then (4.11) and (4.12) hold with $\theta = 0$ and $\lambda_0 := \frac{1}{2\|H'\|_\infty}$.

As far as the approximating paths W_h are concerned, Lemma 4.1 implies that (4.4) and (4.5) will hold, with $k = 2$, if ρ_h and W_h satisfy

$$(4.13) \qquad \sup_{0 \leq t - s \leq \rho_h} |W_h(t) - W_h(s)| \leq \lambda_0 h \quad \text{and} \quad \lim_{h \to 0} \frac{h^2}{\rho_h} = 0.$$

If W_h is smooth, then

$$\sup_{0 \leq t - s \leq \rho_h} |W_h(t) - W_h(s)| \leq \|\dot{W}_h\|_\infty \rho_h.$$

Let $\omega : [0, \infty) \to [0, \infty)$ be the modulus of continuity for W. For many standard approximations of W, there exists some increasing function $h \mapsto \eta_h$ satisfying $\lim_{h \to 0^+} \eta_h = 0$ and some $C > 0$ such that

$$(4.14) \qquad \|\dot{W}_h\|_\infty \leq C \frac{\omega(\eta_h)}{\eta_h}.$$

For example, W_h may be the piecewise linear interpolation of W with step-size η_h, or the convolution of W with a standard mollifier supported in an interval of radius η_h. Then the first part of (4.13) may be replaced with the slightly stronger assumption

$$(4.15) \qquad \frac{C \omega(\eta_h) \rho_h}{h \eta_h} \leq \lambda_0.$$

To be more explicit, suppose that $W \in C^\alpha([0, T], \mathbb{R}^m)$ and, for some $\gamma > 0$, $\eta_h = (\rho_h)^\gamma$. Then (4.15) will hold if ρ_h is defined by

$$\lambda := \frac{C [W]_{\alpha, T} (\rho_h)^{1 - \gamma + \alpha \gamma}}{h} \leq \lambda_0.$$

This yields

$$\frac{h^2}{\rho_h} \approx (\rho_h)^{1 - 2\gamma + 2\alpha\gamma},$$

so that (4.13) will be satisfied if

$$0 < \gamma < \frac{1}{2(1 - \alpha)}.$$

If $\alpha > \frac{1}{2}$, then γ is allowed to be 1, and in particular, it is natural to define W_h to be the piecewise linear interpolation of W on a partition of step-size $\eta_h = \rho_h$. Notice also that paths in C^α for such α have quadratic variation equal to 0.

However, for $\alpha \leq \frac{1}{2}$, γ is forced to be less than 1, and so we must make W_h a milder approximation. The work in the subsequent sections suggests that choosing $\gamma = \frac{1}{2}$ gives the best rate of convergence regardless of the regularity of the path W.

4.1.2. A second order example.

Verifying (4.4) and (4.5) is more complicated for finite difference approximations of second order equations. Rather than stating very general assumptions on F_h or H_h, we perform these calculations for a specific scheme. More examples can be formed by adapting the results of [20, 21].

Assume for simplicity that $d = 1$, $H \in C^3(\mathbb{R}, \mathbb{R}^m)$, and that F depends only on u_{xx}, and define, for some $\epsilon_h > 0$,

$$H_h(p, q, \Delta \zeta) := H \left(\frac{p + q}{2} \right) \Delta \zeta \quad \text{and}$$

$$F_h(X) = F(X) + \epsilon_h X \quad \text{for } X = D_{h,1}^2 u \text{ and } (p, q) = D_{h,1} u.$$

Note that the ellipticity condition (3.1) means that F is increasing, and so a routine calculation shows that S_h, W_h, and ρ_h satisfy (4.4) if

$$\rho_h := \lambda h^2 \quad \text{with } \lambda \leq \frac{1}{2 \|F'\|_\infty}$$

and

$$(4.16) \qquad \|\dot{W}_h\|_\infty \leq \frac{2}{\|H'\|_\infty} \cdot \frac{\epsilon_h}{h}.$$

Now let $\Phi_h \in C(I, C^3(\mathbb{R}))$ and $\phi \in C^3(\mathbb{R})$ be as in (4.5). Observe that it is possible to find such a solution because of the added regularity for H, and that

$$\sup_{h>0}(\|\Phi_{h,xx}\|_\infty + \|\Phi_{h,xxx}\|_\infty) < \infty.$$

Then, for some $C > 0$ depending only on $\|H'\|_\infty$, and for all $\rho \in (0, \lambda h^2)$,

$$\left| \Phi_h(x, t+\rho) - \Phi_h(x, t) - \sum_{i=1}^m H^i(\Phi_{h,x}(x, t))(W_h(t+\rho) - W_h(t)) \right|$$

$$\leq C\left(\|\Phi_{h,xx}\|_\infty \max_{t \leq s \leq t+\rho} |W_h(s) - W_h(t)|^2 \right) \leq C\lambda \|\Phi_{h,xx}\|_\infty (\epsilon_h)^2 \rho.$$

The estimates (4.8) and (4.9) then imply that, for s_h and t_h as in (4.5),

$$\left| S_h(t_h, s_h; W_h)\Phi_h(\cdot, s_h)(x) - \Phi_h(x, t_h) - (t_h - s_h)F(\phi_{xx}(x, t)) \right|$$

$$\leq C\rho_h \cdot \left(\|\Phi_{h,xx}\|_\infty \epsilon_h + \|\Phi_{h,xxx}\|_\infty h + \|\Phi_{h,xx}(\cdot, s_h) - \phi_{xx}\|_\infty \right),$$

and so (4.5) holds if $\lim_{h\to 0} \epsilon_h = 0$. This, in turn, requires that

$$\lim_{h\to 0} h\|\dot{W}_h\|_\infty = 0,$$

or that W_h satisfies (4.14) with η_h such that

$$\lim_{h\to 0} \frac{h\omega(\eta_h)}{\eta_h} = 0.$$

Taking $W \in C^\alpha([0, T], \mathbb{R}^m)$ and $\eta_h = (\rho_h)^\gamma = \lambda^\gamma h^{2\gamma}$ for some $\gamma > 0$ as a concrete example, this leads once more to the restriction

$$0 < \gamma < \frac{1}{2(1-\alpha)}.$$

4.2. *Other approximations.*

4.2.1. *Stability for* (1.1). The proof of Theorem 4.1 is a generalization of the argument that (1.1) is stable with respect to perturbations in the data. In fact, Theorem 4.1 recovers these stability properties.

Suppose that

$$(4.17) \quad \begin{cases} u_0^\epsilon \in C^{0,1}(\mathbb{R}^d), & W^\epsilon, W \in C([0, T]; \mathbb{R}^m), \\ H^\epsilon, H \in C^2(\mathbb{R}^d; \mathbb{R}^m), & F^\epsilon, F \text{ satisfy (3.1)}, \\ \text{and} \quad \lim_{\epsilon\to 0}(\|u_0^\epsilon - u_0\|_\infty, \|W^\epsilon - W\|_\infty, \|H^\epsilon - H\|_{C^2}, \|F^\epsilon - F\|_\infty) = 0, \end{cases}$$

and let $u^\epsilon \in BUC(\mathbb{R}^d \times [0, T])$ be the unique solution of

$$(4.18) \quad \begin{cases} du^\epsilon = F^\epsilon(D^2 u^\epsilon, Du^\epsilon)\,dt + \sum_{i=1}^m H^{i,\epsilon}(Du^\epsilon) \circ dW^{i,\epsilon} & \text{in } \mathbb{R}^d \times (0, T] \text{ and} \\ u^\epsilon(\cdot, 0) = u_0^\epsilon & \text{on } \mathbb{R}^d. \end{cases}$$

THEOREM 4.2. *Assume* (4.17) *and let* u^ϵ *and* u *solve respectively* (4.18) *and* (1.1). *Then, as* $\epsilon \to 0$, u^ϵ *converges locally uniformly to* u.

PROOF. The comparison principle implies that the solution operator for (4.18) is contractive, and therefore, it suffices to assume that $u_0^\epsilon = u_0$ for all $\epsilon > 0$.

For $s \leq t$, $\zeta \in C([0, T]; \mathbb{R}^m)$, and $h > 0$, let $S_\epsilon(t, s; \zeta) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ be the solution operator for (4.18) driven by the path ζ instead of W^ϵ. Properties (4.1) and (4.2) are readily verified, and, letting $\rho_h = \rho_\epsilon$ be arbitrary and setting $W_h = W^\epsilon$, (4.4) follows immediately from the comparison principle.

Finally, in view of the uniform bound for $D^2 H^\epsilon$, for any interval $I \subset [0, T]$ and solution $\Phi \in C(I, C^2(\mathbb{R}^d))$ of (3.6), there exists a family of solutions $\Phi^\epsilon \in C(I, C^2(\mathbb{R}^d))$ solving (3.6) with the Hamiltonian H^ϵ and path W^ϵ, converging in $C(I, C^2(\mathbb{R}^d))$ to Φ as $\epsilon \to 0$. This can be seen using the method of characteristics, as in Section 3. Therefore, (4.5) is a consequence of Definition 3.1 and the local uniform convergence of F^ϵ to F. Theorem 4.1 now gives the result. $\square$

4.2.2. *A mixing formula.* It is also possible to derive general Trotter–Kato type mixing formulas for (1.1). Here, we present a specific example. A different approach to the following can be found in the work of Gassiat and Gess [13].

Assume, in addition to (3.1), that

$$F \in C^{1,1}(S^d \times \mathbb{R}^d) \quad \text{and} \quad H \in C^4(\mathbb{R}^d, \mathbb{R}^m),$$

and, for $\zeta \in C([0, T], \mathbb{R}^m)$, let $S_F(t) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ and $S_H(t, s; \zeta) : BUC(\mathbb{R}^d) \to BUC(\mathbb{R}^d)$ be the solution operators for respectively

$$u_t = F(D^2 u, Du) \quad \text{and} \quad du = \sum_{i=1}^m H^i(Du) \circ d\zeta^i.$$

Define

$$S_h(t, s; \zeta) = S_F(t - s) S_H(t, s; \zeta).$$

THEOREM 4.3. *For any sequence of approximating paths $\{W_h\}_{h>0}$ and modulus $h \to \rho_h$ satisfying (4.3), the triple (S_h, W_h, ρ_h) satisfies (4.1)–(4.5).*

PROOF. Properties (4.1)–(4.4) are immediate from the definitions of the above objects. Let $I \subset [0, T]$, $s_h, t_h \in I$, $\Phi_h \in C(I; C^4(\mathbb{R}^d))$, and $\phi \in C^4(\mathbb{R}^d)$ be as in (4.5). Such a solution Φ exists in view of the additional regularity assumed for H.

For any $x \in \mathbb{R}^d$,

$$S_h(t_h, s_h; W_h) \Phi_h(\cdot, s_h)(x) - \Phi_h(x, t_h) = S_F(t_h - s_h) \Phi_h(\cdot, t_h)(x) - \Phi_h(x, t_h).$$

Define $\phi_h := \Phi_h(\cdot, t_h)$, which satisfies

$$R := \sup_{h>0} \|\phi_h\|_{C^4(\mathbb{R}^d)} < \infty \quad \text{and} \quad \lim_{h \to 0} \|\phi_h - \phi\|_{C^2(\mathbb{R}^d)} = 0,$$

and let

$$z_h(x, t) := \phi_h(x) + t F(D^2 \phi_h(x), D\phi_h(x)).$$

Then, for some universal constant $C > 0$, z_h is a viscosity super-solution of

$$z_{h,t} \geq F(D^2 z_h, Dz_h) - C\|F\|_{C^{1,1}(\mathbb{R}^d)} R\rho_h \quad \text{in } \mathbb{R}^d \times [0, \rho_h],$$

so that, for all $\rho \in (0, \rho_h)$,

$$\sup_{x \in \mathbb{R}^d} (S_F(\rho)\phi_h(x) - z_h(x, \rho)) \leq C\|F\|_{C^{1,1}(\mathbb{R}^d)} R\rho_h \rho.$$

A similar argument, using that z_h satisfies an analogous viscosity sub-solution property, gives a lower bound, whence

$$\left|S_F(t_h - s_h)\phi_h(x) - \phi_h(x) - (t_h - s_h)F\big(D^2\phi_h(x), D\phi_h(x)\big)\right| \le C\|F\|_{C^{1,1}(\mathbb{R}^d)} R\rho_h(t_h - s_h).$$

Property (4.5) now follows, with $k = 4$, from the fact that

$$\lim_{h \to 0} F(D^2\phi_h, D\phi_h) = F(D^2\phi, D\phi) \quad \text{uniformly.} \qquad \square$$

4.3. *The proof of Theorem* 4.1. Define

$$u^*(x, t) = \limsup_{h \to 0, (y,s) \to (x,t)} u_h(y, s) \quad \text{and} \quad u_*(x, t) = \liminf_{h \to 0, (y,s) \to (x,t)} u_h(y, s).$$

The functions u^* and u_*, called the half-relaxed limits of u_h, are respectively upper- and lower-semicontinuous. Furthermore, $u_* \le u^*$ on $\mathbb{R}^d \times [0, T]$ and $u_*(\cdot, 0) \le u_0 \le u^*(\cdot, 0)$ on $\mathbb{R}^d$. The goal will be to show that $u_* = u^*$, which yields the local uniform convergence of u_h and the fact that the limit u solves (1.1).

Step 1: *Finiteness of* u^* *and* u_*. Observe that, for any constant $k \in \mathbb{R}$, the function

$$\Phi_h(x, t) = k + \sum_{i=1}^{m} H^i(0) W_h^i(t)$$

is a smooth solution of (3.6) for all $(x, t) \in \mathbb{R}^d \times [0, T]$. Therefore, in view of (4.4) and (4.5),

$$u_h(x, t) \le \|u_0\|_\infty + \sum_{i=1}^{m} H^i(0) W_h^i(t) + T\big(F(0, 0) + 1\big)$$

for all sufficiently small $h > 0$, and so $u^*(x, t) < \infty$ for all $(x, t) \in \mathbb{R}^d \times [0, T]$. A similar argument gives $u_* > -\infty$.

Step 2: The solution inequalities. In this step, we demonstrate that u^* and u_* satisfy respectively the sub- and super-solution properties in Definition 3.1 for equation (1.1). Only the argument for u^* is presented, since the proof for u_* is similar.

Assume that $(x_0, t_0) \in \mathbb{R}^d \times (0, T]$, $I \ni t_0$, $\psi \in C^1([0, T])$, $\Phi \in C(I, C^k(\mathbb{R}^d))$ solves (3.6) with

$$\max_{j=2,3,\dots,k} \sup_{t \in I} \left\|D^j \Phi(\cdot, t)\right\|_\infty < \infty,$$

and $u^*(x, t) - \Phi(x, t) - \psi(t)$ attains a local maximum at (x_0, t_0). As discussed in Section 3, it may be assumed that this maximum is strict in $\mathbb{R}^d \times I$, and that

(4.19) $$\lim_{|x| \to +\infty} \frac{\Phi(x, t)}{|x|} = +\infty \quad \text{uniformly for } t \in I.$$

The definition of u^* implies that there exist $y_h \in \mathbb{R}^d$ and $s_h \in [0, T]$ such that

$$\lim_{h \to 0} \big(y_h, s_h, u_h(y_h, s_h)\big) = \big(x_0, t_0, u^*(x_0, t_0)\big).$$

The method of characteristics and the fact that $\lim_{h \to 0} \|W_h - W\|_\infty = 0$ yield the existence of a subinterval of I containing t_0, relabeled as I for convenience, such that, for all $h > 0$, there exists a solution $\Phi_h \in C(I, C^k(\mathbb{R}^d))$ of

$$d\Phi_h = \sum_{i=1}^{m} H^i(D\Phi_h) \circ dW_h \quad \text{in } \mathbb{R}^d \times I \quad \text{and} \quad \Phi_h(\cdot, t_0) = \Phi(\cdot, t_0) \quad \text{in } \mathbb{R}^d$$

that satisfies (4.19) uniformly in h, and Φ_h converegs to Φ in $C(I, C^k(\mathbb{R}^d))$ as $h \to 0$. It follows that

$$u_h(x, t) - \Phi_h(x, t) - \psi(t)$$

attains a global maximum at $(\hat{y}_h, \hat{s}_h)$ over $\mathbb{R}^d \times \overline{I}$ such that $\{\hat{y}_h\}_{h>0}$ is bounded. This gives, in particular,

$$u_h(y_h, s_h) - \Phi_h(y_h, s_h) - \psi(s_h) \leq u_h(\hat{y}_h, \hat{s}_h) - \Phi_h(\hat{y}_h, \hat{s}_h) - \psi(\hat{s}_h).$$

Let $(\hat{x}, \hat{t})$ be a limit point of the sequence $\{(\hat{y}_h, \hat{s}_h)\}_{h>0}$. Taking $h \to 0$ along the appropriate subsequence above results in the inequality

$$u^*(x_0, t_0) - \Phi(x_0, t_0) - \psi(t_0) \leq u^*(\hat{x}, \hat{t}) - \Phi(\hat{x}, \hat{t}) - \psi(\hat{t}).$$

The strictness of the original maximum then implies that $\lim_{h \to 0}(\hat{y}_h, \hat{s}_h) = (x_0, t_0)$.

Because $|\mathcal{P}_h| \leq \rho_h \xrightarrow{h \to 0} 0$, it follows that, for sufficiently small h, there exists $t_n \in \mathcal{P}_h$ such that $t_n < \hat{s}_h \leq t_{n+1}$ and $t_n \in I$. Then, for all $x \in \mathbb{R}^d$,

$$(4.20) \qquad u_h(x, t_n) \leq u_h(\hat{y}_h, \hat{s}_h) + \Phi_h(x, t_n) - \Phi_h(\hat{y}_h, \hat{s}_h) + \psi(t_n) - \psi(\hat{s}_h).$$

Applying the operator $S_h(\hat{s}_h, t_n; W_h)$ to both sides of (4.20), using (4.4) and the fact that $0 < \hat{s}_h - t_n \leq \rho_h$, and rearranging terms yields

$$\frac{\psi(\hat{s}_h) - \psi(t_n)}{\hat{s}_h - t_n} \leq \frac{S_h(\hat{s}_h, t_n; W_h)\Phi_h(\cdot, t_n)(\hat{y}_h) - \Phi_h(\hat{y}_h, \hat{s}_h)}{\hat{s}_h - t_n}.$$

Sending $h \to 0$ and using (4.5) gives $\psi'(t_0) \leq F(D^2\Phi(x_0, t_0), D\Phi(x_0, t_0))$, as desired.

Step 3: Initial data. We now prove that $u^*(x, 0) = u_0(x) = u_*(x, 0)$. Only the first equality is considered, and since $u^*(x, 0) \geq u_0(x)$, it suffices to show that $u^*(x, 0) \leq u_0(x)$.

Let $\phi \in C^k(\mathbb{R}^d)$ be such that

$$R := \max_{j=2,3,\ldots,k} \|D^j\phi\|_\infty < \infty$$

and $u_0 \leq \phi$ on $\mathbb{R}^d$, and let $I \ni 0$ and $\Phi \in C(I, C^k(\mathbb{R}^d))$ be a solution of (3.6) with $\Phi(\cdot, 0) = \phi$. Define $\phi_h \in C^{0,1}(\mathbb{R}^d \times [0, T])$ as in (4.6) with initial condition $\phi_h(\cdot, 0) = \phi$, path W_h, and partition $\mathcal{P}_h$. Then (4.4) and (4.5) yield, for some $C > 0$ depending only on R and $\|DF\|_\infty$, and for any $(y, s) \in \mathbb{R}^d \times I$ and sufficiently small h,

$$u_h(y, s) \leq \phi_h(y, s) \leq \Phi(y, s) + Cs.$$

Sending $(y, s) \to (x, 0)$ and $h \to 0$, this becomes $u^*(x, 0) \leq \phi(x)$, completing the argument since ϕ was arbitrary.

Step 4: The comparison principle. In view of the comparison principle (3.14), $u^*(x, t) \leq u_*(x, t)$ for all $(x, t) \in \mathbb{R}^d \times [0, T]$. Therefore $u^* = u_*$, and the result is proved.

5. The pathwise estimate. The remaining sections focus on deriving quantitative error estimates for schemes in the first-order setting. We will henceforth always assume that H satisfies (3.2) and u_0 satisfies (3.3) (and thus, without loss of generality, H satisfies (3.4)). Also, in addition to (4.1) and (4.2), the schemes in this part of the paper will be required to satisfy the following quantitative versions of (4.4) and (4.5): for some $\lambda_0 > 0$,

$$(5.1) \qquad \text{if } u_1 \leq u_2 \text{ and } \operatorname{osc}(\zeta, s, t) \leq \lambda_0 h, \quad \text{then } S_h(t, s; \zeta)u_1 \leq S_h(t, s; \zeta)u_2,$$

and

$$
(5.2) \quad \begin{cases} \text{there exists } C = C_L > 0 \text{ such that, if } \zeta \in C([0, T], \mathbb{R}^m), \ \Phi \in C(I, C^{1,1}(\mathbb{R}^d)) \\ \text{is a solution of } d\Phi = \sum_{i=1}^{m} H^i(D\Phi) \circ d\zeta^i \text{ in } \mathbb{R}^d \times I, \text{ and } \mathrm{osc}(\zeta, s, t) \leq \lambda_0 h, \text{ then} \\ \left\| S_h(t, s; \zeta)\Phi(\cdot, s) - \Phi(\cdot, t) \right\|_\infty \leq C \left\| D^2 \Phi \right\|_\infty h^2. \end{cases}
$$

This is motivated by the properties obtained in Lemma 4.1 for the finite difference approximations discussed in Section 4.1.1.

Fix a partition

$$
\mathcal{P} = \{0 = t_0 < t_1 < t_2 < \cdots < t_N = T\}
$$

of $[0, T]$, set $(\Delta t)_n := t_{n+1} - t_n$, and let $\zeta : [0, T] \to \mathbb{R}^m$ be any continuous path satisfying

$$
(5.3) \quad \begin{cases} \zeta(0) = 0, \quad \zeta \text{ is affine on } [t_n, t_{n+1}] \text{ for every } n = 0, 1, 2, \ldots, N - 1, \text{ and} \\ \max_{n=0,1,2,\ldots,N-1} |\zeta(t_{n+1}) - \zeta(t_n)| \leq \lambda_0 h. \end{cases}
$$

In this section, we obtain an estimate for the error between the viscosity solution v of

$$
(5.4) \qquad v_t = \sum_{i=1}^{m} H^i(Dv)\dot{\zeta}^i(t) \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad v(\cdot, 0) = u_0 \quad \text{on } \mathbb{R}^d
$$

and the approximate solution $v_h(\cdot; \zeta, \mathcal{P})$ given by (4.6), which, for convenience, we define again here:

$$
(5.5) \qquad \begin{cases} v_h(\cdot, 0; \zeta, \mathcal{P}) := u_0, \\ v_h(\cdot, t; \zeta, \mathcal{P}) := S_h(t, t_n; \zeta)v_h(\cdot, t_n; \zeta, \mathcal{P}) \\ \qquad \text{for } n = 0, 1, \ldots, N - 1 \text{ and } t \in (t_n, t_{n+1}]. \end{cases}
$$

THEOREM 5.1. *Assume* (3.2), (3.3) *and* (3.4). *Then there exists* $C = C_L > 0$ *such that, if* S_h *satisfies* (4.1), (4.2), (5.1), *and* (5.2), ζ *and* $\mathcal{P}$ *satisfy* (5.3), *and* v *and* v_h *are as in* (5.4) *and* (5.5) *with* $\|Du_0\|_\infty \leq L$, *then, for all* $\epsilon, h > 0$,

$$
\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |v_h(x, t; \zeta, \mathcal{P}) - v(x, t)|
$$

$$
\leq \frac{1}{\epsilon} \sum_{n=0}^{N-1} (\Delta t_n)^2 + C\sqrt{N}h + \max_{s,t\in[0,T]} \left\{ C |\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right\}.
$$

Before proving Theorem 5.1, we state some regularity estimates for v and v_h. First, the monotonicity of the scheme operator S_h, the comparison principle for (6.1), and the translation invariance of the solution operators for each immediately yield the Lipschitz bounds

$$
(5.6) \qquad \|Dv\|_\infty, \|Dv_h\|_\infty \leq L.
$$

The regularity of v_h and v in the time variable is established by the next result.

LEMMA 5.1. *Assume* (3.2), (3.3) *and* (3.4). *There exists* $C = C_L > 0$ *such that, for all* $(x, s, t) \in \mathbb{R}^d \times [0, T] \times [0, T]$ *with* $s < t$,

$$
(5.7) \qquad |v(x, t) - v(x, s)| \leq C\mathrm{osc}(\zeta, s, t)
$$

and, for all $m, n \in \{0, 1, 2, \ldots, N\}$ *with* $m < n$,

$$
(5.8) \qquad |v_h(x, t_n; \zeta, \mathcal{P}) - v_h(x, t_m; \zeta, \mathcal{P})| \leq C(h\sqrt{n - m} + \mathrm{osc}(\zeta, t_m, t_n)).
$$

PROOF.   The bound (5.7) follows from the cancellation estimates presented in Proposition 7.2 of [31]. To prove (5.8), observe first that, in view of Lemma 3.2(c), there exists $C = C_L > 0$ such that, for all $z \in \mathbb{R}^d$ and $\delta > 0$,

$$L|z| \leq \Phi_\delta(z, t_m, t_m; \zeta) + C\delta.$$

Then (5.6) yields, for all $x, y \in \mathbb{R}^d$,

(5.9)
$$\begin{aligned}
&v_h(x, t_m; \zeta, \mathcal{P}) \\
&\quad \leq v_h(y, t_m; \zeta, \mathcal{P}) + L|x - y| \leq v_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x - y, t_m, t_m; \zeta) + C\delta.
\end{aligned}$$

Keeping y fixed, we then apply the operator $\prod_{k=m}^{n-1} S_h(t_{k+1}, t_k; \zeta, \mathcal{P})$ to the left- and right-hand sides of the inequality (5.9), which is preserved because of the monotonicity of this operator implied by (5.1) and (5.3). According to (5.5), the left-hand side becomes $v_h(x, t_n; \zeta, \mathcal{P})$. Iteratively using (5.2) to compare the right-hand side to $\Phi_\delta(x - y, t_n, t_m; \zeta, \mathcal{P})$ yields, in view of Lemma 3.2(b),

$$v_h(x, t_n; \zeta, \mathcal{P}) \leq v_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x - y, t_n, t_m; \zeta) + C\big(\delta + (n - m)\|D^2\Phi_\delta\|_\infty h^2\big)$$

$$\leq v_h(y, t_m; \zeta, \mathcal{P}) + \Phi_\delta(x - y, t_n, t_m; \zeta) + C\left(\delta + \frac{(n - m)h^2}{\delta}\right),$$

as long as $\mathrm{osc}(\zeta, t_m, t_n) \leq \delta$. Setting $x = y$ gives

$$v_h(x, t_n; \zeta, \mathcal{P}) - v_h(x, t_m; \zeta, \mathcal{P}) \leq C \inf\left\{\delta + \frac{(n - m)h^2}{\delta} : \delta \geq \mathrm{osc}(\zeta, t_m, t_n)\right\}.$$

If $\mathrm{osc}(\zeta, t_m, t_n) \leq h\sqrt{n - m}$, then the right-hand side is optimized by choosing $\delta = h\sqrt{n - m}$. Otherwise, setting $\delta = \mathrm{osc}(\zeta, t_m, t_n)$ gives the result, since in this case,

$$\frac{(n - m)h^2}{\delta} = \frac{(n - m)h^2}{\mathrm{osc}(\zeta, t_m, t_n)} \leq h\sqrt{n - m}.$$

The lower bound for $v_h(\cdot, t_n; \zeta, \mathcal{P}) - v_h(\cdot, t_m; \zeta, \mathcal{P})$ is proved similarly.   □

PROOF OF THEOREM 5.1.   Throughout the proof, to simplify the presentation, we set $v_h(x, t) := v_h(x, t; \zeta, \mathcal{P})$. Fix a constant $\overline{C} = \overline{C}_L > 0$ to be determined later, and let $\alpha, \mu : [0, T] \to \mathbb{R}$ be the nondecreasing, lower-semicontinuous, piecewise constant functions defined by $\alpha(0) = \mu(0) = 0$ and

$$\alpha(s) - \alpha(t_n) := \big[(\Delta t)_n\big]^2 \quad \text{and} \quad \mu(s) - \mu(t_n) = \overline{C}h^2$$

$$\text{for } n = 0, 1, 2, \ldots, N - 1 \text{ and } s \in (t_n, t_{n+1}].$$

Choose $\epsilon > 0$ and

(5.10)
$$\delta > \max\left\{2\lambda_0 h, \max_{s,t \in [0,T]}\left\{\overline{C}|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon}\right\}\right\},$$

and define the auxiliary function $\Psi : [0, T] \times [0, T] \to \mathbb{R}$ by

(5.11) $\Psi(s, t) = \sup_{x, y \in \mathbb{R}^d}\big\{v_h(x, s) - v(y, t) - \Phi_\delta(x - y, s, t; \zeta)\big\} - \frac{|s - t|^2}{2\epsilon} - \frac{\mu(s)}{\delta} - \frac{\alpha(s)}{\epsilon},$

where Φ_δ is the "distance function" given in (3.12).

Step 1: We first prove that, if $\overline{C}$ is sufficiently large, then

(5.12)
$$\max_{[0,T]^2} \Psi = \max\left\{\max_{s \in [0,T]} \Psi(s, 0), \max_{t \in [0,T]} \Psi(0, t)\right\}.$$

Assume for the sake of contradiction that, for some $\sigma > 0$, $\Psi(s, t) - \sigma t$ attains its maximum in $[0, T] \times [0, T]$ at $(\hat{s}, \hat{t})$ with $\hat{s} > 0$ and $\hat{t} > 0$.

The first observation is that, for some $M = M_L > 0$, the supremum in (5.11) may be restricted to $x, y \in \mathbb{R}^d$ satisfying $|x - y| \leq M\delta$. This is because, for any $s, t \in [0, T]$ and for some $C' = C'_L > 0$,

$$\sup_{x,y \in \mathbb{R}^d} \{v_h(x, s) - v(y, t) - \Phi_\delta(x - y, s, t; \zeta)\} \geq \sup_{x \in \mathbb{R}^d} \{v_h(x, s) - v(x, t)\} - C'\delta,$$

while, if $|x - y| > M\delta$, then (5.6) and Lemma 3.2(c) give, for some $C = C_L > 0$,

$$v_h(x, s) - v(y, t) - \Phi_\delta(x - y, s, t; \zeta)$$

$$\leq \sup_{x \in \mathbb{R}^d} \{v_h(x, s) - v(x, t)\} + L|x - y| - \frac{|x - y|^2}{2(C + 1)\delta} + C\delta$$

$$\leq \sup_{x \in \mathbb{R}^d} \{v_h(x, s) - v(x, t)\} - \frac{M^2\delta}{4(C + 1)} + (C + (C + 1)L^2)\delta$$

$$< \sup_{x \in \mathbb{R}^d} \{v_h(x, s) - v(x, t)\} - C'\delta,$$

where the last inequality holds if M is sufficiently large.

As a result, if $\overline{C}$ is large enough, then $(\hat{s}, \hat{t}) \in U_{\delta/2}(W)$. To verify this, we rearrange terms in the inequality $\Psi(\hat{s}, \hat{s}) \leq \Psi(\hat{s}, \hat{t})$ and use Lemmas 3.2(a) and 5.1 to obtain, for some $C = C_L > 0$,

$$\frac{|\hat{s} - \hat{t}|^2}{2\epsilon} \leq \sup_{|x-y| \leq M\delta} \{v_h(y, \hat{s}) - v(y, \hat{t}) + \Phi_\delta(x - y, \hat{s}, \hat{s}; \zeta) - \Phi_\delta(x - y, \hat{s}, \hat{t}; \zeta)\}$$

$$\leq C\operatorname{osc}(\zeta, \hat{s}, \hat{t}).$$

Consequently,

$$\overline{C}\operatorname{osc}(\zeta, \hat{s}, \hat{t}) \leq \max_{s,t \in [0,T]} \left\{ \overline{C}\operatorname{osc}(\zeta, s, t) - \frac{|s - t|^2}{2\epsilon} \right\} + \frac{|\hat{s} - \hat{t}|^2}{2\epsilon}$$

$$\leq \max_{s,t \in [0,T]} \left\{ \overline{C}|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right\} + C\operatorname{osc}(\zeta, \hat{s}, \hat{t}) \leq \delta + C\operatorname{osc}(\zeta, \hat{s}, \hat{t}),$$

so that

$$\operatorname{osc}(\zeta, \hat{s}, \hat{t}) \leq \frac{\delta}{\overline{C} - C} < \frac{\delta}{2} \quad \text{if } \overline{C} > C + 2.$$

Now, if $\hat{n} \in \{0, 1, 2, \ldots, N - 1\}$ is the integer satisfying $t_{\hat{n}} < \hat{s} \leq t_{\hat{n}+1}$, then the linearity of ζ on $[t_{\hat{n}}, t_{\hat{n}+1}]$ implies that

$$|\zeta(\hat{s}) - \zeta(t_{\hat{n}})| \leq \lambda_0 h < \frac{\delta}{2},$$

and so the triangle inequality yields $(t_{\hat{n}}, \hat{t}) \in U_\delta(\zeta)$. This, in turn, means that $(s, \hat{t}) \in U_\delta(\zeta)$ for all $s \in [t_{\hat{n}}, \hat{s}]$.

We next use the definition of pathwise viscosity solutions to establish the inequality

(5.13)
$$\frac{\hat{s} - \hat{t}}{\epsilon} \geq \sigma.$$

In view of Lemma 3.2(c), for any $x \in \mathbb{R}^d$, the function

$$y \mapsto v(y, t) + \Phi_\delta(x - y, \hat{s}, t; \zeta)$$

attains a global minimum over $\mathbb{R}^d$. Definition 3.2 and Lemma 3.2(d) then imply that

$$t \mapsto \inf_{y \in \mathbb{R}^d} \{v(y, t) + \Phi_\delta(x - y, \hat{s}, t; \zeta)\}$$

is nondecreasing on $I := \{t \in [0, T] : (\hat{s}, t) \in U_\delta(\zeta)\}$, and therefore, so is

$$\phi(t) := \inf_{x, y \in \mathbb{R}^d} \{v(y, t) - v_h(x, \hat{s}) + \Phi_\delta(x - y, \hat{s}, t; \zeta)\}.$$

Since $\phi(t) + \frac{|\hat{s} - t|^2}{2\epsilon} + \sigma t$ attains a minimum at $\hat{t} \in I$, (5.13) follows.

On the other hand, we obtain a contradiction by using (5.1) and (5.2) to show that

(5.14)
$$\frac{\hat{s} - \hat{t}}{\epsilon} \leq 0.$$

The first step is to prove that, for each $y \in \mathbb{R}^d$, the function

$$a(s) := \sup_{x \in \mathbb{R}^d} \{v_h(x, s) - \Phi_\delta(x - y, s, \hat{t}; \zeta)\} - \frac{\mu(s)}{\delta}$$

satisfies

$$\max_{[t_{\hat{n}}, t_{\hat{n}+1}]} a = a(t_{\hat{n}}).$$

Indeed, if this were not the case, then, for some $s^* \in [t_{\hat{n}}, t_{\hat{n}+1}]$ and sufficiently small $\beta > 0$,

$$a(t_{\hat{n}}) \leq a(s^*) - \beta(s^* - t_{\hat{n}}).$$

Lemma 3.2(c) implies that the supremum in the definition of $a(s^*)$ is attained for some $x^* \in \mathbb{R}^d$, and so it follows that, for all $x \in \mathbb{R}^d$,

$$v_h(x, t_{\hat{n}}) \leq v_h(x^*, s^*) + \Phi_\delta(x - y, t_{\hat{n}}, \hat{t}; \zeta) - \Phi_\delta(x^* - y, s^*, \hat{t}; \zeta)$$

(5.15)
$$- \frac{\mu(s^*) - \mu(t_{\hat{n}})}{\delta} - \beta(s^* - t_{\hat{n}}).$$

In view of (5.1) and the fact that $\mathrm{osc}(\zeta, t_{\hat{n}}, s^*) \leq \lambda_0 h$, the operator $S_h(s^*, t_{\hat{n}}; \zeta)$ is monotone. Applying it to both sides of the inequality (5.15), setting $x = x^*$, rearranging terms, and using (5.2) and Lemma 3.2(b) and (d) yield

$$\frac{\overline{C} h^2}{\delta} + \beta(s^* - t_{\hat{n}}) = \frac{\mu(s^*) - \mu(t_{\hat{n}})}{\delta} + \beta(s^* - t_{\hat{n}}) \leq C \|D^2 \Phi\|_\infty h^2 \leq \frac{C h^2}{\delta}.$$

This results in a contradiction as long as $\overline{C} \geq C$.

As a consequence,

$$\psi(s) := \sup_{x, y \in \mathbb{R}^d} \{v_h(x, s) - v(y, \hat{t}) - \Phi_\delta(x - y, s, \hat{t}; \zeta)\} - \frac{\mu(s)}{\delta}$$

attains its maximum in $[t_{\hat{n}}, t_{\hat{n}+1}]$ at $t_{\hat{n}}$, and therefore, because $\psi(s) - \frac{|s - \hat{t}|^2}{2\epsilon} - \frac{\alpha(s)}{\epsilon}$ attains a maximum at $\hat{s}$,

$$\psi(t_{\hat{n}}) - \frac{|t_{\hat{n}} - \hat{t}|^2}{2\epsilon} - \frac{\alpha(t_{\hat{n}})}{\epsilon} \leq \psi(\hat{s}) - \frac{|\hat{s} - \hat{t}|^2}{2\epsilon} - \frac{\alpha(\hat{s})}{\epsilon} \leq \psi(t_{\hat{n}}) - \frac{|\hat{s} - \hat{t}|^2}{2\epsilon} - \frac{\alpha(\hat{s})}{\epsilon},$$

which, after rearranging terms, yields (5.14). Together with (5.13), this establishes (5.12).

Step 2: The next claim is that, for some $C = C_L > 0$,

$$\max_{\{0\} \times [0, T] \cup [0, T] \times \{0\}} \Psi \leq C(\delta + \sqrt{N} h) + \max_{s, t \in [0, T]} \left\{ C |\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right\}.$$

Assume that Ψ attains its maximum at $(\hat{s}, \hat{t})$, with either $\hat{s} = 0$ or $\hat{t} = 0$.

If $\hat{s} = \hat{t} = 0$, then Lemmas 3.2(c) and 5.1 yield $C = C_L > 0$ such that

$$\Psi(0, 0) = \sup_{x,y \in \mathbb{R}^d} \{u_0(x) - u_0(y) - \Phi_\delta(x - y, 0, 0; \zeta)\}$$

$$\leq \sup_{x,y \in \mathbb{R}^d} \left\{ L|x - y| - \frac{1}{2(C+1)\delta}|x - y|^2 \right\} + C\delta \leq \left(C + \frac{(C+1)L^2}{2} \right)\delta.$$

Assume now that $\hat{s} = 0$. Then, in view of Lemmas 3.2(c) and 5.1,

$$\Psi(0, \hat{t}) = \sup_{|x-y| \leq M\delta} \{u_0(x) - v(y, \hat{t}) - \Phi_\delta(x - y, 0, \hat{t}; \zeta)\} - \frac{\hat{t}^2}{2\epsilon}$$

$$\leq C\delta + \sup_{|x-y| \leq M\delta} \{u_0(y) - v(y, \hat{t}) - \Phi_\delta(x - y, 0, \hat{t}; \zeta)\} - \frac{\hat{t}^2}{2\epsilon}$$

$$\leq C\delta + \max_{t \in [0,T]} \left(C\operatorname{osc}(\zeta, 0, t) - \frac{t^2}{2\epsilon} \right) = C\delta + \max_{s,t \in [0,T]} \left(C|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right).$$

Finally, if $\hat{t} = 0$, then Lemma 5.1 gives

$$\Psi(\hat{s}, 0) \leq \sup_{|x-y| \leq M\delta} \{v_h(x, \hat{s}) - u_0(y) - \Phi_\delta(x - y, \hat{s}, 0; \zeta)\} - \frac{\hat{s}^2}{2\epsilon}$$

$$\leq C\delta + \sup_{x \in \mathbb{R}^d} \{v_h(x, \hat{s}) - u_0(x)\} - \frac{\hat{s}^2}{2\epsilon}$$

$$\leq C(\delta + \sqrt{N}h) + \max_{s,t \in [0,T]} \left\{ C|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right\}.$$

Step 3. Combining the previous two steps and rearranging terms yields, for all $(x, t) \in \mathbb{R}^d \times [0, T]$,

$$v_h(x, t) - v(x, t)$$

$$\leq \frac{1}{\epsilon} \sum_{n=0}^{N-1} (\Delta t_n)^2 + C\left(\delta + \frac{Nh^2}{\delta} + \sqrt{N}h \right) + \max_{s,t \in [0,T]} \left\{ C|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right\}.$$

The inequality is optimized by setting

$$\delta := \max \left\{ C\sqrt{N}h, \max_{s,t \in [0,T]} \left(C|\zeta(s) - \zeta(t)| - \frac{|s - t|^2}{2\epsilon} \right) \right\}$$

for a sufficiently large constant $C = C_L > 0$, which clearly satisfies (5.10). This finishes the proof of the upper bound for $v_h - v$, and the lower bound is proved similarly. $\square$

6. Convergence rates.

In this section, the pathwise estimate from Theorem 5.1 is used to obtain a rate of convergence for schemes approximating solutions of the Hamilton–Jacobi equation

$$(6.1) \qquad du = \sum_{i=1}^{m} H^i(Du) \circ dW^i \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d.$$

It will always be assumed, as in Section 5, that

$$(6.2) \qquad \begin{cases} H \text{ and } u_0 \text{ satisfy } (3.2), (3.3), \text{ and } (3.4), \\ \text{and the scheme operator } S_h \text{ satisfies } (4.1), (4.2), (5.1), \text{ and } (5.2). \end{cases}$$

We first examine the setting in which W is a fixed, deterministic path, and then some extensions are presented in the case where W is a Brownian motion. Following Section 4, we define $u_h := v_h(\cdot; W_h, \mathcal{P}_h)$, with v_h as in (5.5), for an appropriate family of approximating paths $\{W_h\}_{h>0}$ and partitions $\{\mathcal{P}_h\}_{h>0}$. Let v be the viscosity solution of

$$(6.3) \qquad v_t = \sum_{i=1}^{m} H^i(Dv)\dot{W}_h \quad \text{in } \mathbb{R}^d \times (0, T] \quad \text{and} \quad v(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}^d.$$

The error $u_h - u$ is then controlled by using Theorem 5.1 and Lemma 3.3 to estimate respectively the differences $u_h - v$ and $v - u$.

6.1. *A fixed continuous path.* Fix $W \in C([0, T]; \mathbb{R}^m)$, and let $\omega : [0, \infty) \to [0, \infty)$ be its modulus of continuity. Define ρ_h implicitly by

$$(6.4) \qquad \lambda := \frac{\omega((\rho_h)^{1/2})}{(\rho_h)^{1/2}} < \lambda_0$$

and set $\mathcal{P}_h := \{n\rho_h \wedge T\}_{n=0}^{N}$, where N is the smallest integer for which $N\rho_h \wedge T = T$.

Recall from Section 4.1.1 that taking W_h to be the piecewise linear interpolation of W over the partition $\mathcal{P}_h$ may not, in general, yield a convergent scheme. Instead, we set

$$M_h := \lfloor (\rho_h)^{-1/2} \rfloor$$

and define W_h by

$$(6.5) \qquad W_h(t) := W(kM_h\rho_h) + \left(\frac{W((k+1)M_h\rho_h) - W(kM_h\rho_h)}{M_h\rho_h} \right)(t - kM_h\rho_h)$$

for $k \in \mathbb{N}_0$ and $t \in [kM_h\rho_h, (k+1)M_h\rho_h)$. Observe that the approximating path W_h satisfies (4.14) with $\eta_h = (\rho_h)^{1/2}$.

Now set $u_h := v_h(\cdot; W_h, \mathcal{P}_h)$, with v_h as in (5.5), and let v be the solution of (6.3).

THEOREM 6.1.   *Assume* (6.2) *and let* u_h *and* u *be as described above. Then there exists* $C = C_{L,\lambda} > 0$ *such that*

$$(6.6) \qquad \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - u(x,t)| \leq C(1+T)\omega((\rho_h)^{1/2}).$$

As an example, assume that $W \in C^\alpha([0, T], \mathbb{R}^m)$ and set

$$\lambda := [W]_{\alpha,T} \frac{(\rho_h)^{(1+\alpha)/2}}{h}.$$

Then, as long as $\lambda < \lambda_0$, the scheme converges with a rate of order $(\rho_h)^{\alpha/2} \approx h^{\alpha/(1+\alpha)}$.

PROOF OF THEOREM 6.1.   First, notice that, in view of (6.4), W_h satisfies (5.3). In particular, for some $C = C_L > 0$,

$$\max_{s,t\in[0,T]} \left(C|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\epsilon} \right)$$

$$\leq C\lambda_0 h + \max_{n\in\mathbb{N}_0} \left(Cn\lambda_0 h - \frac{n^2\rho_h^2}{2\epsilon} \right) \leq C\lambda_0 h + \frac{(C\lambda_0 h)^2 \epsilon}{2(\rho_h)^2}.$$

Theorem 5.1 then gives, for any $\epsilon > 0$,

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]}\big|u_h(x,t)-v(x,t)\big| \leq \frac{N(\rho_h)^2}{\epsilon}+C\sqrt{N}h+\frac{(C\lambda_0 h)^2\epsilon}{2(\rho_h)^2}$$

$$\leq \frac{T\rho_h}{\epsilon}+C\sqrt{T}\frac{h}{\sqrt{\rho_h}}+\frac{(C\lambda_0 h)^2\epsilon}{2(\rho_h)^2}.$$

Upon choosing $\epsilon = \sqrt{T}\frac{(\rho_h)^{3/2}}{h}$, this becomes

$$(6.7)\qquad \max_{(x,t)\in\mathbb{R}^d\times[0,T]}\big|u_h(x,t)-v(x,t)\big| \leq C\sqrt{T}\frac{h}{\sqrt{\rho_h}} = C\sqrt{T}\omega\big((\rho_h)^{1/2}\big).$$

Notice that the error term takes the form $\sqrt{\frac{h^2}{\rho_h}}$, which is consistent with the discussion in Section 4.1.1.

Lemma 3.3 then implies that

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]}\big|u_h(x,t)-u(x,t)\big| \leq C\big(\sqrt{T}\omega\big((\rho_h)^{1/2}\big)+\omega(M_h\rho_h)\big),$$

and the result is proved in view of the choice of M_h. $\square$

6.2. *Brownian paths.* For the rest of the paper, we investigate schemes for which W is a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$. The expectation and variance with respect to $\mathbf{P}$ are denoted by respectively $\mathbf{E}$ and $\mathbf{Var}$. To simplify the presentation, it is assumed that $m = 1$, so that W is one-dimensional, although all three schemes below can be adapted to the case when $m > 1$.

6.2.1. *Regular partitions.* Theorem 6.1 may be applied in this situation by using the fact that oscillations of Brownian paths are controlled by the Lévy modulus of continuity. More precisely,

$$(6.8)\qquad \mathbf{P}\bigg(\limsup_{\delta\to 0}\ \sup_{\delta\leq t\leq T-\delta}\frac{|W(t)-W(t+\delta)|}{\sqrt{2\delta|\log\delta|}}=1\bigg)=1.$$

THEOREM 6.2. *Assume* (6.2), *let* ρ_h *be defined implicitly by*

$$(6.9)\qquad \lambda := \frac{(\rho_h)^{3/4}|\log\rho_h|^{1/2}}{h} < \lambda_0,$$

and let u_h, $\mathcal{P}_h$, *and* W_h *be as in the previous subsection. Then there exists a deterministic constant* $C = C_{L,\lambda} > 0$ *such that*

$$\mathbf{P}\bigg(\limsup_{h\to 0}\ \sup_{(x,t)\in\mathbb{R}^d\times[0,T]}\frac{|u_h(x,t)-u(x,t)|}{(\rho_h)^{1/4}|\log\rho_h|^{1/2}}\leq C(1+T)\bigg)=1.$$

PROOF. Define $M_h := \lfloor(\rho_h)^{-1/2}\rfloor$ and $K_h := \lfloor T/(M_h\rho_h)\rfloor$. The definitions of W_h and λ give

$$\max_{n=0,1,2,\ldots,N-1}\frac{|W_h(n\rho_h)-W_h((n+1)\rho_h)|}{h}$$

$$=\max_{k=0,1,2,\ldots,K_h}\frac{|W(kM_h\rho_h)-W((k+1)M_h\rho_h|}{M_h h}$$

$$= \lambda \max_{k=0,1,2,\ldots,K_h} \frac{|W(kM_h\rho_h) - W((k+1)M_h\rho_h)|}{M_h(\rho_h)^{3/4}|\log\rho_h|^{1/2}}$$

$$\leq \lambda \frac{\max_{|s-t|\leq(\rho_h)^{1/2}} |W(s) - W(t)|}{(\rho_h)^{1/4}(1 - (\rho_h)^{1/2})|\log\rho_h|^{1/2}}.$$

Therefore, in view of (6.8), for any $\delta > 0$,

$$\mathbf{P}\left(\max_{n=0,1,2,\ldots,N-1} \frac{|W_h(n\rho_h) - W_h((n+1)\rho_h)|}{h} \leq \lambda\frac{1+\delta}{1 - (\rho_h)^{1/2}} \text{ for sufficiently small } h\right) = 1.$$

Taking $\delta \in (0, \lambda_0/\lambda - 1)$, this implies that

$$\mathbf{P}\left(\limsup_{h\to 0} \max_{n=0,1,2,\ldots,N-1} \frac{|W_h(n\rho_h) - W_h((n+1)\rho_h)|}{h} < \lambda_0\right) = 1,$$

so that, for some $h_0 > 0$,

$$\mathbf{P}\left(|W_h(n\rho_h) - W_h((n+1)\rho_h)| \leq \lambda_0 h \text{ for all } 0 < h < h_0 \text{ and } n = 0, 1, 2, \ldots, N_h - 1\right) = 1.$$

Shrinking h_0, if necessary, it may be concluded from (6.6) and (6.8) that

$$\mathbf{P}\left(\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x, t) - u(x, t)| \leq CT(\rho_h)^{1/4}|\log\rho_h|^{1/2} \text{ for all } 0 < h < h_0\right) = 1. \qquad \square$$

Observe that (6.9) implies that $\lim_{h\to 0} \frac{\log\rho_h}{\log h} = \frac{4}{3}$, so that the convergence rate in Theorem 6.2 can be rewritten as

$$(6.10) \qquad \limsup_{h\to 0} \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \frac{|u_h(x, t) - u(x, t)|}{h^{1/3}|\log h|^{1/3}} \leq C(1 + T).$$

6.2.2. *Random partitions.* For the next scheme, the partitions $\mathcal{P}_h$ are defined using a sequence of stopping times adapted to the filtration $\mathcal{F}_t$ of the Brownian motion W. By choosing the stopping times carefully to control the maximal oscillations of the Brownian paths, it is possible to recover the error estimate from Theorem 6.2.

For $h > 0$, define $\eta_h := h^{1/3}|\log h|^{-2/3}$, set $T_0 = T_0(h) := 0$, and, for $k \in \mathbb{N}_0$,

$$T_{k+1} = T_{k+1}(h) := \inf\{t > T_k(h) : \mathrm{osc}(W, T_k(h), t) > \eta_h\} \quad \text{and}$$

$$\tau_{k+1} = \tau_{k+1}(h) := T_{k+1}(h) - T_k(h).$$

Observe that $\{T_k\}_{k=0}^\infty$ is an increasing sequence of stopping times, and, for each fixed k, $h \to T_k(h)$ decreases as $h \to 0$. Therefore, by the strong Markov property for Brownian motion, for each fixed h, $\{\tau_k(h)\}_{k=1}^\infty$ is a collection of independent, identically distributed random variables. As a result, for any integer $\ell > 0$, there exists a constant $c_\ell > 0$ such that, for all k,

$$\mathbf{E}\left[\tau_k(h)^\ell\right] = c_\ell(\eta_h)^{2\ell}.$$

Indeed, it is well known that the first exit time of a Brownian motion from a bounded interval has finite moments of any order. The exact formula follows from the scaling properties of Brownian motion, so that

$$c_\ell := \mathbf{E}[\inf\{t > 0 : \mathrm{osc}(W, 0, t^{1/\ell}) > 1\}].$$

Let W_h be the piecewise interpolation of W over the partition $\{0 = T_0(h) < T_1(h) < T_2(h) < \cdots\}$. That is,

$$W_h(t) := W(T_k(h)) + \frac{W(T_{k+1}(h)) - W(T_k(h))}{\tau_{k+1}(h)}(t - T_k(h))$$

whenever $T_k(h) \leq t < T_{k+1}(h)$.

Define

$$M_h := \left\lceil \frac{\eta_h}{\lambda_0 h} \right\rceil = \left\lceil (\lambda_0 h^{2/3} |\log h|^{2/3})^{-1} \right\rceil,$$

$t_0 = t_0(h) := 0$, and, whenever $k = 0, 1, 2, \ldots$ and $k M_h \leq n < (k+1) M_h$,

$$t_n = t_n(h) := T_k(h) + (n - k M_h) \frac{\tau_{k+1}(h)}{M_h} \quad \text{and}$$

$$\Delta t_n = \Delta t_n(h) := t_{n+1}(h) - t_n(h) = \frac{\tau_{k+1}(h)}{M_h}.$$

Also set

$$K_h := \sup\{k \in \mathbb{N}_0 : T_k(h) \leq T\} \quad \text{and} \quad N_h := \sup\{n \in \mathbb{N}_0 : t_n(h) \leq T\},$$

and note that $h \mapsto K_h$ increases as $h \to 0$.

We have defined the path W_h, which is piecewise linear over the partition

$$\mathcal{P}_h := \{0 = t_0(h) < t_1(h) < t_2(h) < \cdots < t_{N_h}(h) \leq T\},$$

in such a way that (5.3) holds for $\zeta = W_h$. Indeed, if $n = 0, 1, 2, \ldots, N - 1$ and k is such that $k M_h \leq t_n < t_{n+1} \leq (k+1) M_h$, then

$$\left| W_h(t_{n+1}) - W_h(t_n) \right| = \frac{|W(T_{k+1}) - W(T_k)|}{M_h} \leq \lambda_0 h.$$

Finally, set $u_h := v_h(\cdot; W_h, \mathcal{P}_h)$ and let u be the stochastic viscosity solution of (6.1).

THEOREM 6.3. *Assume* (6.2), *and let u_h and u be as described above. Then there exists a deterministic constant $C = C_L > 0$ such that*

$$\mathbf{P}\left(\limsup_{h \to 0} \max_{(x,t) \in \mathbb{R}^d \times [0,T]} \frac{|u_h(x,t) - u(x,t)|}{h^{1/3} |\log h|^{1/3}} \leq C(1 + T) \right) = 1.$$

We proceed with a series of lemmas that indicate how to control the various terms appearing in the estimate from Theorem 5.1.

LEMMA 6.1.

$$\mathbf{P}\left(\limsup_{h \to 0} K_h \eta_h^2 \leq \frac{T}{c_1} \right) = 1.$$

PROOF. Fix α and β such that $1 < \beta^{2/3} < \alpha$, and define $h_m := \beta^{-m}$. Note that

$$\lim_{m \to \infty} \frac{\eta_{h_{m+1}}}{\eta_{h_m}} = \frac{1}{\beta^{1/3}}.$$

The monotonicity of K_h and η_h implies that

(6.11)
$$\mathbf{P}\left(\sup_{h_{m+1} \leq h < h_m} K_h \eta_h^2 > \frac{\alpha T}{c_1} \right) \leq \mathbf{P}\left(K_{h_{m+1}} > \frac{\alpha T}{c_1 \eta_{h_m}^2} \right).$$

Set

$$k_m := \left\lceil \frac{\alpha T}{c_1 \eta_{h_m}^2} \right\rceil,$$

so that

$$k_m c_1 \eta_{h_{m+1}}^2 \geq \alpha T \left(\frac{\eta_{h_{m+1}}}{\eta_{h_m}} \right)^2 \xrightarrow{m \to \infty} \alpha \beta^{-2/3} T > T,$$

and therefore, for any fixed $\gamma > 0$ and all sufficiently large m, $k_m c_1 \eta_{h_{m+1}}^2 \geq (1 + \gamma) T$.

Define $\sigma^2 := c_2 - c_1^2$, so that $\mathbf{Var}(\tau_k(h)) = \sigma^2 \eta_h^4$ for all k and h. Continuing (6.11) and applying Markov's inequality yields, for some fixed positive constant $C > 0$ and for all sufficiently large m,

$$\mathbf{P}\left(K_{h_{m+1}} > \frac{\alpha T}{c_1 \eta_{h_m}^2} \right) = \mathbf{P}\left(\sum_{k=1}^{k_m} \tau_k(h_{m+1}) \leq T \right) \leq \mathbf{P}\left(\sum_{k=1}^{k_m} (\tau_k(h_{m+1}) - c_1 \eta_{h_{m+1}}^2) \leq -\gamma T \right)$$

$$\leq \frac{k_m \sigma^2 \eta_{h_{m+1}}^4}{\gamma^2 T^2} \leq C\beta^{-2m/3}.$$

The Borel–Cantelli lemma applied to the events

$$E_m := \left\{ \sup_{h_{m+1} \leq h < h_m} K_h \eta_h^2 > \frac{\alpha T}{c_1} \right\}$$

gives

$$\mathbf{P}\left(\limsup_{h \to 0} K_h \eta_h^2 > \frac{\alpha T}{c_1} \right) = \mathbf{P}\left(\limsup_{m \to \infty} E_m \right) = 0,$$

and we may conclude upon sending $\alpha \to 1^+$. $\quad \square$

LEMMA 6.2.

$$\mathbf{P}\left(\limsup_{h \to 0} \frac{1}{h \eta_h} \sum_{n=0}^{N-1} (\Delta t_n)^2 \leq \frac{T \lambda_0 c_2}{c_1} \right) = 1.$$

PROOF. Fix α and β satisfying $1 < \beta^{7/3} < \alpha$ and set $h_m := \beta^{-m}$. If, for some m, $h_{m+1} \leq h < h_m$, then

$$\sum_{n=0}^{N_h-1} (\Delta t_n(h))^2 \leq \sum_{k=1}^{K_h+1} M_h \left(\frac{\tau_k(h)}{M_h} \right)^2 \leq \lambda_0 \frac{h_m}{\eta_{h_{m+1}}} \sum_{k=1}^{K_{h_{m+1}}+1} \tau_k(h_m)^2.$$

Fix $m_0 \in \mathbb{N}$ and define the event

$$E_{m_0} := \left\{ K_{h_{m+1}} + 1 \leq \hat{K}_m := \left\lceil \frac{\alpha T}{c_1 \eta_{h_{m+1}}^2} \right\rceil \text{ for all } m \geq m_0 \right\}.$$

In view of Lemma 6.1, $\lim_{m_0 \to \infty} \mathbf{P}(E_{m_0}) = 1$.

Now, for any $m \geq m_0$,

$$\mathbf{P}\left(\left\{ \sup_{h_{m+1} \leq h < h_m} \frac{1}{h \eta_h} \sum_{n=0}^{N_h-1} (\Delta t_n(h))^2 > \frac{\alpha^2 T \lambda_0 c_2}{c_1} \right\} \cap E_{m_0} \right)$$

$$\leq \mathbf{P}\left(\sum_{k=1}^{\hat{K}_m} \tau_k(h_m)^2 > \frac{\alpha^2 T c_2 h_{m+1} (\eta_{h_{m+1}})^2}{c_1 h_m} \right)$$

$$= \mathbf{P}\left(\sum_{k=1}^{\hat{K}_m} (\tau_k(h_m)^2 - c_2 \eta_{h_m}^4) > \frac{\alpha^2 T c_2 h_{m+1} (\eta_{h_{m+1}})^2}{c_1 h_m} - \hat{K}_m c_2 \eta_{h_m}^4 \right)$$

$$\leq \mathbf{P}\left(\sum_{k=1}^{\hat{K}_m}(\tau_k(h_m)^2 - c_2\eta_{h_m}^4) > \frac{\alpha T c_2\eta_{h_m}^2}{c_1}\left(\frac{\alpha h_{m+1}(\eta_{h_{m+1}})^2}{h_m(\eta_{h_m})^2} - \frac{\eta_{h_m}^2}{\eta_{h_{m+1}}^2}\right) - c_2\eta_{h_m}^4\right).$$

Since

$$\lim_{m\to\infty}\left(\frac{\alpha h_{m+1}\eta_{h_{m+1}}^2}{h_m\eta_{h_m}^2} - \frac{\eta_{h_m}^2}{\eta_{h_{m+1}}^2}\right) = \frac{\alpha}{\beta^{5/3}} - \beta^{2/3} > 0,$$

it follows that, for some fixed $\gamma > 0$, all sufficiently large m_0, and all $m > m_0$,

$$\mathbf{P}\left(\left\{\sup_{h_{m+1}\leq h<h_m}\frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\} \cap E_{m_0}\right)$$

$$\leq \mathbf{P}\left(\sum_{k=1}^{\hat{K}_m}(\tau_k(h_m)^2 - c_2\eta_{h_m}^4) > \gamma\eta_{h_m}^2\right).$$

Set $\sigma^2 := c_4 - c_2^2 > 0$. Then Markov's inequality gives, for some constant $C > 0$ independent of m,

$$\mathbf{P}\left(\left\{\sup_{h_{m+1}\leq h<h_m}\frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\} \cap E_{m_0}\right)$$

$$\leq \frac{\hat{K}_m\sigma^2\eta_{h_m}^4}{\gamma^2} \leq C\eta_{h_m}^2 \leq C\beta^{-2m/3}.$$

An application of the Borel–Cantelli lemma for the events

$$\left\{\sup_{h_{m+1}\leq h<h_m}\frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\} \cap E_{m_0}$$

yields

$$\mathbf{P}\left(\left\{\limsup_{h\to0}\frac{1}{h\eta_h}\sum_{n=0}^{N_h-1}(\Delta t_n(h))^2 > \frac{\alpha^2 T\lambda_0 c_2}{c_1}\right\} \cap E_{m_0}\right) = 0.$$

Sending $m_0 \to \infty$ and then $\alpha \to 1^+$ finishes the proof. $\square$

LEMMA 6.3. *For any deterministic constant $C > 0$,*

$$\mathbf{P}\left(\limsup_{\epsilon\to0}\frac{\max_{s,t\in[0,T]}\{C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon}\}}{\epsilon^{1/3}|\log\epsilon|^{2/3}} \leq \frac{4C^{4/3}}{3^{2/3}}\right) = 1.$$

PROOF. Let $1 < \beta < \alpha$. If, for some $\delta > 0$,

(6.12) $\operatorname{osc}(W, k\delta, (k+1)\delta) \leq \sqrt{2\beta}\delta^{1/2}|\log\delta|^{1/2}$ for all $k = 0, 1, 2, \ldots, \left\lceil\frac{T}{\delta}\right\rceil$,

then

$$\max_{s,t\in[0,T]}\left\{C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon}\right\}$$

$$\leq \sqrt{2\beta}C\delta^{1/2}|\log\delta|^{1/2} + \max_{n\in\mathbb{N}_0}\left\{\sqrt{2\beta}C\delta^{1/2}|\log\delta|^{1/2}n - \frac{n^2\delta^2}{2\epsilon}\right\}$$

$$\leq \sqrt{2\beta}C\delta^{1/2}|\log\delta|^{1/2} + \beta C^2\frac{|\log\delta|}{\delta}\epsilon.$$

Taking $\delta := C^{2/3} 3^{-1/3} \beta^{1/3} \epsilon^{2/3} |\log \epsilon|^{1/3}$ yields, for some deterministic function $c(\epsilon) \xrightarrow{\epsilon \to 0} 0$,

$$\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon} \right\} \leq \frac{4C^{4/3}}{3^{2/3}} \beta^{2/3} \epsilon^{1/3} |\log \epsilon|^{2/3} (1 + c(\epsilon)),$$

and, therefore, if ϵ is sufficiently small,

$$\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon} \right\} \leq \frac{4C^{4/3}}{3^{2/3}} \alpha^{2/3} \epsilon^{1/3} |\log \epsilon|^{2/3}.$$

Define

$$\epsilon_m := \alpha^{-m} \quad \text{and} \quad \delta_m := \frac{C^{2/3} \beta^{1/3} \epsilon_m^{2/3} |\log \epsilon_m|^{1/3}}{3^{1/3}},$$

and note that

$$\lim_{m \to \infty} \frac{\epsilon_m^{1/3} |\log \epsilon_m|^{2/3}}{\epsilon_{m+1}^{1/3} |\log \epsilon_{m+1}|^{2/3}} = \alpha^{1/3}.$$

It follows that, for sufficiently large m,

$$\mathbf{P}\left(\sup_{\epsilon_{m+1} \leq \epsilon < \epsilon_m} \frac{\max_{s,t \in [0,T]} \{C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon}\}}{\epsilon^{1/3} |\log \epsilon|^{2/3}} > \frac{4C^{4/3}\alpha}{3^{2/3}} \right)$$

$$\leq \mathbf{P}\left(\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon_m} \right\} > \frac{4C^{4/3}\alpha \epsilon_{m+1}^{1/3} |\log \epsilon_{m+1}|^{2/3}}{3^{2/3}} \right)$$

$$\leq \mathbf{P}\left(\max_{s,t \in [0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon_m} \right\} > \frac{4C^{4/3}\alpha^{2/3} \epsilon_m^{1/3} |\log \epsilon_m|^{2/3}}{3^{2/3}} \right)$$

$$\leq \mathbf{P}\left(\text{osc}(W, k\delta_m, (k+1)\delta_m) > \sqrt{2\beta} \delta_m^{1/2} |\log \delta_m|^{1/2} \text{ for some } k = 0, 1, 2, \ldots, \left\lceil \frac{T}{\delta_m} \right\rceil \right)$$

$$\leq \left\lceil \frac{T}{\delta_m} \right\rceil \mathbf{P}\left(\max_{[0,1]} W - \min_{[0,1]} W > \sqrt{2\beta} |\log \delta_m|^{1/2} \right)$$

$$\leq 2 \left\lceil \frac{T}{\delta_m} \right\rceil \mathbf{P}\left(\max_{[0,1]} W > \sqrt{2\beta} |\log \delta_m|^{1/2} \right)$$

$$\leq CT \delta_m^{\beta - 1} \leq CT \alpha^{-\gamma m} \quad \text{for } \gamma = \frac{2}{3}(\beta - 1) > 0.$$

The symmetry and scaling properties of Brownian motion, as well as the reflection principle, were all used above. In particular, since the processes

$$\left\{ t \mapsto \max_{s \in [0,t]} W(s) - W(t) \right\} \quad \text{and} \quad |W|$$

are identically distributed, so are the random variables

$$\max_{[0,1]} W - \min_{[0,1]} W \quad \text{and} \quad \max_{[0,1]} |W| = \max\left\{ \max_{[0,1]} W, -\min_{[0,1]} W \right\}.$$

The Borel–Cantelli lemma implies that

$$\mathbf{P}\left(\limsup_{\epsilon \to 0} \frac{\max_{s,t \in [0,T]} \{C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon}\}}{\epsilon^{1/3} |\log \epsilon|^{2/3}} > \frac{4C^{4/3}\alpha}{3^{2/3}} \right) = 0,$$

and sending $\alpha \to 1^+$ gives the result. $\quad \square$

PROOF OF THEOREM 6.3. Let v be the solution of (6.3). Then Lemma 3.3 gives, for some $C = C_L > 0$,

$$\sup_{(x,t)\in\mathbb{R}^d\times[0,T]} |v(x,t) - u(x,t)| \le C|W(t) - W_h(t)|$$

$$\le C \max_{k=0,1,2,\dots,K_h} \mathrm{osc}(W, T_k(h), T_{k+1}(h))$$

$$\le C \frac{h^{1/3}}{|\log h|^{2/3}}.$$

Next, define $\epsilon_h := \frac{h}{|\log h|}$ and recall the pathwise estimate from Theorem 5.1:

$$\max_{(x,t)\in\mathbb{R}^d\times[0,T]} |u_h(x,t) - v(x,t)|$$

$$\le \frac{1}{\epsilon_h} \sum_{n=0}^{N_h-1} (\Delta t_n(h))^2 + C\sqrt{N_h}h + \max_{s,t\in[0,T]} \left\{ C|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\epsilon_h} \right\}.$$

From the definitions of N_h, M_h, and K_h, and from Lemma 6.1, it follows that, for some $C = C_L > 0$, with probability one, for all sufficiently small h,

$$C\sqrt{N_h}h \le C\sqrt{(K_h + 1)M_h}h \le CTh^{1/3}|\log h|^{1/3}.$$

Meanwhile, Lemma 6.2 yields $C = C_L > 0$ such that, with probability one, for all sufficiently small h,

$$\frac{1}{\epsilon_h} \sum_{n=0}^{N_h-1} (\Delta t_n(h))^2 = \frac{1}{h\eta_h} \sum_{n=0}^{N_h-1} (\Delta t_n)^2 \cdot h^{1/3}|\log h|^{1/3} \le CTh^{1/3}|\log h|^{1/3}.$$

In view of the definition of W_h,

$$\max_{0\le t\le T} |W_h(t) - W(t)| \le \max_{k=0,1,2,\dots,K_h} \mathrm{osc}(W, T_k(h), T_{k+1}(h)) \le \eta_h,$$

so that, with probability one, for all h,

$$\frac{\max_{s,t\in[0,T]}\{C|W_h(s) - W_h(t)| - \frac{|s-t|^2}{2\epsilon_h}\}}{h^{1/3}|\log h|^{1/3}}$$

$$\le \frac{\max_{s,t\in[0,T]}\{C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon_h}\}}{h^{1/3}|\log h|^{1/3}} + \frac{C}{|\log h|},$$

while Lemma 6.3 implies that, with probability one and for all sufficiently small h,

$$\max_{s,t\in[0,T]} \left\{ C|W(s) - W(t)| - \frac{|s-t|^2}{2\epsilon_h} \right\} \le C\epsilon_h^{1/3}|\log \epsilon_h|^{2/3} \le Ch^{1/3}|\log h|^{1/3}.$$

Combining all terms in the estimate finishes the proof. □

6.2.3. *Scaled random walks and convergence in law.* The point of view for the preceding approximations was pathwise; that is, the schemes converged for **P**-almost every sample path of Brownian motion. Here, the strategy is to use independent Rademacher random variables to build an object that converges to the solution of (1.1) in distribution. This construction has the advantage that it is simple to implement numerically.

Fix a probability space $(\mathcal{A}, \mathcal{G}, \mathbb{P})$, not necessarily related to $(\Omega, \mathcal{F}, \mathbf{P})$, and let $\{\xi_n\}_{n=1}^{\infty} : \mathcal{A} \to \{-1, 1\}$ be independent and identically distributed with

$$\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = \frac{1}{2}.$$

Define ρ_h by

$$\lambda := \frac{(\rho_h)^{3/4}}{h} \leq \lambda_0,$$

and, as before, set $M_h = \lfloor (\rho_h)^{-1/2} \rfloor$, $\mathcal{P}_h := \{t_n\}_{n=0}^{N} = \{n\rho_h \wedge T\}_{n=0}^{N}$, $W_h(0) = 0$ and, for $k \in \mathbb{N}_0$ and $t \in [kM_h\rho_h, (k+1)M_h\rho_h)$,

$$W_h(t) := W_h(kM_h\rho_h) + \frac{\xi_k}{\sqrt{M_h\rho_h}}(t - kM_h\rho_h).$$

The path W_h is a parabolically scaled simple random walk, and therefore, as is well known, as $h \to 0$, W_h converges to the Wiener process B in distribution. More precisely, if μ is the Wiener measure on $X := C([0, T], \mathbb{R})$ and μ_h is the probability measure on X induced by W_h, then μ_h converges weakly to μ as $h \to 0$, that is, for any bounded continuous function $\phi : X \to \mathbb{R}$,

$$\lim_{h \to 0} \int_X \phi \, d\mu_h = \int_X \phi \, d\mu.$$

Define $u_h := v_h(\cdot; W_h, \mathcal{P}_h) \in BUC(\mathbb{R}^d \times [0, T])$ and let $v \in BUC(\mathbb{R}^d \times [0, T])$ be the solution of (6.3).

THEOREM 6.4. *Assume (6.2) and let u_h and u be as described above. As $h \to 0$, u_h converges to u in distribution.*

PROOF. Observe first that

$$\left|W_h(t_{n+1}) - W_h(t_n)\right| = \sqrt{\frac{\rho_h}{M_h}} \leq (\rho_h)^{3/4} \leq \lambda_0 h,$$

so that W_h satisfies (5.3). Then (6.6) becomes, for some $C = C_{L,\lambda} > 0$,

$$(6.13) \qquad \max_{(x,t) \in \mathbb{R}^d \times [0,T]} \left|u_h(x, t) - v(x, t)\right| \leq C(1 + T)(\rho_h)^{1/4} = C(1 + T)h^{1/3}.$$

Lemma 3.3 implies that the map

$$S : X = C([0, T], \mathbb{R}) \ni \zeta \mapsto v \in BUC(\mathbb{R}^d \times [0, T]) =: Y,$$

where v is the solution of (5.4), is uniformly continuous. Let $\tilde{v}_h$ and v be the push-forwards by S of respectively μ_h and μ, that is, for any measurable $\psi : Y \to \mathbb{R}$,

$$\int_Y \psi \, dv = \int_X \psi \circ S \, d\mu,$$

with the analogous relation holding for v_h and v. It is clear that $\tilde{v}_h$ converges weakly to v. On the other hand, if v_h is the measure on $BUC(\mathbb{R}^d \times [0, T])$ induced by u_h, then (6.13) and Slutzky's theorem imply that, as $h \to 0$, v_h converges weakly to v. $\square$

REFERENCES

[1] BARLES, G. and JAKOBSEN, E. R. (2002). On the convergence rate of approximation schemes for Hamilton–Jacobi–Bellman equations. *M2AN Math. Model. Numer. Anal.* **36** 33–54. MR1916291 https://doi.org/10.1051/m2an:2002002

[2] BARLES, G. and JAKOBSEN, E. R. (2005). Error bounds for monotone approximation schemes for Hamilton–Jacobi–Bellman equations. *SIAM J. Numer. Anal.* **43** 540–558. MR2177879 https://doi.org/10.1137/S003614290343815X

[3] BARLES, G. and JAKOBSEN, E. R. (2007). Error bounds for monotone approximation schemes for parabolic Hamilton–Jacobi–Bellman equations. *Math. Comp.* **76** 1861–1893. MR2336272 https://doi.org/10.1090/S0025-5718-07-02000-5

[4] BARLES, G. and SOUGANIDIS, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.* **4** 271–283. MR1115933

[5] BUCKDAHN, R. and MA, J. (2001). Stochastic viscosity solutions for nonlinear stochastic partial differential equations. I. *Stochastic Process. Appl.* **93** 181–204. MR1828772 https://doi.org/10.1016/S0304-4149(00)00093-4

[6] BUCKDAHN, R. and MA, J. (2001). Stochastic viscosity solutions for nonlinear stochastic partial differential equations. II. *Stochastic Process. Appl.* **93** 205–228. MR1831830 https://doi.org/10.1016/S0304-4149(00)00092-2

[7] CAFFARELLI, L. A. and SOUGANIDIS, P. E. (2008). A rate of convergence for monotone finite difference approximations to fully nonlinear, uniformly elliptic PDEs. *Comm. Pure Appl. Math.* **61** 1–17. MR2361302 https://doi.org/10.1002/cpa.20208

[8] CARUANA, M., FRIZ, P. K. and OBERHAUSER, H. (2011). A (rough) pathwise approach to a class of nonlinear stochastic partial differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **28** 27–46. MR2765508 https://doi.org/10.1016/j.anihpc.2010.11.002

[9] COURANT, R., FRIEDRICHS, K. and LEWY, H. (1967). On the partial difference equations of mathematical physics. *IBM J. Res. Develop.* **11** 215–234. MR0213764 https://doi.org/10.1147/rd.112.0215

[10] CRANDALL, M. G., ISHII, H. and LIONS, P.-L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **27** 1–67. MR1118699 https://doi.org/10.1090/S0273-0979-1992-00266-5

[11] CRANDALL, M. G. and LIONS, P.-L. (1984). Two approximations of solutions of Hamilton–Jacobi equations. *Math. Comp.* **43** 1–19. MR0744921 https://doi.org/10.2307/2007396

[12] FRIZ, P. K., GASSIAT, P., LIONS, P.-L. and SOUGANIDIS, P. E. (2017). Eikonal equations and pathwise solutions to fully non-linear SPDEs. *Stoch. Partial Differ. Equ. Anal. Comput.* **5** 256–277. MR3640072 https://doi.org/10.1007/s40072-016-0087-9

[13] GASSIAT, P. and GESS, B. (2019). Regularization by noise for stochastic Hamilton–Jacobi equations. *Probab. Theory Related Fields* **173** 1063–1098. MR3936151 https://doi.org/10.1007/s00440-018-0848-7

[14] GUBINELLI, M., TINDEL, S. and TORRECILLA, I. Controlled viscosity solutions of fully nonlinear rough PDEs. Available at arXiv:1403.2832 [math.PR].

[15] HOEL, H., KARLSEN, K. H., RISEBRO, N. H. and STORRØSTEN, E. B. (2019). Numerical methods for conservation laws with rough flux. *Stoch PDE: Anal Comp.* https://doi.org/10.1007/s40072-019-00145-7

[16] ISHII, H. (1985). Hamilton–Jacobi equations with discontinuous Hamiltonians on arbitrary open sets. *Bull. Fac. Sci. Engrg. Chuo Univ.* **28** 33–77. MR0845397

[17] JAKOBSEN, E. R. (2004). On error bounds for approximation schemes for non-convex degenerate elliptic equations. *BIT* **44** 269–285. MR2093506 https://doi.org/10.1023/B:BITN.0000039390.33444.f2

[18] JAKOBSEN, E. R. (2006). On error bounds for monotone approximation schemes for multi-dimensional Isaacs equations. *Asymptot. Anal.* **49** 249–273. MR2270893

[19] KRYLOV, N. V. (2015). On the rate of convergence of finite-difference approximations for elliptic Isaacs equations in smooth domains. *Comm. Partial Differential Equations* **40** 1393–1407. MR3355497 https://doi.org/10.1080/03605302.2015.1029074

[20] KUO, H.-J. and TRUDINGER, N. S. (1996). Positive difference operators on general meshes. *Duke Math. J.* **83** 415–433. MR1390653 https://doi.org/10.1215/S0012-7094-96-08314-3

[21] KUO, H. J. and TRUDINGER, N. S. (1992). Discrete methods for fully nonlinear elliptic equations. *SIAM J. Numer. Anal.* **29** 123–135. MR1149088 https://doi.org/10.1137/0729008

[22] LIONS, P.-L. and PERTHAME, B. (1987). Remarks on Hamilton–Jacobi equations with measurable time-dependent Hamiltonians. *Nonlinear Anal.* **11** 613–621. MR0886652 https://doi.org/10.1016/0362-546X(87)90076-9

[23] LIONS, P.-L. and SOUGANIDIS, P. E. Forthcoming book on fully nonlinear stochastic partial differential equations.

[24] LIONS, P.-L. and SOUGANIDIS, P. E. (1998). Fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.* **326** 1085–1092. MR1647162 https://doi.org/10.1016/S0764-4442(98)80067-0

[25] LIONS, P.-L. and SOUGANIDIS, P. E. (1998). Fully nonlinear stochastic partial differential equations: Non-smooth equations and applications. *C. R. Acad. Sci. Paris Sér. I Math.* **327** 735–741. MR1659958 https://doi.org/10.1016/S0764-4442(98)80161-4

[26] LIONS, P.-L. and SOUGANIDIS, P. E. (2000). Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.* **331** 783–790. MR1807189 https://doi.org/10.1016/S0764-4442(00)01597-4

[27] LIONS, P.-L. and SOUGANIDIS, P. E. (2000). Fully nonlinear stochastic pde with semilinear stochastic dependence. *C. R. Acad. Sci. Paris Sér. I Math.* **331** 617–624. MR1799099 https://doi.org/10.1016/S0764-4442(00)00583-8

[28] NUNZIANTE, D. (1990). Uniqueness of viscosity solutions of fully nonlinear second order parabolic equations with discontinuous time-dependence. *Differential Integral Equations* **3** 77–91. MR1014727

[29] SEEGER, B. (2018). Homogenization of pathwise Hamilton–Jacobi equations. *J. Math. Pures Appl.* (9) **110** 1–31. MR3744918 https://doi.org/10.1016/j.matpur.2017.07.012

[30] SEEGER, B. (2018). Perron's method for pathwise viscosity solutions. *Comm. Partial Differential Equations* **43** 998–1018. MR3909032 https://doi.org/10.1080/03605302.2018.1488262

[31] SOUGANIDIS, P. E. Fully nonlinear first- and second-order stochastic partial differential equations. Preprint. Available at arXiv:1809.01748 [math.AP].

[32] SOUGANIDIS, P. E. (1985). Approximation schemes for viscosity solutions of Hamilton–Jacobi equations. *J. Differential Equations* **59** 1–43. MR0803085 https://doi.org/10.1016/0022-0396(85)90136-6

[33] SOUGANIDIS, P. E. (1985). Max-min representations and product formulas for the viscosity solutions of Hamilton–Jacobi equations with applications to differential games. *Nonlinear Anal.* **9** 217–257. MR0784388 https://doi.org/10.1016/0362-546X(85)90062-8

[34] SOUGANIDIS, P. E. and YIP, N. K. (2004). Uniqueness of motion by mean curvature perturbed by stochastic noise. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** 1–23. MR2037245 https://doi.org/10.1016/S0294-1449(03)00029-5

[35] TURANOVA, O. (2015). Error estimates for approximations of nonhomogeneous nonlinear uniformly elliptic equations. *Calc. Var. Partial Differential Equations* **54** 2939–2983. MR3412399 https://doi.org/10.1007/s00526-015-0890-6