MATHEMATICS OF COMPUTATION
VOLUME 52, NUMBER 186
APRIL 1989, PAGES 615–645

Convergence of a Random Particle Method to Solutions of the Kolmogorov Equation $u_t = \nu u_{xx} + u(1-u)$

By Elbridge Gerry Puckett*

Dedicated to Professor Eugene Isaacson on the occasion of his 70th birthday

Abstract. We study a random particle method for solving the reaction-diffusion equation $u_t = \nu u_{xx} + f(u)$ which is a one-dimensional analogue of the random vortex method. It is a fractional step method in which $u_t = \nu u_{xx}$ is solved by random walking the particles while $u_t = f(u)$ is solved with a numerical ordinary differential equation solver such as Euler's method. We prove that the method converges when $f(u) = u(1-u)$, i.e. the Kolmogorov equation, and that when the time step Δt is $O(\sqrt[4]{N^{-1}})$ the rate of convergence is like $\ln N \cdot \sqrt[4]{N^{-1}}$ where N denotes the number of particles. Furthermore, we show that this rate of convergence is uniform as the diffusion coefficient ν tends to 0. Thus, travelling waves with arbitrarily steep wavefronts may be modeled without an increase in the computational cost. We also present the results of numerical experiments including the use of second-order time discretization and second-order operator splitting and use these results to estimate the expected value and standard deviation of the error.

1. Introduction. In this paper we study a random particle method due to Chorin [9] for approximating solutions of the one-dimensional reaction-diffusion equation,

$$u_t = \nu u_{xx} + f(u), \tag{1.1a}$$

$$u(x,0) = u^0(x), \tag{1.1b}$$

where the forcing function, $f(u)$, satisfies

$$f(0) = f(1) = 0, \tag{1.2a}$$

$$f(u) > 0 \qquad \text{for } 0 < u < 1, \tag{1.2b}$$

$$f'(u) \le 1 \qquad \text{for } 0 \le u \le 1. \tag{1.2c}$$

We call this method the *random gradient method.* Algorithms based on this method have been used to solve Nagumo's equation [33] and the Hodgkin-Huxley equations

Received May 20, 1987; revised July 6, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 35–04, 35K57, 60J15, 60J65, 65–04, 65C05, 65C20, 65M15, 65N15.

Key words and phrases. Kolmogorov equation, particle method, random vortex method, random walk.

*This work was part of the author's Ph.D. Dissertation in the Mathematics Department at the University of California, Berkeley. Portions of this work were performed under the auspices of the U.S. Department of Energy at Lawrence Berkeley Laboratory under contract number DE-AC03-76SF00098 and Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

[34]. We prove the convergence of the random gradient method to solutions of the Kolmogorov equation,

$$u_t = \nu u_{xx} + u(1-u), \tag{1.3a}$$

$$u(x,0) = u^0(x), \tag{1.3b}$$

subject to the constraints

$$0 \le u^0(x) \le 1, \tag{1.3c}$$

$$\lim_{x \to -\infty} u^0(x) = 1, \tag{1.3d}$$

$$\lim_{x \to +\infty} u^0(x) = 0. \tag{1.3e}$$

Our work follows that of Roberts [31] who proved the convergence of a random particle method to Burgers' equation, $u_t + uu_x = \nu u_{xx}$. Related theoretical work includes [6], [13], [21], [22], [28], [29], [32]. A review of particle methods which use random walks to model diffusion may be found in [15].

Our interest in the random gradient method is primarily motivated by the fact that it is a one-dimensional analogue of the random vortex method [7] for approximating solutions of the Navier-Stokes equations. We hope that a thorough examination of the errors obtained when using the random gradient method will yield a greater understanding of the error inherent in using the random vortex method, particularly the error due to the random walk. In order to motivate the subsequent discussion, we list here the most important characteristics that these two methods have in common.

(i) Both are *particle methods*, with the particles representing point concentrations of some derivative of the solution. (The gradient of u in the case of the random gradient method, vorticity in the case of the random vortex method.)

(ii) Both are *splitting* or *fractional step methods*. That is, the equation to be solved is split into two evolution equations, each of which is solved separately. This process is coupled by using the solution obtained after solving one of the evolution equations as the initial data for the other.

(iii) In both methods one of the fractional steps is the heat equation, $u_t = \nu \Delta u$. In each method the numerical solution to the heat equation is obtained by random walking the particles.

(iv) Finally, in both methods the second of the fractional steps is a nonlinear evolution equation. In the case of the random gradient method this is the *reaction equation* $u_t = f(u)$, whereas for the random vortex method it is the Euler equations.

Similar analogies may be drawn between the present method and the vortex sheet method [8] for approximating solutions of the Prandtl boundary layer equations. Numerical estimates of the convergence rate for the random vortex method have been given by Roberts [30] while convergence proofs for the method in the absence of boundaries may be found in [16], [26]. Theoretical work on the vortex method solution of the Euler equations includes [1], [3], [4], [12], [17], [19], [20], [23].

In our treatment of the random gradient method particles are not permitted to divide in two when their strengths surpass some critical value, as was originally proposed by Chorin. This greatly simplifies the convergence proof. Difficulties

which arise while trying to prove convergence for the algorithm with particle creation are very similar to those which arise when attempting to prove convergence of the random vortex method in the presence of boundaries. In this case, particle creation corresponds to the creation of vorticity, an important phenomenon in fluid flow. Hald [22] has proven the convergence of a method with particle creation for solving a one-dimensional diffusion equation with thermal convection.

For a random variable Z let $E[Z]$ denote the expected value of Z and $\mathrm{var}(Z)$ its variance. The main result of this paper is the following.

THEOREM 1.1. *Assume* $0 < \nu \leq 1$. *Fix* $T > 0$ *and choose a time step* $0 < \Delta t \leq 1$ *such that* $T = k\Delta t$ *for some integer* k. *Let* $u(x,T)$ *be the solution at time* T *of* (1.3a–e) *with initial data* u^0, *and let* $\tilde{u}^k(x)$ *be the corresponding computed solution with initial data* $\tilde{u}^0$. *Let* $N \geq 10$ *denote the number of particles used to generate* $\tilde{u}^k$, *and assume that* $\Delta t = O(\sqrt[4]{N}^{-1})$. *Then there exist positive constants* C_0, C_1 *and* C_2, *independent of* ν, Δt, *and* N, *such that*

$$E(\|u(T) - \tilde{u}^k\|_{L^1}) \leq \left(1 + \frac{T}{C_0}\right)\left[e^T\|u^0 - \tilde{u}^0\|_{L^1} + C_1\sqrt{\nu}\Delta t + C_2\frac{\ln N}{\sqrt[4]{N}}\right], \tag{1.4}$$

$$\mathrm{var}(\|u(T) - \tilde{u}^k\|_{L^1}) \leq \left(1 + \frac{T}{C_0}\right)\left[e^T\|u^0 - \tilde{u}^0\|_{L^1} + C_1\sqrt{\nu}\Delta t + C_2\frac{\ln N}{\sqrt[4]{N}}\right]^2. \tag{1.5}$$

In order to prove this theorem, several assumptions regarding u^0 and $\tilde{u}^0$ have been made. In addition to satisfying the constraints (1.3c–e), it has been assumed that u^0 is continuously differentiable on $\mathbf{R}$ and $u_x^0 \in L^1(\mathbf{R}) \cap L^\infty(\mathbf{R})$. The approximate initial data $\tilde{u}^0$ is a step function approximation to u^0 and is required to be monotonically decreasing. All hypotheses are listed in Subsection 8.1. We first prove the theorem for $\nu = 1$ and then use a simple scaling argument to demonstrate the validity of the result for $\nu \leq 1$. One of the most important consequences of Theorem 1.1 is that the error is independent of the diffusion coefficient, or 'viscosity', ν. Thus, solutions with arbitrarily steep wavefronts may be modeled without any increase in the computational cost.

The details of the random gradient method are developed in Section 2, beginning with some notation and followed by the algorithm itself. In Subsection 2.3 the class, $\mathbf{S}$, of permissible starting approximations $\tilde{u}^0$ is defined and several preliminary lemmas are proved. Most of the error analysis is written in the language of solution operators. This notation is introduced in Subsection 3.1. A brief account of the proof may be found in Subsection 3.2, together with a description of how the details are divided among Sections 4–7. In Section 8 we put the various parts together and prove the theorem. Finally, in Section 9, we use the numerical method to compute a known exact solution of the Kolmogorov equation. This permits us to compare the convergence rate predicted by the proof with that obtained during an actual calculation.

2. A Description of the Random Gradient Method. We begin this section with the introduction of some notation and a description of the algorithm. This is

followed by a discussion of the difficulties that are encountered for nonmonotonic initial data and the proof of several basic facts that hold for monotonic initial data.

2.1. *Step Function Notation.* We will denote the numerical approximation of a function which is intended to be obtained on a computer by the symbol "~". Thus, $\tilde{u}(x,t)$ denotes an approximation to the solution $u(x,t)$ of Eq. (1.1a,b). We will use the term step function to refer to any piecewise constant function of $x \in \mathbf{R}$ that has a finite number of discontinuities.

In the random gradient method, $\tilde{u}$ is a step function approximation to u. Consequently, knowledge of the position of each discontinuity and of the amount of each jump is all that is required in order to know $\tilde{u}$. It is convenient to think of $\tilde{u}$ at a given time t as being represented by N particles. Each particle has associated with it a position on the x-axis and a *strength* or *weight*, the particle's position being a point at which $\tilde{u}$ is discontinuous and its strength being the amount by which $\tilde{u}$ changes at that point. The position of the ith particle at time $t = j\Delta t$ will be denoted by X_i^j and its weight by w_i^j. We denote the computed solution after j time steps as $\tilde{u}^j(x)$ and write

$$\tilde{u}^j(x) = \sum_{i=1}^{N} H(X_i^j - x) w_i^j, \tag{2.1}$$

where $H(x)$ is the Heaviside function

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \tag{2.2}$$

We assume that the particles have been labeled so that for each j,

$$X_1^j \leq X_2^j \leq \cdots \leq X_N^j. \tag{2.3}$$

This may require a relabeling of the particles at each time step, since random walking the particles can result in a different ordering of the particle positions. This is simply a notational convenience and has no effect on the actual details of the convergence proof.

Let $\tilde{u}_i^j = \tilde{u}^j(X_i^j)$ denote the value of $\tilde{u}^j$ at the ith particle position. For future reference we note that

$$\tilde{u}_i^j = \sum_{r=1}^{N} H(X_r^j - X_i^j) w_r^j = \sum_{r \geq i} w_r^j. \tag{2.4}$$

Consequently, the strength of the ith particle is given by $w_i^j = \tilde{u}_i^j - \tilde{u}_{i+1}^j$. The variable N will always be used to denote the number of particles present in the flow; N is fixed for a given run of the numerical method.

2.2. *The Algorithm.* We begin the random gradient method by determining a step function approximation $\tilde{u}^0$ to the exact initial data u^0. Given the computed solution $\tilde{u}^j$ at time $j\Delta t$, the solution at time $(j+1)\Delta t$ is obtained in two distinct steps:

Step I. The first step is the numerical solution of $u_t = f(u)$. For fixed x, this is an ordinary differential equation (ODE) in t with initial data $\tilde{u}^j(x)$. The solution of this equation can easily be obtained using any explicit ODE solver. In the convergence proof that follows we will assume that Euler's method is used. It

should be noted, however, that the analysis (Section 6) carries through for higher-order Runge-Kutta methods as well. Furthermore, there are some cases in which $u_t = f(u)$ may be solved exactly. For example, in Eq (1.3a,b), $f(u) = u(1-u)$, in which case the reaction equation $u_t = u(1-u)$ may be solved exactly.

When the solution of the ODE is obtained using Euler's method, the value of the intermediate solution at the point x is given by

$$\tilde{v}^{j+1}(x) = \tilde{u}^j(x) + \Delta t f(\tilde{u}^j(x)). \tag{2.5}$$

Here, Δt is the time step and the variable $\tilde{v}$ has been used to denote the solution after one half of a two part fractional step. Since $\tilde{u}^j$ is a step function, so is $\tilde{v}^{j+1}$, the height of the step above the point x having been increased or decreased by the amount $\Delta t f(\tilde{u}^j(x))$. This is equivalent to altering the weights w_i^j so that the new weights w_i^{j+1} satisfy

$$\tilde{v}^{j+1}(x) = \sum_{i=1}^{N} H(X_i^j - x) w_i^{j+1}. \tag{2.6}$$

A simple formula can be derived for the w_i^{j+1}. Let $\tilde{v}_i^{j+1} = \tilde{v}^{j+1}(X_i^j)$ and define $\tilde{u}_{N+1}^j = 0$. Then

$$\begin{aligned} w_i^{j+1} &= \tilde{v}_i^{j+1} - \tilde{v}_{i+1}^{j+1} = \tilde{u}_i^j - \tilde{u}_{i+1}^j + \Delta t[f(\tilde{u}_i^j) - f(\tilde{u}_{i+1}^j)] \\ &= w_i^j + \Delta t[f(\tilde{u}_i^j) - f(\tilde{u}_{i+1}^j)]. \end{aligned} \tag{2.7}$$

Boundary Conditions. The function $\tilde{v}^{j+1}$ automatically satisfies the boundary condition (1.3e) since $H(X_i^j - x) = 0$ for all $x > X_i^j$. Furthermore, by summing over the w_i^j and using (1.2a) it is easy to show that $\sum w_i^j = 1$ implies $\sum w_i^{j+1} = 1$. Since Step II does not alter the particle strengths (as will be seen below), it follows that if $\sum w_i^0 = 1$, then the sum of the particle strengths is a conserved quantity in the random gradient method. In other words, $\tilde{u}^j$ satisfies the boundary condition (1.3d) at each time step if $\tilde{u}^0$ does initially.

Step II. It remains to solve the heat equation $u_t = \nu u_{xx}$ with initial data $\tilde{v}^{j+1}$. First select N random numbers $\eta_1, \eta_2, \ldots, \eta_N$ from a Gaussian distribution with mean 0 and variance $2\nu\Delta t$. The position of the ith particle X_i^j is then altered by the amount η_i to obtain $X_i^{j+1} = X_i^j + \eta_i$. Thus,

$$\tilde{u}^{j+1}(x) = \sum_{i=1}^{N} H(X_i^j + \eta_i - x) w_i^{j+1} = \sum_{i=1}^{N} H(X_i^{j+1} - x) w_i^{j+1}. \tag{2.8}$$

2.3. *Restriction to Monotonic Initial Data.* In order to prove the convergence of this method to solutions of (1.3a-e), we have found it necessary to assume that the initial approximation $\tilde{u}^0$ is monotonic. This is due to the following reason. If one allows particle weights with different signs, then some realizations of the $\eta_1, \ldots, \eta_N$ will result in $\tilde{u}^{j+1}(x) < 0$ for some x. This is true even if $0 \le \tilde{u}^j \le 1$ (and hence $0 \le \tilde{v}^{j+1} \le 1$) everywhere. (See Fig. 3 of [21].) Not only are such negative solutions incorrect (solutions of (1.3a-e) always lie in [0,1]; see Section 5), but solutions of $u_t = u(1-u)$ with negative initial data blow up in finite time. This can lead to particle strengths which increase without bound, further degrading the numerical solution. In [21] Hald encountered precisely this same problem and also found it

necessary to assume that the initial data is monotonic. Given these considerations, we start by defining the class **S** of acceptable starting approximations.

Definition 2.1. Let **S** be the class of all monotonically decreasing step functions $\tilde{u}$ which satisfy $\tilde{u}(-\infty) = 1$ and $\tilde{u}(\infty) = 0$. Thus, $\tilde{u} \in \mathbf{S}$ if and only if $\tilde{u}$ can be written in the form

$$\tilde{u}(x) = \sum_{i=1}^{N} H(X_i - x)w_i \tag{2.9}$$

where the weights $w_1, \ldots, w_N$ satisfy

$$0 < w_i \leq 1, \tag{2.10a}$$

$$\sum_{i=1}^{N} w_i = 1. \tag{2.10b}$$

We now show that the random gradient method maps the class **S** into itself, thereby avoiding the difficulties described above.

Assumption. Here and for the remainder of this paper we assume $f(u) = u(1-u)$. While many of the theorems that follow hold for general f, this assumption greatly simplifies the exposition.

LEMMA 2.2. *Fix $\Delta t \leq 1$ and assume that $\tilde{u}^j \in \mathbf{S}$. Let $\tilde{v}^{j+1}$ and $\tilde{u}^{j+1}$ be given by* (2.6) *and* (2.8), *respectively. Then $\tilde{v}^{j+1} \in \mathbf{S}$ and $\tilde{u}^{j+1} \in \mathbf{S}$.*

Proof. First we show $\tilde{v}^{j+1} \in \mathbf{S}$. Recall that $\sum w_i^j = 1$ implies $\sum w_i^{j+1} = 1$. Furthermore,

$$w_i^{j+1} = w_i^j + \Delta t(f(\tilde{u}_i^j) - f(\tilde{u}_{i+1}^j)) = w_i^j[1 + \Delta t(1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j))] > 0 \tag{2.11}$$

since $w_i^j[1 + \Delta t(1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j))]$ is the product of two positive quantities. This can be seen as follows. By assumption, $w_i^j > 0$. Since $0 \leq \tilde{u}_i^j \leq 1$ for all i and since $\tilde{u}_i^j = 1$ only if $i = 1$, we have

$$-1 < 1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j) \leq 1. \tag{2.12}$$

Finally, $w_i^{j+1} > 0$ and $\sum w_i^{j+1} = 1$ together imply $w_i^{j+1} \leq 1$ for all i. Thus, $\tilde{v}^{j+1} \in \mathbf{S}$ as claimed. Since $\tilde{v}^{j+1} \in \mathbf{S}$ and since an alteration of the particle positions has no effect on the weights, it follows that $\tilde{u}^{j+1} \in \mathbf{S}$ as well. □

COROLLARY 2.3. *Fix $\Delta t \leq 1$ and assume that $\tilde{u}^0 \in \mathbf{S}$. Then for all $j \geq 1$, we have $\tilde{v}^j \in \mathbf{S}$ and $\tilde{u}^j \in \mathbf{S}$.*

Assumption. Throughout the remainder of this paper we assume that $\Delta t \leq 1$.

One final fact will be established in this section, a bound on the particle strengths w_i^j. By (2.10a) and Corollary 2.3, $w_i^j \leq 1$ for all i, j. However, one needs to know that $Nw_i^j = O(1)$ as $N \to \infty$. If the strengths are initially chosen so that $w_i^0 = O(N^{-1})$, then this is a consequence of the following lemma.

LEMMA 2.4. *For $\tilde{u}^0 \in \mathbf{S}$ let $\tilde{u}^j = \sum H(X_i^j - x)w_i^j$ be the computed solution at time $T = j\Delta t$. Then for all i, the particle strengths w_i^j satisfy*

$$w_i^j \leq e^T w_i^0. \tag{2.13}$$

Proof. From (2.11) and (2.12) we see that $w_i^{j+1} = w_i^j[1 + \Delta t(1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j))] \leq w_i^j(1 + \Delta t)$. The inequality in (2.13) follows immediately. □

3. Solution Operator Notation and an Outline of the Proof. The primary purpose of this section is to develop a notation with which to discuss the error. We then present an outline of the convergence proof. We begin by assuming that $\nu = 1$. This makes the exposition simpler. This restriction will be removed at the end of Section 8.

3.1. *Solution Operator Notation.* Define F_t, the *exact solution operator* for the Kolmogorov equation, by $F_t u^0(x) = u(x,t)$, where $u(x,t)$ is the solution to (1.3a,b) at time t. Note that if $t = j\Delta t$, then $u(x,t) = F_{\Delta t}^j u^0$, where the superscript j indicates the jth power of the operator $F_{\Delta t}$. The *reaction operator* R_t and the *diffusion operator* D_t are defined similarly. Thus, $R_t u^0$ is the solution at time t to the reaction equation with initial data u^0,

$$u_t = u(1-u), \tag{3.1a}$$

$$u(x,0) = u^0(x), \tag{3.1b}$$

and $D_t u^0$ is the solution at time t to the heat equation with initial data u^0,

$$u_t = u_{xx}, \tag{3.2a}$$

$$u(x,0) = u^0(x). \tag{3.2b}$$

Let $u(x)$ be an arbitrary piecewise continuous function. Define the *approximate reaction operator* $\tilde{R}_{\Delta t}$ by

$$\tilde{R}_{\Delta t} u(x) = u(x) + \Delta t u(x)(1-u(x)). \tag{3.3}$$

In other words, for each fixed $x \in \mathbf{R}$, $\tilde{R}_{\Delta t} u(x)$ is simply the Euler's method approximation after one time step to the solution of (3.1a,b) with initial data $u(x)$. Similarly, for an arbitrary step function $\tilde{u}$ of the form (2.9) we define the *approximate diffusion operator* $\tilde{D}_{\Delta t}$ by

$$\tilde{D}_{\Delta t}\tilde{u}(x) = \sum_{i=1}^{N} H(X_i + \eta_i - x) w_i \tag{3.4}$$

where $\eta_1, \eta_2, \ldots, \eta_N$ are N independent random numbers chosen from a Gaussian distribution with mean 0 and variance $2\Delta t$. Thus, $\tilde{v}^j = \tilde{R}_{\Delta t}\tilde{u}^{j-1}$, $\tilde{u}^j = \tilde{D}_{\Delta t}\tilde{v}^j$, and $\tilde{u}^j = (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^j \tilde{u}^0$.

3.2. *An Outline of the Proof.* Let $\tilde{u}^0 \in \mathbf{S}$ be a step function approximation to the initial data u^0. The L^1 difference at time $T = k\Delta t$ between the exact solution of (1.3a,b) and our approximate solution may be divided into three distinct components,

$$\begin{aligned} \|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 \le{}& \|F_{\Delta t}^k u^0 - (D_{\Delta t}R_{\Delta t})^k u^0\|_1 \\ &+ \|(D_{\Delta t}R_{\Delta t})^k u^0 - (D_{\Delta t}R_{\Delta t})^k \tilde{u}^0\|_1 \\ &+ \|(D_{\Delta t}R_{\Delta t})^k \tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1. \end{aligned} \tag{3.5}$$

The first term on the right is called the *splitting error.* It is the error due to the fractional step or *exact operator splitting.* In Section 5 we prove that this error is $O(\Delta t)$,

$$\|F_{\Delta t}^k u^0 - (D_{\Delta t}R_{\Delta t})^k u^0\|_1 \le C_1 \Delta t. \tag{3.6}$$

The second term on the right is the error due to our approximation of the initial data u^0 by the step function $\tilde{u}^0$. In Section 4 we show that the operators R_t and D_t are stable in the L^1 norm and hence, that

$$\|(D_{\Delta t}R_{\Delta t})^k u^0 - (D_{\Delta t}R_{\Delta t})^k \tilde{u}^0\|_1 \leq e^T \|u^0 - \tilde{u}^0\|_1.$$

The third term on the right in (3.5) is the error due to the numerical approximation of the solutions to Eqs. (3.1a,b) and (3.2a,b). That is, the error that results from approximating the exact operators $R_{\Delta t}$ and $D_{\Delta t}$ by the approximate operators $\tilde{R}_{\Delta t}$ and $\tilde{D}_{\Delta t}$. Assume N and Δt have been chosen so that for some constant C_0, $\Delta t = C_0 \sqrt[4]{N}^{-1}$. Since the effect of the operator $\tilde{D}_{\Delta t}$ is random, the bound on this error takes the form

$$P\left(\|(D_{\Delta t}R_{\Delta t})^k \tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 \geq \gamma C_2 \frac{\ln N}{\sqrt[4]{N}}\right) \leq \frac{4T}{C_0} N^{-\gamma}, \tag{3.7}$$

where $\gamma \geq 1$ is an arbitrary real number. To prove this, we use the L^1 stability of the operators D_t and R_t to divide this error into $2k$ pieces,

$$\begin{aligned} &\|(D_{\Delta t}R_{\Delta t})^k \tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 \\ &\quad \leq e^T \sum_{j=0}^{k-1} \|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 + e^T \sum_{j=1}^{k} \|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1. \end{aligned} \tag{3.8}$$

Let $t = j\Delta t$. In Section 6 we prove that for each j,

$$P(\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 \geq \gamma B_1(t)\sqrt{\ln N}(\Delta t)^2) \leq N^{-5\gamma/4}. \tag{3.9}$$

This estimate is based on the fact that Euler's method has local truncation error $O((\Delta t)^2)$, that $R_t\tilde{u}^j(x) = \tilde{R}_{\Delta t}\tilde{u}^j(x)$ for all $|x| > \max|X_i^j|$, and on a probabilistic bound for the X_i^j. In Section 7 we prove

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1 \geq \gamma B_2(t)\frac{\ln N}{\sqrt{N}}\right) \leq 3N^{-5\gamma/4}. \tag{3.10}$$

The proof is based on the pointwise estimate $P(|D_{\Delta t}\tilde{v}^j(x) - \tilde{D}_{\Delta t}\tilde{v}^j(x)| \geq C\alpha) \leq e^{-2N\alpha^2}$, where $\alpha > 0$ is arbitrary. Using (3.9) and (3.10) in (3.8) we obtain (3.7).

4. The Exact Solution Operators R_t and D_t. In this section we develop some of the basic properties of the operators R_t and D_t, the principal result being that both operators are stable in the L^1 norm. We then use this fact to examine the propagation of the error which is induced by approximating the initial data with a step function.

4.1. *The Exact Reaction Operator R_t.* It is a simple matter to check that the function defined by

$$R_t u^0(x) = \frac{u^0(x)e^t}{1 + (e^t - 1)u^0(x)} \tag{4.1}$$

is a solution of the reaction equation (3.1a,b). The L^1 stability of R_t is an immediate consequence** of having an exact expression for $R_t u^0$.

**It has been pointed out by a referee that the L^1 stability of R_t also follows from the fact that $u(1-u)$ is Lipschitz continuous.

LEMMA 4.1 (L^1 Stability of R_t). *Let u and v be measurable functions on $\mathbb{R}$ such that $0 \le u, v \le 1$ and $\|u-v\|_1 < \infty$. Then for any time $t > 0$, $\|R_t u - R_t v\|_1 \le e^t \|u - v\|_1$.*

During the course of proving that the error due to exact operator splitting is small (Section 5) it will be necessary to bound $R_t(u)_x$. By differentiating (4.1) with respect to x we obtain the following result.

LEMMA 4.2. *Let $u \in C^1(\mathbb{R})$ and assume that $0 \le u \le 1$. Then for $p = 1, \infty$ and any time $t > 0$,*

$$\|u_x\|_p < \infty \quad \textit{implies} \quad \|(R_t u)_x\|_p \le e^t \|u_x\|_p.$$

4.2. *The Exact Diffusion Operator D_t.* Define the heat kernel $G(x,t)$ by

$$G(x,t) = \frac{e^{-x^2/4t}}{\sqrt{4\pi t}}. \tag{4.2}$$

Occasionally, when there is no possibility of confusion, we will write $G_t(x)$ instead of $G(x,t)$. The solution of the heat equation (3.2a,b) is given by $u(x,t) = (G_t * u^0)(x)$ where $*$ denotes convolution. Hence $D_{\Delta t} u^0 = G_{\Delta t} * u^0$. A basic result from the theory of partial differential equations is that the diffusion operator D_t maps $L^p(\mathbb{R})$ onto $L^p(\mathbb{R})$ for $1 \le p \le \infty$ (see [14]). In particular, we have the following fact.

LEMMA 4.3. *Let u be any measurable function of $x \in \mathbb{R}$ and let $1 \le p \le \infty$. Then for any time $t > 0$,*

$$\|u\|_p < \infty \quad \textit{implies} \quad \|D_t u\|_p \le \|u\|_p.$$

Remark. For any bounded, differentiable function u on $\mathbb{R}$ which satisfies $u_x \in L^1$ we have $(G_t * u)_x = G_t * u_x$. Consequently, it follows from Lemma 4.3 that $\|(D_t u)_x\|_p \le \|u_x\|_p$ as well.

The L^1 stability of D_t is an immediate consequence of Lemma 4.3 and the linearity of D_t.

COROLLARY 4.4 (L^1 Stability of D_t). *Let $t \ge 0$. Then for any bounded measurable functions u, v defined on $\mathbb{R}$ such that $\|u-v\|_1 < \infty$ we have $\|D_t u - D_t v\|_1 \le \|u - v\|_1$.*

4.3. *The Error Due to Approximating the Initial Data.* The L^1 stability of the operators $R_{\Delta t}$ and $D_{\Delta t}$ allows us to bound the error that occurs as a result of approximating the initial data u^0 with a step function $\tilde{u}^0$. This is accomplished by examining the second term on the right in (3.5). This expression can be bounded in terms of the initial error by repeatedly applying Corollary 4.4 to functions of the form $u = R_{\Delta t}(D_{\Delta t} R_{\Delta t})^j u^0$, $v = R_{\Delta t}(D_{\Delta t} R_{\Delta t})^j \tilde{u}^0$ and Lemma 4.1 to functions of the form $u = (D_{\Delta t} R_{\Delta t})^j u^0$, $v = (D_{\Delta t} R_{\Delta t})^j \tilde{u}^0$. We note that the hypotheses of Lemma 4.1 require $0 \le u, v \le 1$. This follows from the simple fact that

$$0 \le u \le 1 \quad \text{implies} \quad 0 \le R_t u(x) \le 1 \text{ and } 0 \le D_t u(x) \le 1. \tag{4.3}$$

We state this result in a somewhat more general form here for future reference.

THEOREM 4.5. *Let u and v be bounded measurable functions defined on* $\mathbf{R}$ *satisfying* $0 \le u,\ v \le 1$. *Then*

$$\|(D_{\Delta t}R_{\Delta t})^k u - (D_{\Delta t}R_{\Delta t})^k v\|_1 \le e^{k\Delta t}\|u - v\|_1.$$

5. The Error Due to Exact Operator Splitting. In this section we prove (3.6). The key idea is to show that $w(x,t) = F_t u^0(x) - D_t R_t u^0(x)$ is a solution of

$$w_t = w_{xx} + a(x,t)w + b(x,t), \tag{5.1a}$$

$$w(x,0) = w^0(x) \tag{5.1b}$$

with $w^0(x) \equiv 0$ and $\|b(\cdot,t)\|_1 = O(t)$. Hence, there exists a constant $A > 0$ such that

$$\|w(\cdot,T)\|_1 \le e^{AT}\|w^0\|_1 + e^{AT}\int_0^T \|b(\cdot,t)\|_1\,dt \tag{5.2}$$

for all times $T > 0$. By setting $T = \Delta t$, it follows that the splitting error after one time step is $O(\Delta t^2)$,

$$\|F_{\Delta t}u^0 - D_{\Delta t}R_{\Delta t}u^0\|_1 \le C(\Delta t)^2. \tag{5.3}$$

This inequality, together with the L^1 stability of the operator $F_{\Delta t}$, yields (3.6).

The ideas we use to prove (3.6) have been used by Roberts [31] to establish an analogous result for operator splitting applied to Burgers' equation. Similar results in the L^2 norm have been obtained by Beale and Majda [2] for viscous splitting of the Navier-Stokes equations.

5.1. *Some Properties of Solutions to the Kolmogorov Equation.* Solutions of the nonlinear reaction-diffusion equation (1.1a,b) satisfy a maximum principle in much the same way as do solutions of linear parabolic differential equations. Here we state the maximum principle for solutions of the Kolmogorov equation (1.3a,b). For a proof see [5].

LEMMA 5.1 (Maximum Principle). *Let u, v be bounded solutions of* (1.3a,b) *on* $\Omega = \mathbf{R} \times [0,T]$ *with initial data u^0, v^0, respectively. Suppose that $v^0(x) \le u^0(x)$ for all $x \in \mathbf{R}$. Then $v(x,t) \le u(x,t)$ for all $(x,t) \in \Omega$.*

For bounded and sufficiently smooth initial data, solutions of (1.3a,b) exist, are unique, and possess bounded derivatives. In particular, solutions u of (1.3a,b) with $u^0 \in C^1(\mathbf{R})$, $0 \le u^0 \le 1$, and $\|u_x^0\|_\infty < \infty$ satisfy

$$\|u_x(t,\cdot)\|_\infty \le e^t\|u_x^0\|_\infty, \tag{5.4}$$

a fact which we shall have occasion to use. For details we refer the reader to [5] or [35]. Next we state the conditions under which solutions of (5.1a,b) satisfy (5.2).

LEMMA 5.2. *Fix $T > 0$ and let $a(x,t)$ and $b(x,t)$ be bounded, continuous functions on the strip $\Omega = \mathbf{R} \times [0,T]$ such that $a \in C^1(\mathbf{R})$, a_x is bounded in Ω, and $b \in L^1(\Omega)$. Suppose $w(x,t)$ is a solution of* (5.1a,b) *in $\mathbf{R} \times (0,T]$ with $w^0 \in L^1(\mathbf{R})$, and that w and w_x are bounded in Ω. Then w satisfies* (5.2) *for $A = \sup_\Omega\{0, a(x,t)\}$.*

Remark. The L^1 stability of solutions to (1.3a,b) follows from (5.2) and (5.4) provided we restrict ourselves to solutions with initial data that are bounded between 0 and 1 and have bounded first derivative.

5.2. *The Splitting Error.* We are now ready to prove (5.3). This is accomplished in two steps. The first step consists of using (5.2) to show that the L^1 norm of the function $w(x,t) = F_t u^0 - D_t R_t u^0$ is $O(t^2)$. If one regards $D_{\Delta t} R_{\Delta t} u^0$ as a numerical approximation to $F_{\Delta t} u^0$ after one time step of length Δt, then this is simply the statement that the local truncation error is of order $(\Delta t)^2$. In other words, our numerical scheme (approximating $F_{\Delta t} u^0$ by $D_{\Delta t} R_{\Delta t} u^0$) is consistent.

THEOREM 5.3. *Let $u(x,t)$ be a solution of the Kolmogorov equation* (1.3a,b) *with initial data $u^0 \in C^1(\mathbf{R})$. Assume that $0 \le u^0 \le 1$ and that $u_x^0 \in L^1(\mathbf{R}) \cap L^\infty(\mathbf{R})$. Then* (5.3) *holds with*

$$C = e^{3\Delta t} \left\{ e^{\Delta t} \|u_x^0\|_\infty + 4 \frac{\sqrt{2\Delta t}}{\sqrt{\pi}} \right\} \|u_x^0\|_1. \tag{5.5}$$

Proof. We will show that $w(x,t) = F_t u^0(x) - D_t R_t u^0(x)$ satisfies a differential equation of the form (5.1a,b) on $\Omega = \mathbf{R} \times [0, \Delta t]$ with initial data $w^0(x) \equiv 0$ and where $b(x,t)$ in (5.1a) satisfies $\|b\|_{L^1(\Omega)} = O(\Delta t^2)$. Then (5.3) is a consequence of (5.2). Let $u(x,t) = F_t u^0(x)$ and $v(x,t) = R_t u^0(x)$. Then $w(x,t) = u(x,t) - (G * v)(x,t)$, where $G(x,t)$ is the heat kernel (4.2). By differentiating w with respect to t and using the fact that $v_t = v(1-v)$ and $G_t = G_{xx}$ one can show that w satisfies (5.1a,b) with $a = 1 - (u + G * v)$ and $b = G * v^2 - (G * v)^2$.

It follows from Lemma 5.1 that $\|u(\cdot,t)\|_\infty \le 1$ for all time $t \ge 0$. Furthermore, by (5.4), $\|u_x(\cdot,t)\|_\infty \le e^t \|u_x^0\|_\infty$. Identical estimates hold for v and $G * v$. For, by (4.3), $\|v(\cdot,t)\|_\infty \le 1$ and hence, $\|(G * v)\|_\infty \le 1$ for all $t \ge 0$. (When it is convenient to do so we will often suppress mention of the variable t.) By Lemma 4.2, $\|v_x\|_\infty \le e^t \|u_x^0\|_\infty$ and, upon writing $(G * v)_x = (G * v_x)$, one finds that $\|(G*v)_x\|_\infty \le e^t \|u_x^0\|_\infty$. Thus, a, a_x, b, w and w_x are bounded continuous functions on Ω.

It remains to show that $\|b\|_{L^1(\Omega)} = O(\Delta t^2)$. To this end, Lemma 5.2 will once again be used, this time applied to the function $b = G*v^2 - (G*v)^2$. Differentiating b with respect to t and remembering that $v_t = v(1-v)$ and $G_t = G_{xx}$, one finds that b satisfies

$$b_t = b_{xx} + 2b + c(x,t), \tag{5.6a}$$

$$b(x,0) = 0 \tag{5.6b}$$

with $c = 2[(G*v_x)^2 + (G*v)(G*v^2) - G*v^3]$. Noting that v and v_x are bounded and continuous in Ω, it follows that b, b_x, and c are as well. Now consider $t \in [0, \Delta t]$. We claim that

$$\int_0^t \int_{-\infty}^{\infty} |c(x,s)|\, dx\, ds \le 2Cte^{-2t}, \tag{5.7}$$

where C is given by (5.5). This follows from $\|c(\cdot,t)\|_1 \le 2Ce^{-2t}$ whenever $t \le \Delta t$, a fact which we now prove. By Lemmas 4.2 and 4.3,

$$\|(G * v_x)^2\|_1 \le \|G * v_x\|_\infty \|G * v_x\|_1 \le \|v_x\|_\infty \|v_x\|_1 \le e^{2t} \|u_x^0\|_\infty \|u_x^0\|_1.$$

To estimate the remaining portion of c, we write

$$\begin{aligned}
&\|(G*v)(G*v^2)-(G*v^3)\|_1 \\
&\quad\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)G(x-z)|v(z)v^2(y)-v^3(z)|\,dy\,dz\,dx \\
&\quad\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)G(x-z)v(z)(v(y)+v(z))|v(y)-v(z)|\,dy\,dz\,dx \\
&\quad= 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)G(x-z)\left|\int_0^1 v_x(z+\theta(y-z))\,d\theta\right||y-z|\,dy\,dz\,dx \\
&\quad\le 2\|v_x(\cdot,t)\|_1\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(y)G(z)|y-z|\,dy\,dz \\
&\quad\le 4\frac{\sqrt{2t}}{\sqrt{\pi}}e^t\|u_x^0\|_1,
\end{aligned}$$

where we have used the fact that $0\le v\le 1$. Thus,

$$\|c(\cdot,t)\|_1\le 2\left\{e^{2t}\|u_x^0\|_1\|u_x^0\|_\infty+\frac{4\sqrt{2t}}{\sqrt{\pi}}e^t\|u_x^0\|_1\right\}\le 2Ce^{-2t},$$

whereby (5.7) holds for $t\in[0,\Delta t]$ as claimed. Since $b(x,t)$ is a solution of (5.6a,b), it now follows from Lemma 5.2 that $\|b(\cdot,t)\|_1\le 2Ct$. Applying Lemma 5.2 once more, this time to $w(x,t)$ on $\Omega=\mathbf{R}\times[0,\Delta t]$, we obtain (5.3). □

Continuing to think of $D_{\Delta t}R_{\Delta t}u^0$ as a numerical approximation to $F_{\Delta t}u^0$, we may now use consistency (Theorem 5.3) together with stability (see the remark after Lemma 5.2) to prove that for fixed time $T=k\Delta t$, $(D_{\Delta t}R_{\Delta t})^k u^0\to F_{\Delta t}^k u^0$ as $\Delta t\to 0$. The proof proceeds precisely as it would for a numerical method; at a given time step j we use stability and consistency to reduce the error at time $j\Delta t$ into the error at time $(j-1)\Delta t$ plus a term of order $(\Delta t)^2$. Thus, the error at time $T=k\Delta t$ is the sum of k terms, each of order $(\Delta t)^2$ plus the error due to the initial approximation (which in our case is 0). The only detail remaining is that, in order to apply the stability and consistency results, we must check that the functions obtained at the intermediate time steps satisfy the appropriate hypotheses.

THEOREM 5.4. *Let u be a solution of the Kolmogorov equation* (1.3a,b) *with initial data $u^0\in C^1(\mathbf{R})$ such that $0\le u^0\le 1$ and $u_x^0\in L^1(\mathbf{R})\cap L^\infty(\mathbf{R})$. Then* (3.6) *holds where, for $T=k\Delta t$,*

$$C_1=Te^{4T}\left\{e^T\|u_x^0\|_\infty+\frac{4\sqrt{2\Delta t}}{\sqrt{\pi}}\right\}\|u_x^0\|_1. \tag{5.8}$$

Proof. Let $u^j=F_{\Delta t}^j u^0$. By the maximum principle, $0\le u^j\le 1$ for each j and, since $u_x^0\in L^\infty(\mathbf{R})$, (5.4) implies $(u^j)_x\in L^\infty(\mathbf{R})$. Similarly, let $v^j=(D_{\Delta t}R_{\Delta t})^j u^0$. Repeated application of (4.3) yields $0\le v^j\le 1$. Furthermore, Lemmas 4.2 and 4.3 imply that $\|(v^j)_x\|_\infty\le e^{j\Delta t}\|u_x^0\|_\infty$ and $\|(v^j)_x\|_1\le e^{j\Delta t}\|u_x^0\|_1$. Hence, $(v^j)_x\in L^1(\mathbf{R})\cap L^\infty(\mathbf{R})$. We can now use the L^1 stability of $F_{\Delta t}$ (one must verify, by induction, that $u^{j-1}-v^{j-1}\in L^1$) and Theorem 5.3 to obtain

$$\begin{aligned}
&\|F_{\Delta t}^j u^0-(D_{\Delta t}R_{\Delta t})^j u^0\|_1 \\
&\quad\le \|F_{\Delta t}u^{j-1}-F_{\Delta t}v^{j-1}\|_1+\|F_{\Delta t}v^{j-1}-D_{\Delta t}R_{\Delta t}v^{j-1}\|_1 \\
&\quad\le e^{\Delta t}\|u^{j-1}-v^{j-1}\|_1+C_{j-1}(\Delta t)^2,
\end{aligned}$$

where the constant C_{j-1} is given by (5.5) with u^0 replaced by v^{j-1}. Hence,

$$\|F_{\Delta t}^k u^0 - (D_{\Delta t} R_{\Delta t})^k u^0\|_1 \le \sum_{j=0}^{k-1} e^{(k-1-j)\Delta t} C_j (\Delta t)^2 \le \max_{j<k} C_j T e^T \Delta t.$$

Now use the L^1 and L^∞ bounds on $(v^j)_x$ found above to obtain

$$\max_{j<k} C_j = \max_{j<k} e^{3\Delta t} \left\{ e^{\Delta t} \|(v^j)_x\|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \right\} \|(v^j)_x\|_1$$

$$\le e^{3T} \left\{ e^T \|u_x^0\|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \right\} \|u_x^0\|_1. \quad \square$$

6. The Error Due to the Approximate Reaction Operator $\tilde{R}_{\Delta t}$. In this section we prove that the inequality in (3.9) holds for $\gamma \ge 1$ where the constant $B_1 = B_1(t)$ depends only on the initial data $\tilde{u}^0$ and the time $t = j\Delta t$. The proof is based on the following two points:

(i) Given any $L > 0$ such that for all $i, X_i^j \in (-L, L)$ then, by (2.1), $x < -L$ implies $\tilde{u}^j(x) = 1$ and $x > L$ implies $\tilde{u}^j(x) = 0$. Therefore, $R_{\Delta t}\tilde{u}^j(x) = \tilde{R}_{\Delta t}\tilde{u}^j(x)$ for $|x| > L$ and the L^1 estimate of the error is reduced to an estimate over the interval $(-L, L)$,

$$\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 = \int_{-L}^{L} |R_{\Delta t}\tilde{u}^j(x) - \tilde{R}_{\Delta t}\tilde{u}^j(x)|\, dx = \|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_{(-L,L)}.$$

(ii) For fixed x, $\tilde{R}_{\Delta t}$ is simply Euler's method for approximating the solution of an ODE, and hence the local truncation error is known to be $O(\Delta t^2)$. This fact can be exploited to obtain a bound for $|R_{\Delta t}\tilde{u}^j(x) - \tilde{R}_{\Delta t}\tilde{u}^j(x)|$ which is uniform in x.

Together, (i) and (ii) imply $\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 = \|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_{(-L,L)} \le$ const $2L(\Delta t)^2$. In general, however, the size of the interval $(-L, L)$ cannot be given a deterministic bound. For the particle positions X_i^j are random variables, yet L has been chosen so that $|X_i^j| < L$ for all i. Consequently, the most that one can hope for is to find the probability that L is a given size. This is accomplished by examining the movement of the particles.

Recall that X_i^j is obtained from X_i^{j-1} by adding a normally distributed random variable η_i^j with mean 0 and variance $2\Delta t$ to X_i^{j-1}. The movement of the particles is thus governed by the distribution of the η_i^j,

$$P(\eta_i^j < x) = \frac{1}{\sqrt{4\pi\Delta t}} \int_{-\infty}^{x} e^{-s^2/4\Delta t}\, ds = \phi\left(\frac{x}{\sqrt{2\Delta t}}\right),$$

where ϕ is the probability distribution function for a Gaussian distribution with mean 0 and variance 1 (see Chung [10, p. 100]). By writing $X_i^j = X_i^0 + \eta_i^1 + \cdots + \eta_i^j$ and noting that $\eta_i^1 + \cdots + \eta_i^j$ is a normally distributed random variable with mean 0 and variance $2j\Delta t$ we obtain the following result.

LEMMA 6.1. *Let $K > 0$ by chosen so that $X_i^0 \in (-K, K)$ for all i. Then for all $\alpha > 0$,*

$$P(|X_i^j| > K + \alpha) < 2\phi\left(\frac{-\alpha}{\sqrt{2j\Delta t}}\right).$$

It is well known that $\phi(x)$ decreases at an exponential rate as $x \to -\infty$. This allows us to compute a bound on $\phi(x)$ and hence on the probability that the particles lie outside a given interval.

LEMMA 6.2. *For any* $x < 0$,

$$\phi(x) \le \frac{1}{|x|\sqrt{2\pi}} e^{-x^2/2}.$$

Proof. Since $s \le x < 0$, we have $s/x \ge 1$ and

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2}\, ds \le \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \frac{s}{x} e^{-s^2/2}\, ds = \frac{1}{|x|\sqrt{2\pi}} e^{-x^2/2}. \quad \square$$

We now establish a probability inequality for $\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1$ which depends on the parameter α. An appropriate choice of α then yields (3.9).

THEOREM 6.3. *Let* $\tilde{u}^0 \in \mathbf{S}$ *and let* $K > 0$ *be such that* $X_i^0 \in (-K, K)$ *for all* i. *Then for all* $\alpha > 0$,

$$P\left(\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 > \frac{\sqrt{3}}{9}(K+\alpha)(\Delta t)^2\right) \le \frac{2N\sqrt{j\Delta t}}{\alpha\sqrt{\pi}} e^{-\alpha^2/4j\Delta t}. \tag{6.1}$$

Proof. Set $L = K + \alpha$ and assume that $-L < X_i^j < L$, $i = 1, \ldots, N$. Using the fact that $0 \le \tilde{u}^j(x) \le 1$ (and hence $0 \le R_t\tilde{u}^j(x) \le 1$) we find

$$|R_{\Delta t}\tilde{u}^j(x) - \tilde{R}_{\Delta t}\tilde{u}^j(x)| \le (\sqrt{3}/18)(\Delta t)^2$$

for all $x \in \mathbf{R}$. It follows that

$$\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 = \int_{-L}^{L} |R_{\Delta t}\tilde{u}^j(x) - \tilde{R}_{\Delta t}\tilde{u}^j(x)|\, dx \le \frac{\sqrt{3}}{9} L(\Delta t)^2. \tag{6.2}$$

This estimate is valid as long as our assumption that $-L < X_i^j < L$ holds. Thus, by Lemmas 6.1 and 6.2,

$$\begin{aligned} P\left(\|R_{\Delta t}\tilde{u}^j - \tilde{R}_{\Delta t}\tilde{u}^j\|_1 > \frac{\sqrt{3}}{9} L(\Delta t)^2\right) &\le \sum_{i=1}^{N} P(|X_i^j| > L) \\ &\le \frac{2N\sqrt{j\Delta t}}{\alpha\sqrt{\pi}} e^{-\alpha^2/4j\Delta t}. \quad \square \end{aligned}$$

We now set $\alpha = 3\gamma\sqrt{t}\sqrt{\ln N}$ in (6.1) to obtain (3.9).

COROLLARY 6.4. *Let* $\tilde{u}^0 \in \mathbf{S}$ *be generated by* $N \ge 3$ *particles, let* $t = j\Delta t$, *and let* $K > 0$ *be chosen so that for all* i, $X_i^0 \in (-K, K)$. *Then for each* $j = 1, 2 \ldots$, *and all* $\gamma \ge 1$ *the inequality in* (3.9) *holds with*

$$B_1(t) = \frac{\sqrt{3}}{9}(K + 3\sqrt{t}). \tag{6.3}$$

7. The L^1 Convergence of the Approximate Diffusion Operator $\tilde{D}_{\Delta t}$. In this section we prove that, whenever $\tilde{u}^0 \in \mathbf{S}$, inequality (3.10) holds for all $\gamma \ge 1$ where the constant $B_2 = B_2(t)$ depends only on the initial data $\tilde{u}^0$ and the time

$t = j\Delta t$. Setting $\gamma = 1$ it follows that, given any $\varepsilon > 0$, we can find $N_0 = N_0(\varepsilon)$ such that for all $N \geq N_0$,

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1 \geq B_2(t)\frac{\ln N}{\sqrt{N}}\right) \leq \varepsilon.$$

Thus, by using sufficiently many particles, one can guarantee that the error due to approximating $D_{\Delta t}$ by $\tilde{D}_{\Delta t}$ is small with arbitrarily high probability. It is in this sense that the approximate diffusion operator $\tilde{D}_{\Delta t}$ converges to the exact diffusion operator $D_{\Delta t}$. Similar results hold in the L^2 norm ([22], [27]) and the sup norm ([18]).

Note that the convergence rate as stated here is $O(\ln N/\sqrt{N})$. The true convergence rate is probably $O(1/\sqrt{N})$, with the factor $\ln N$ being a spurious term introduced by the analysis. Furthermore, note that the rate of convergence does not depend on the time step Δt. Thus, since all of the other sources of error behave like $O(\Delta t)$ and since it is considerably cheaper to halve the time step than to quadruple the number of particles, this quickly becomes the dominant source of error. This feature is common to all numerical methods which use random walks or some other form of random sampling.

The results in this section are based on the work of Roberts. Most of the reasoning is identical to the argument in Section 4 of [31]. The main difference between Roberts' convergence proof for the approximate diffusion operator and the proof here is that in his work the particle strengths are constant in time, whereas here the particle strengths are random variables. This difference manifests itself primarily in Lemma 7.4 where, in order to establish a pointwise bound on the difference between $D_{\Delta t}\tilde{v}^j$ and $\tilde{D}_{\Delta t}\tilde{v}^j$, it is first necessary to bound the particle strengths.

7.1. *The Underlying Probability Space Ω and a Brief Outline of the Argument.* Implicit in inequalities (3.7), (3.9), and (3.10) is the existence of a probability space (Ω, Σ, μ) over which the respective errors are random variables. We can construct this probability space in the following way. Let $(\Omega, \Sigma, \mu) = (\prod_{j=1}^k \Omega_j, \prod_{j=1}^k \Sigma_j, \prod_{j=1}^k \mu_j)$, where $\Omega_j = \mathbf{R}^N, \Sigma_j$ is the Euclidean Borel field on $\mathbf{R}^N$, and $\mu_j = \mu_j(\Delta t)$ is Gaussian measure on $\mathbf{R}^N$ with mean 0 and variance $2\Delta t$. Here, N is the number of particles, Δt the time step, and $k\Delta t$ the *final* time at which we wish to examine the error; N, Δt, and k are all fixed. There is a simple one-to-one correspondence between elements of Ω and a given run of the random gradient method: each $\omega \in \Omega$ corresponds to one realization of the random walks, $\omega = (\eta_1^1, \ldots, \eta_N^1, \ldots, \eta_1^k, \ldots, \eta_N^k)$. The component space $(\Omega_j, \Sigma_j, \mu_j)$ has been chosen so that there is a one-to-one correspondence between an element $\omega_j \in \Omega_j$ and the N random numbers used at the jth time step to random walk the particles, $\eta_1^j, \ldots, \eta_N^j$.

Now let Z be a random variable on (Ω, Σ, μ) with $E[Z] < \infty$, let $\Omega_j^* = \prod_{l \neq j} \Omega_l$, $\Sigma_j^* = \prod_{l \neq j} \Sigma_l, \mu_j^* = \prod_{l \neq j} \mu_l$, and let ω_j^* denote an element of Ω_j^*. By Fubini's theorem we have

$$E[Z] = \int_{\Omega_j^*} \int_{\Omega_j} Z(\omega_j, \omega_j^*)\mu_j(d\omega_j)\mu_j^*(d\omega_j^*) = \int_{\Omega_j^*} E_{\Omega_j}[Z]\mu_j^*(d\omega_j^*), \tag{7.1}$$

where

$$E_{\Omega_j}[Z] \equiv \int_{\Omega_j} Z(\omega_j, \omega_j^*)\mu_j(d\omega_j)$$

is the conditional expectation of Z with respect to the Borel field $\{\Omega_j\} \times \Sigma_j^* \subset \Sigma$. (See Chung [10] for a discussion of conditional expectation, conditional probability, and an explanation of the basic facts regarding these concepts which we use below.) Define

$$P_{\Omega_j}(Z \geq \varepsilon) \equiv E_{\Omega_j}[H(Z - \varepsilon)]. \tag{7.2}$$

Since $H(Z - \varepsilon)$ is the characteristic function for the event $A \equiv Z \geq \varepsilon$, it follows that $P_{\Omega_j}(A)$ is the conditional probability of A with respect to $\{\Omega_j\} \times \Sigma_j^*$. Note that $P_{\Omega_j}(A)$ is a random variable which belongs to $\{\Omega_j\} \times \Sigma_j^*$, and that it depends on ω_j^* but not on ω_j. Furthermore, note that when $Z = \|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1$, the random variable $P_{\Omega_j}(A)$ does not depend on the random walks taken after time $j\Delta t$ and hence, an equivalent way of writing $P_{\Omega_j}(A)$ is

$$P(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1 \geq \varepsilon \mid \text{ the past up to time } (j-1)\Delta t).$$

Our proof of (3.10) proceeds as follows. We first divide the real line into two pieces, $(-L, L)$ and $(-L, L)^c = (-\infty, -L) \cup (L, \infty)$, where $L > 0$ is free to be chosen as we wish. We then write

$$\begin{aligned} P(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1 \geq \varepsilon) \leq\ & P(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \geq \varepsilon_1) \\ & + P(\|D_{\Delta t}\tilde{v} - \tilde{D}_{\Delta t}\tilde{v}\|_{(-L,L)^c} \geq \varepsilon_2), \end{aligned}$$

where $\varepsilon = \varepsilon_1 + \varepsilon_2$ and $\|\cdots\|_{(-L,L)}$ (resp. $\|\cdots\|_{(-L,L)^c}$) denotes the L^1 norm over $(-L, L)$ (resp. $(-L, L)^c$). To estimate the error on the tails $(-L, L)^c$ we bound the error over $(-\infty, -L)$ (resp. (L, ∞)) under the assumption that all of the particles lie in the interval $(-B, B) \subset (-L, L)$ at times $(j-1)\Delta t$ and $j\Delta t$. We then use Lemmas 6.1 and 6.2 to estimate the probability that this assumption holds. To estimate the error over the interval $(-L, L)$ we show that for all $\omega_j^* \in \Omega_j^*$

$$P_{\Omega_j}(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \geq \varepsilon_1) \leq \delta, \tag{7.3}$$

where δ depends on N, Δt, $t = j\Delta t$, and the initial data $\tilde{u}^0$ but not on ω_j^*. In other words, we show that

$$P(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \geq \varepsilon \mid \text{the past up to time } (j-1)\Delta t) \leq \delta$$

for all pasts up to time $(j-1)\Delta t$. This inequality follows from a probability inequality of exponential type for the pointwise error at N points in $(-L, L)$.

7.2. *Pointwise Estimates.* In this section we investigate the size of the pointwise error $|D_{\Delta t}\tilde{v}^j(x) - \tilde{D}_{\Delta t}\tilde{v}^j(x)|$. The principal result is that

$$P_{\Omega_j}(|D_{\Delta t}\tilde{v}^j(x) - \tilde{D}_{\Delta t}\tilde{v}^j(x)| \geq \alpha N\bar{w}) \leq 2e^{-2N\alpha^2} \tag{7.4}$$

holds uniformly for all $x \in \mathbf{R}$ where $\bar{w} = \max_i w_i^j$. We begin by showing that the expected value of the random walk at time $j\Delta t$ is precisely the exact solution of the heat equation with initial data $\tilde{v}^j$.

LEMMA 7.1. *Fix $\tilde{u}^0 \in \mathbf{S}$. Then for all $x \in \mathbf{R}$,*

$$E_{\Omega_j}[\tilde{D}_{\Delta t}\tilde{v}^j(x)] = D_{\Delta t}\tilde{v}^j(x).$$

Proof. Suppose that $\tilde{v}^j(x) = H(X - x)$ where $X \in \mathbf{R}$ is arbitrary. Then, by (3.4), $\tilde{D}_{\Delta t}\tilde{v}^j(x) = H(X + \eta - x)$ where η is a Gaussian random variable with mean 0 and variance $2\Delta t$. Therefore, since η and X are independent, we find

$$E_{\Omega_j}[\tilde{D}_{\Delta t}\tilde{v}^j(x)] = \frac{1}{\sqrt{4\pi\Delta t}} \int_{-\infty}^{\infty} H(X + \eta - x) e^{-\eta^2/4\Delta t}\, d\eta$$
$$= (G_{\Delta t} * \tilde{v}^j)(x) = D_{\Delta t}\tilde{v}^j(x).$$

In general, $\tilde{v}^j$ is of the form given by (2.6) and $\tilde{D}_{\Delta t}\tilde{v}^j$ is of the form given by (3.4). The lemma now follows from the linearity of E and $D_{\Delta t}$ and the independence of the X_i^{j-1} and the η_i^j. □

We now estimate the size of the pointwise error due to approximating $D_{\Delta t}\tilde{v}^j(x)$ by $\tilde{D}_{\Delta t}\tilde{v}^j(x)$. Following Roberts [31], we use a probability inequality of exponential type due to Hoeffding [24]. Similar inequalities may be found in Loève [25].

LEMMA 7.2 (Hoeffding [24, p. 16]). *Let $Z_1, \ldots, Z_N$ be N independent random variables satisfying $0 \le Z_i \le 1$. Then for all $\alpha > 0$,*

$$P\left(\frac{1}{N}\sum_{i=1}^{N} Z_i - \frac{1}{N}\sum_{i=1}^{N} E[Z_i] \ge \alpha\right) \le e^{-2N\alpha^2}.$$

Applying this lemma twice, once to the Z_i and once to the $1 - Z_i$, and then using the fact that $P(|X - Y| \ge \alpha) \le P(X - Y \ge \alpha) + P(Y - X \ge \alpha)$ for all X, Y, we obtain the following more useful form of this inequality.

COROLLARY 7.3. *Let $Z_1, \ldots, Z_N$ be N independent random variables satisfying $0 \le Z_i \le 1$. Then for all $\alpha > 0$,*

$$P\left(\left|\frac{1}{N}\sum_{i=1}^{N} Z_i - \frac{1}{N}\sum_{i=1}^{N} E[Z_i]\right| \ge \alpha\right) \le 2e^{-2N\alpha^2}.$$

Now define $\bar{w}$ as above and note that for each fixed $x \in \mathbf{R}$,

$$Z_i = H(X_i + \eta_i - x) w_i \bar{w}^{-1}$$

satisfies the hypotheses of Corollary 7.3. Furthermore,

$$(N\bar{w})^{-1}\tilde{D}_{\Delta t}\tilde{v}^j(x) = N^{-1}\sum Z_i$$

and hence, by Lemma 7.1, $(N\bar{w})^{-1}D_{\Delta t}\tilde{v}^j(x) = N^{-1}\sum E[Z_i]$. Applying Corollary 7.3 to the Z_i we obtain the following result.

LEMMA 7.4. *Let $\tilde{u}^0 \in \mathbf{S}$ and let $\bar{w} = \max_i w_i^j$. Then for all $\alpha > 0$ the inequality* (7.4) *holds uniformly for all $x \in \mathbf{R}$.*

Remark. If $w_i^0 = O(N^{-1})$, then this estimate depends exclusively on the parameter α, the number of particles N, and the time $t = j\Delta t$. For, by Lemma 2.4, $N\bar{w}$ is $O(e^t)$ for any $\tilde{v}^j$ which has been generated by the random gradient method from initial data with particle strengths that are $O(N^{-1})$.

7.3. *The L^1 Convergence of $\tilde{D}_{\Delta t}$.* We will now use the above pointwise estimate to derive the error bound in the L^1 norm. We begin by establishing a bound of the form given by (7.3).

THEOREM 7.5. *Let $\tilde{u}^0 \in \mathbf{S}$ be generated by N particles and let $\bar{w} = \max_i w_i^j$. Then for all real α, $L > 0$,*

$$P_{\Omega_j}\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \geq 2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right]\right) \leq 2Ne^{-2N\alpha^2}. \tag{7.5}$$

Proof. For the duration of this proof we drop the superscript j from $\tilde{v}^j$ and write $\tilde{v}$. By Corollary 2.3, $\tilde{v} \in \mathbf{S}$ and hence $D_{\Delta t}\tilde{v}$ and $\tilde{D}_{\Delta t}\tilde{v}$ are monotonically decreasing functions bounded between 0 and 1. Therefore, we can find a sequence $a_1, a_2, \ldots, a_N$ with $-L = a_1 < a_2 < \cdots < a_N = L$ such that $|D_{\Delta t}\tilde{v}(a_{r-1}) - D_{\Delta t}\tilde{v}(a_r)| < \sqrt{N}^{-1}$ for $r = 2, \ldots, N$. For each r let $\theta(a_r) = |D_{\Delta t}\tilde{v}(a_r) - \tilde{D}_{\Delta t}\tilde{v}(a_r)|$. Since $D_{\Delta t}\tilde{v}$ and $\tilde{D}_{\Delta t}\tilde{v}$ are monotone decreasing functions of x, it follows that for each $x \in (a_{r-1}, a_r)$,

$$\begin{aligned} D_{\Delta t}\tilde{v}(x) - \tilde{D}_{\Delta t}\tilde{v}(x) &\leq |D_{\Delta t}\tilde{v}(a_{r-1}) - D_{\Delta t}\tilde{v}(a_r)| + \theta(a_r) \\ &\leq \frac{1}{\sqrt{N}} + \max\{\theta(a_{r-1}), \theta(a_r)\}. \end{aligned}$$

Similarly, one can show that $-(D_{\Delta t}\tilde{v}(x) - \tilde{D}_{\Delta t}\tilde{v}(x)) \leq \sqrt{N}^{-1} + \max\{\theta(a_{r-1}), \theta(a_r)\}$ and hence, for $x \in (a_{r-1}, a_r)$,

$$|D_{\Delta t}\tilde{v}(x) - \tilde{D}_{\Delta t}\tilde{v}(x)| \leq \frac{1}{\sqrt{N}} + \max\{\theta(a_{r-1}), \theta(a_r)\}.$$

This yields the following estimate for the L^1 norm over the interval $(-L, L)$:

$$\begin{aligned} \|D_{\Delta t}\tilde{v} - \tilde{D}_{\Delta t}\tilde{v}\|_{(-L,L)} &\leq \sum_{j=2}^{N}(a_r - a_{r-1})\left[\frac{1}{\sqrt{N}} + \max\{\theta(a_{r-1}), \theta(a_r)\}\right] \\ &\leq 2L\left[\frac{1}{\sqrt{N}} + \max_r \theta(a_r)\right]. \end{aligned}$$

The function $\Theta = \max_r \theta(a_r)$ is a random variable which depends on the $\eta_1, \ldots, \eta_N$. The probability that the error over the interval $(-L, L)$ is greater than $2L[\sqrt{N}^{-1} + \alpha N\bar{w}]$ can be estimated in terms of the probability that $\Theta > \alpha N\bar{w}$. To see this, note that by the last inequality above,

$$\begin{aligned} \|D_{\Delta t}\tilde{v} - \tilde{D}_{\Delta t}\tilde{v}\|_{(-L,L)} &\leq 2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right] \\ &\Rightarrow 2L\left[\frac{1}{\sqrt{N}} + \Theta\right] \geq 2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right] \Leftrightarrow \Theta \geq \alpha N\bar{w}. \end{aligned}$$

Therefore, since $\Theta \geq \alpha N\bar{w}$ implies $|D_{\Delta t}\tilde{v}(a_r) - \tilde{D}_{\Delta t}\tilde{v}(a_r)| \geq \alpha N\bar{w}$ for some a_r, we can use Lemma 7.4 to obtain

$$\begin{aligned} &P_{\Omega_j}\left(\|D_{\Delta t}\tilde{v} - \tilde{D}_{\Delta t}\tilde{v}\|_{(-L,L)} \geq 2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right]\right) \\ &\quad \leq P_{\Omega_j}(\exists r\colon |D_{\Delta t}\tilde{v}(a_r) - \tilde{D}_{\Delta t}\tilde{v}(a_r)| \geq \alpha N\bar{w}) \\ &\quad \leq \sum_{r=1}^{N} P_{\Omega_j}(|D_{\Delta t}\tilde{v}(a_r) - \tilde{D}_{\Delta t}\tilde{v}(a_r)| \geq N\bar{w}\alpha) \leq 2Ne^{-2N\alpha^2}. \quad \square \end{aligned}$$

Using (7.1), (7.2), and (7.5), we can now establish a bound which holds over the entire space (Ω, Σ, μ).

COROLLARY 7.6. *Let $\tilde{u}^0 \in \mathbf{S}$ be generated by N particles and let $\bar{w} = \max_i w_i^j$. Then for all $\alpha, L > 0$,*

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \geq 2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right]\right) \leq 2Ne^{-2N\alpha^2}. \tag{7.6}$$

The next step is to prove a probability inequality for the error over the tails $(-L,L)^c$. Note that we are still free to choose L. Let $K > 0$ be chosen so that at time $t = 0$ all the particles lie in $(-K,K)$. Let $L = K + 2\beta$ where $\beta > 0$ is an arbitrary parameter. The idea, due to Roberts [31], is to estimate the error as a function of β under the assumption that at times $(j-1)\Delta t$ and $j\Delta t$ the particles remain in the interval $(-K-\beta, K+\beta)$. This reduces the problem to that of finding the probability that the particles are in this interval at the $(j-1)$st and jth time steps, a problem which may be solved with Lemmas 6.1 and 6.2.

THEOREM 7.7. *Assume that $\tilde{u}^0 \in \mathbf{S}$ is generated by N particles, all of which lie in the interval $(-K,K)$. Denote the time by $t = j\Delta t$. Let $\beta > 0$ and $L = K + 2\beta$. Then*

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)^c} \geq \frac{2\sqrt{\Delta t}}{\sqrt{\pi}}e^{-\beta^2/4\Delta t}\right) \leq \frac{4N\sqrt{t}}{\beta\sqrt{\pi}}e^{-\beta^2/4t}. \tag{7.7}$$

Proof. Let $B = K + \beta$. We first show that for $(a,b) = (-\infty, -L), (L, \infty)$,

$$\begin{aligned} &\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(a,b)} \\ &\quad \geq \frac{\sqrt{\Delta t}}{\sqrt{\pi}}e^{-\beta^2/4\Delta t} \Rightarrow \exists i \text{ such that } X_i^{j-1} \text{ or } X_i^j \notin (-B,B). \end{aligned} \tag{7.8}$$

To begin, let $(a,b) = (-\infty, -L)$. We prove the contrapositive of (7.8). Therefore, assume

$$-B < X_i^{j-1}, X_i^j < B \quad \forall i \in \{1, \ldots, N\}.$$

By the triangle inequality,

$$\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-\infty,-L)} \leq \|D_{\Delta t}\tilde{v}^j - 1\|_{(-\infty,-L)} + \|1 - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-\infty,-L)}$$

and hence, it suffices to show that

$$\|D_{\Delta t}\tilde{v}^j - 1\|_{(-\infty,-L)} < \frac{\sqrt{\Delta t}}{\sqrt{\pi}}e^{-\beta^2/4\Delta t} \quad \text{and} \quad \|1 - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-\infty,-L)} = 0.$$

The second statement is an immediate consequence of the fact that $\tilde{D}_{\Delta t}\tilde{v}^j(x) = 1$ for all $x < -L < -B$. To prove the first statement, we use $0 \leq 1 - D_{\Delta t}\tilde{v}^j(x) \leq 1 - D_{\Delta t}H(-B-x)$ for all $x \in \mathbf{R}$ to show

$$\|1 - D_{\Delta t}\tilde{v}^j\|_{(-\infty,-L)} \leq \|D_{\Delta t}H\|_{(-\infty,-\beta)} = \|G_{\Delta t} * H\|_{(-\infty,-\beta)} < \frac{\sqrt{\Delta t}}{\sqrt{\pi}}e^{-\beta^2/4\Delta t}.$$

A similar argument can be used to prove (7.8) with $(a,b) = (L, \infty)$. All that remains is to estimate the probability that the right-hand side of (7.8) is true. Since $\phi(x)$

is an increasing function of x, Lemmas 6.1 and 6.2 imply

$$P(\exists i\colon X_i^{j-1} \text{ or } X_i^j \notin (-B,B)) \le \sum_{i=1}^N P(|X_i^{j-1}| > K+\beta) + \sum_{i=1}^N P(|X_i^j| > K+\beta)$$
$$\le 2N\phi\left(\frac{-\beta}{\sqrt{2(j-1)\Delta t}}\right) + 2N\phi\left(\frac{-\beta}{\sqrt{2j\Delta t}}\right)$$
$$\le \frac{4N\sqrt{j\Delta t}}{\beta\sqrt{\pi}} e^{-\beta^2/4j\Delta t}. \quad \square$$

Corollary 7.6, Theorem 7.7, and an appropriate choice of α and β now yields (3.10).

THEOREM 7.8. *Let $\tilde{u}^0 \in \mathbf{S}$ be generated by $N \ge 3$ particles, each with weight $w_i^0 = N^{-1}$, and assume that all of these particles lie in $(-K,K)$. Denote the time by $t = j\Delta t$. Then for all $\gamma \ge 1$, the inequality in* (3.10) *holds with*

$$B_2(t) = 2\left[(K+6\sqrt{t})(1+2e^t) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\right]. \tag{7.9}$$

Proof. Fix $\gamma \in \mathbf{R}$ with $\gamma \ge 1$ and let $\alpha = 3\sqrt{\gamma \ln N}/\sqrt{8N}$, $\beta = 3\sqrt{\gamma t \ln N}$, and $L = K + 2\beta$. Referring to (7.6), we have $2Ne^{-2N\alpha^2} = 2Ne^{-\frac{9}{4}\gamma \ln N} \le 2N^{-\frac{5}{4}\gamma}$. Furthermore, since γ and $\ln N$ are both ≥ 1,

$$2L\left[\frac{1}{\sqrt{N}} + \alpha N\bar{w}\right] = 2(K+6\sqrt{\gamma t}\sqrt{\ln N})\left(\frac{1}{\sqrt{N}} + \frac{3\sqrt{\gamma}}{\sqrt{8}}\frac{\sqrt{\ln N}}{\sqrt{N}} N\bar{w}\right)$$
$$\le 2\gamma(K+6\sqrt{t})(1+2e^t)\frac{\ln N}{\sqrt{N}},$$

where we have used Lemma 2.4, together with $w_i^0 = N^{-1}$, to deduce that $N\bar{w} \le e^t$. Consequently, (7.6) becomes

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)} \ge 2\gamma(K+6\sqrt{t})(1+2e^t)\frac{\ln N}{\sqrt{N}}\right) \le 2N^{-\frac{5}{4}\gamma}. \tag{7.10}$$

With our choice of β the right-hand side of (7.7) becomes

$$\frac{4N\sqrt{t}}{\beta\sqrt{\pi}} e^{-\beta^2/4t} = \frac{4N}{3\sqrt{\gamma\pi}} \frac{N^{-\frac{9}{4}\gamma}}{\sqrt{\ln N}} \le N^{-\frac{5}{4}\gamma}.$$

Furthermore, since $j \ge 1$ and $\gamma \ge 1$, we have

$$\frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{-\beta^2/4\Delta t} = \frac{\sqrt{\Delta t}}{\sqrt{\pi}} N^{-\frac{9}{4}j\gamma} \le \gamma \frac{\sqrt{\Delta t}}{\sqrt{\pi}} \frac{1}{\sqrt{N}}.$$

Substituting these two inequalities into (7.7) yields

$$P\left(\|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_{(-L,L)^c} \ge 2\gamma \frac{\sqrt{\Delta t}}{\sqrt{\pi}} \frac{1}{\sqrt{N}}\right) \le N^{-\frac{5}{4}\gamma}.$$

Combining this estimate with (7.10) above yields (3.10). $\square$

8. Convergence of the Random Gradient Method. We now prove the convergence of the random gradient method. We begin by showing that if the hypotheses listed in Subsection 8.1 below hold, then for all $\gamma \geq 1$,

$$P\left(\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 \geq \gamma\left[e^T\|u^0 - \tilde{u}^0\|_1 + C_1\Delta t + C_2\frac{\ln N}{\sqrt[4]{N}}\right]\right) \leq \frac{4T}{C_0}N^{-\gamma}, \tag{8.1}$$

where the constants C_1 and C_2 depend only on $u^0, \tilde{u}^0$, and the time $T = k\Delta t$. The most important hypothesis here is that for some constant $C_0, \Delta t = C_0/\sqrt[4]{N}$. This has the effect of balancing the error due to the time step (temporal discretization) with the error due to the number of particles (spatial discretization). This inequality tells us that the probability of the error being greater than

$$\gamma[e^T\|u^0 - \tilde{u}^0\|_1 + C_1\Delta t + C_2 \ln N/\sqrt[4]{N}]$$

decreases exponentially as a function of γ. This allows us to find the bounds for the expected value and the variance of the error given in (1.4) and (1.5). Both of these estimates follow from (8.1) and the well-known fact (see Chung [10, p. 42]) that for any random variable $Z \geq 0$ and any real number $a > 0$,

$$E[Z] \leq a\left(1 + \sum_{r=1}^{\infty} P(Z \geq ra)\right). \tag{8.2}$$

8.1. *The Hypotheses.* Throughout this section let $T = k\Delta t$ denote the time at which we wish to compare the computed solution with the exact solution. We assume that the following hypotheses hold:

Hypothesis A_1. In addition to (1.3c–e) the exact initial data u^0 satisfies $u^0 \in C^1(\mathbf{R})$, $0 \leq u^0(x) \leq 1$ for all $x \in \mathbf{R}$, and $u_x^0 \in L^1(\mathbf{R}) \cap L^\infty(\mathbf{R})$.

Hypothesis A_2. The approximate initial data $\tilde{u}^0$ satisfies $\tilde{u}^0 \in \mathbf{S}$ (see Subsection 2.3), $\tilde{u}^0$ is generated by $N \geq 10$ particles, and the initial weights satisfy $w_i^0 = N^{-1}$.

Hypothesis A_3. The computational parameters N and Δt have been chosen so that for some constant C_0 we have

$$\Delta t = \frac{C_0}{\sqrt[4]{N}}. \tag{8.3}$$

We also assume that the constant $K > 0$ has been chosen so that the variation of $\tilde{u}^0$ lies in $(-K, K)$,

$$|X_i^0| < K, \qquad i = 1, \ldots, N. \tag{8.4}$$

8.2. *A Bound on the Probability Distribution of the Error.* The proof of (8.1) is accomplished in two steps. In the first step (Theorem 8.1) we use the estimates from Sections 6 and 7 to establish the probability inequality (3.7) for the error due to the approximate solution operators. In the second step (Theorem 8.2) we use the deterministic bounds from Sections 4 and 5 to control the remaining sources of error.

THEOREM 8.1. *Assume that Hypotheses* A_1, A_2, *and* A_3 *hold and let* B_1 *be given by* (6.3) *and* B_2 *by* (7.9). *Then the inequality in* (3.7) *holds for all* $\gamma \geq 1$ *with*

$$C_2 = \{B_1(T)C_0^2 + B_2(T)\}\frac{Te^T}{C_0}. \tag{8.5}$$

Proof. Let $Z_j = \|(D_{\Delta t}R_{\Delta t})^j\tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^j\tilde{u}^0\|_1$, $V_j = \|R_{\Delta t}\tilde{u}^{j-1} - \tilde{R}_{\Delta t}\tilde{u}^{j-1}\|_1$, and $W_j = \|D_{\Delta t}\tilde{v}^j - \tilde{D}_{\Delta t}\tilde{v}^j\|_1$. Repeated application of Lemma 4.1 and Corollary 4.4 yields $Z_k \le e^T \sum_{j=1}^k (V_j + W_j)$. (This is (3.8).) Since $Z_0 = 0$, $k = T\sqrt[4]{N}/C_0$, and $(\Delta t)^2 = C_0^2/\sqrt{N}$, we have

$$Z_k \ge \gamma C_2 \frac{\ln N}{\sqrt[4]{N}} \Rightarrow \sum_{j=1}^k (V_j + W_j) \ge \frac{\gamma C_2}{e^T}\frac{\ln N}{\sqrt[4]{N}}$$
$$\Rightarrow \exists j \text{ such that } (V_j + W_j) \ge \frac{\gamma C_2}{ke^T}\frac{\ln N}{\sqrt[4]{N}} = \frac{\gamma C_0 C_2}{Te^T}\frac{\ln N}{\sqrt{N}}$$
$$\Rightarrow \exists j \text{ such that } V_j \ge \gamma B_1(T)(\ln N)(\Delta t)^2 \text{ or } W_j \ge \gamma B_2(T)\frac{\ln N}{\sqrt{N}}.$$

We now apply Corollary 6.4 and Theorem 7.8 to obtain

$$P\left(Z_k \ge \gamma C_2 \frac{\ln N}{\sqrt[4]{N}}\right) \le \sum_{j=1}^k P(V_j \ge \gamma B_1 \ln N(\Delta t)^2) + \sum_{j=1}^k P\left(W_j \ge \gamma B_2 \frac{\ln N}{\sqrt{N}}\right)$$
$$\le 4kN^{-\frac{5}{4}\gamma} \le \frac{4T}{C_0}N^{-\gamma}. \quad \square$$

Using Theorem 8.1 and the bounds from Theorems 4.5 and 5.4, we now derive (8.1).

THEOREM 8.2. *Assume that Hypotheses* A_1, A_2, *and* A_3 *hold. Then for all* $\gamma \ge 1$ *the inequality in* (8.1) *holds with* C_1 *given by* (5.8) *and* C_2 *by* (8.5).

Proof. Applying Theorems 4.5 and 5.4 to (3.5), we see that

$$\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1 \le C_1\Delta t + e^T\|u^0 - \tilde{u}^0\|_1 + \|(D_{\Delta t}R_{\Delta t})^k\tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1.$$

Let $h(T, \Delta t, N) = C_1\Delta t + e^T\|u^0 - \tilde{u}^0\|_1 + C_2 \ln N/\sqrt[4]{N}$. Since $\gamma \ge 1$,

$$\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1 \ge \gamma h(T, \Delta t, N)$$
$$\Rightarrow \|(D_{\Delta t}R_{\Delta t})^k\tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1 \ge \gamma C_2\frac{\ln N}{\sqrt[4]{N}}.$$

It now follows from Theorem 8.1 that

$$P(\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1 \ge \gamma h(T, \Delta t, N))$$
$$\le P\left(\|(D_{\Delta t}R_{\Delta t})^k\tilde{u}^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k\tilde{u}^0\|_1 \ge \gamma C_2\frac{\ln N}{\sqrt[4]{N}}\right)$$
$$\le \frac{4T}{C_0}N^{-\gamma}. \quad \square$$

8.3. *The Expected Value and Variance of the Error.* We now prove the bounds on the expected value and the variance of the error in (1.4) and (1.5) for $\nu = 1$. We remove this assumption in Subsection 8.4.

THEOREM 8.3. *Assume that Hypotheses* A_1, A_2, *and* A_3 *hold. Let* C_1 *be given by* (5.8) *and* C_2 *by* (8.5). *Then the inequality in* (1.4) *holds when* $\nu = 1$.

Proof. Define h as above. Setting $\gamma = 1, 2, \dots$ in (8.1) and applying (8.2), we obtain

$$E[\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1] \le h(T, \Delta t, N)\left(1 + \frac{4T}{C_0}\sum_{r=1}^{\infty} N^{-r}\right)$$
$$\le h(T, \Delta t, N)\left(1 + \frac{T}{C_0}\right). \quad \square$$

THEOREM 8.4. *Assume that Hypotheses* A_1, A_2, *and* A_3 *hold. Let* C_1 *be given by* (5.8) *and* C_2 *by* (8.5). *Then the inequality in* (1.5) *holds when* $\nu = 1$.

Proof. Set $Z = \|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1$. Since $\operatorname{var}(Z) = E[Z^2] - E[Z]^2 \le E[Z^2]$, it suffices to bound $E[Z^2]$. From Theorem 8.2 we have

$$P(Z^2 \ge (\gamma h(T, \Delta t, N))^2) = P(Z \ge \gamma h(T, \Delta t, N)) \le \frac{4T}{C_0} N^{-\gamma}.$$

Setting $\gamma = \sqrt{r}$ for each $r \in \{1, 2, \dots\}$ and applying (8.2) we find

$$E[Z^2] \le h(T, \Delta t, N)^2 \left(1 + \frac{4T}{C_0}\sum_{r=1}^{\infty} N^{-\sqrt{r}}\right).$$

To estimate the series on the right set $b = N^{-1}$ and $\beta = -\ln b$. Then

$$\sum_{r=1}^{\infty} b^{\sqrt{r}} \le b + \int_1^{\infty} e^{-\beta\sqrt{x}}\, dx = b - \left[\left(\frac{2}{\beta^2} + \frac{2\sqrt{x}}{\beta}\right) e^{-\beta\sqrt{x}}\right]_1^{\infty}$$
$$= b + \left(\frac{2}{\beta^2} + \frac{2}{\beta}\right) e^{-\beta} \le \frac{5}{2} N^{-1},$$

where we have used the hypothesis that $N \ge 10$ and hence $\beta = \ln N > 2$. Thus,

$$E[Z^2] \le h(T, \Delta t, N)^2 \left(1 + \frac{10T}{C_0} N^{-1}\right) \le \left(1 + \frac{T}{C_0}\right) h(T, \Delta t, N)^2. \quad \square$$

8.4. *Dependence of the Error on Arbitrary* $\nu \le 1$. We will now remove the restriction $\nu = 1$. For arbitrary positive $\nu \le 1$ let u_ν be the solution of Eq. (1.3a,b) with initial data u_ν^0. Define $u(x,t) \equiv u_\nu(\sqrt{\nu}x, t)$. Then u satisfies (1.3a,b) with diffusion coefficient 1 and initial data $u^0(x) \equiv u_\nu^0(\sqrt{\nu}x)$. Note that $\|\partial_x u\|_\infty = \sqrt{\nu}\|\partial_x u_\nu\|_\infty$ and $\|\partial_x u\|_1 = \|\partial_x u_\nu\|_1$.

The random gradient method scales in the same manner. In other words, let $\tilde{u}_\nu^k$ be the random gradient solution of (1.3a,b) at time $T = k\Delta t$ with diffusion coefficient ν. Denote the initial particle positions by $X_i^0(\nu)$. Then for any $k \ge 0$, $\tilde{u}^k(x) \equiv \tilde{u}_\nu^k(\sqrt{\nu}x)$ is the random gradient solution of (1.3a,b) with diffusion coefficient 1 and initial particle positions $X_i^0 = X_i^0(\nu)/\sqrt{\nu}$. This statement follows immediately from the fact that if η is a Gaussian distributed random variable with variance $2\nu\Delta t$, then $\eta/\sqrt{\nu}$ is a Gaussian random variable with variance $2\Delta t$.

It now follows that $\|u_\nu(T) - \tilde{u}_\nu^k\|_1 = \sqrt{\nu}\|u(T) - \tilde{u}^k\|_1$ and hence, by Theorem 8.3,

$$E(\|u_\nu(T) - \tilde{u}_\nu^k\|_1) = \sqrt{\nu}E(\|u(T) - \tilde{u}^k\|_1)$$
$$\le \sqrt{\nu}\left(1 + \frac{T}{C_0}\right)\left[e^T\|u^0 - \tilde{u}^0\|_1 + C_1\Delta t + C_2\frac{\ln N}{\sqrt[4]{N}}\right].$$

It is necessary, however, to investigate the dependence of the constants C_1 and C_2 on ν. From (5.8) we have

$$
\begin{aligned}
C_1 &= Te^{4T}\left\{e^T\|\partial_x u^0\|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}}\right\}\|\partial_x u^0\|_1 \\
&= Te^{4T}\left\{\sqrt{\nu}e^T\|\partial_x u_\nu^0\|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}}\right\}\|\partial_x u_\nu^0\|_1.
\end{aligned} \tag{8.6}
$$

Thus the splitting error is $O(\sqrt{\nu})$. Note that if one is modeling a wave front initially of the form $g_\nu(x) \equiv g(x/\sqrt{\nu})$ for some arbitrary C^1 function g, then, even though $\partial_x g_\nu = O(\nu^{-1/2})$, the constant C_1 remains $O(1)$ owing to the factor $\sqrt{\nu}$ multiplying $\|\partial_x u_\nu^0\|_\infty$ in (8.6) above.

In order to examine the dependence of C_2 on ν, let $K(\nu)$ be chosen so that $-K(\nu) \le X_i^0(\nu) \le K(\nu)$ for all i. Then $K \equiv K(\nu)/\sqrt{\nu}$ satisfies (8.4) and we find

$$
C_2 = \frac{\sqrt{3}}{9}\left(\frac{K(\nu)}{\sqrt{\nu}} + 3\sqrt{T}\right)C_0^2 + 2\left[\left(\frac{K(\nu)}{\sqrt{\nu}} + 6\sqrt{T}\right)(1+2e^T) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\right]\frac{Te^T}{C_0}.
$$

Hence, $C_2' = \sqrt{\nu}C_2$ is bounded uniformly in ν for $\nu \le 1$ as claimed. The conclusion that C_2 is $O(\nu^{-1/2})$ may be misleading however. For example, with waves initially of the form given by g_ν above, one generally chooses the approximate initial data so that $K(\nu) = O(\sqrt{\nu})$. In this case $C_2 = O(1)$.

A similar argument can be used to establish the validity of the bound on the variance of the error in (1.5) for arbitrary $\nu \le 1$.

9. Numerical Results. In order to compare our theoretical bounds with the actual performance of the method, we use it to compute a known exact solution. We also present the results of computations with a second-order solution of (3.1a,b), an exact solution of (3.1a,b) and second-order operator splitting (Strang splitting). These experiments allow us to test the sharpness of our estimates and our understanding of the way various sources of error behave.

9.1. *The Test Problem.* For $\nu = 1$ the Kolmogorov equation (1.3a,b) has a traveling wave solution of the form

$$
u(x,t) = g(x - \alpha t) \tag{9.1}
$$

with speed $\alpha = 5/\sqrt{6}$ and wave form $g(x) = (1 + (\sqrt{2}-1)e^{x/\sqrt{6}})^{-2}$. Our approximation to u^0 was determined by placing N particles, each with weight $w_i^0 = N^{-1}$, at

$$
X_i^0 = \begin{cases} g^{-1}\left(1 - \frac{i}{N}\right), & i = 1,\dots,N-1, \\ g^{-1}\left(\frac{1}{2N}\right), & i = N. \end{cases}
$$

For this choice of $\tilde{u}^0$ we have

$$
\|u^0 - \tilde{u}^0\|_1 = O\left(\frac{1}{N}\right). \tag{9.2}
$$

For u given by (9.1) define the center of the wave at time t to be the point $x_c = x_c(t)$ such that $u(x_c,t) = 1/2$. We measure the error at time $T = 1$ on a grid of 1001 points centered at x_c and spaced a distance $\Delta x = 0.02$ apart. We estimate the expected value of this error by averaging it over twenty independent trials.

Denote this average by $\overline{\text{err}} = \overline{\text{err}}(\Delta t, N)$. The independence of the trials with respect to one another has been achieved by starting each trial with a different, independently chosen seed for the random number generator. We also report the standard deviation of our sample of twenty errors about their mean, denoted by $\sigma = \sigma(\Delta t, N)$. The error in each of the discrete L^1, L^2, and L^∞ norms decreases at roughly the same rate. We present only the results in the L^1 norm.

9.2. *Numerical Results.* Table 1 contains $\overline{\text{err}}$ and σ for a computation with the random gradient method as described in Section 2. The number of particles increases by 4 as one moves to the right along a row, while the time step decreases by 2 as one moves down a column. The average error $\overline{\text{err}}$ roughly decreases by 2 as one moves diagonally down one row and right one column. We therefore conclude that for this problem the proper relationship between Δt and N is

$$\Delta t = O\left(\frac{1}{\sqrt{N}}\right). \tag{9.3}$$

The relation $\Delta t = O(\sqrt[4]{N}^{-1})$ which we have arrived at by theoretical considerations appears to be an underestimate of the dependence of the error on N. In other words, if we set $\Delta t = O(\sqrt[4]{N}^{-1})$, then the errors that depend on N will decrease twice as fast as the errors that depend on Δt, until eventually these latter sources of error dominate all others. The method will still converge, but we will be doing four times as much work*** to get the same results.

TABLE 1

Estimated mean and standard deviation of the error in the L^1 norm.

First-order solution of the ODE $u_t = u(1-u)$.

	Number of Particles					
Δt	1000	4000	16000	64000	256000	1024000
1	.4568 ± .0659	.4361 ± .0334	.4431 ± .0095	.4448 ± .0057	.4435 ± .0020	.4434 ± .0010
2^{-1}	.2345 ± .0473	.2240 ± .0297	.2160 ± .0124	.2220 ± .0063	.2188 ± .0024	.2191 ± .0014
4^{-1}	.1396 ± .0380	.1204 ± .0220	.1106 ± .0099	.1146 ± .0069	.1121 ± .0027	.1118 ± .0013
8^{-1}	.1107 ± .0167	.0692 ± .0170	.0585 ± .0087	.0598 ± .0072	.0568 ± .0029	.0569 ± .0013
16^{-1}	.0977 ± .0204	.0525 ± .0120	.0333 ± .0070	.0328 ± .0053	.0287 ± .0029	.0287 ± .0012
32^{-1}	.0973 ± .0181	.0461 ± .0122	.0250 ± .0063	.0187 ± .0034	.0152 ± .0030	.0142 ± .0015

Fix $\Delta t = 1$ and note that for all N, $\overline{\text{err}}(1, N)$ is within $\sigma(1, N)$ of 0.443, i.e., for $\Delta t = 1, \overline{\text{err}}$ is well within statistical error of being constant. This is because those sources of error due to temporal discretization (e.g., the errors in (3.6) and (6.2)) dominate those errors due to spacial discretization (e.g., (9.2)). These two sources of error are roughly in balance when $\Delta t = 4^{-1}$ and $N = 1000$. This can be seen by noting that if one moves to the right along the row or down the column from this point, then the error remains within $2\sigma(4^{-1}, 1000)$ of being constant.

We can determine the dependence of the error on one of the parameters by letting the other be a small, fixed value and observing how the error behaves as

***We have neglected the cost of sorting the particles at the end of each time step.

a function of the first. For example, for fixed $N = 1024000$, $\overline{\text{err}}$ clearly decreases like $O(\Delta t)$. On the other hand, if we fix $\Delta t = 32^{-1}$ we find that for $N < 64000$ the error decays like $\sqrt{N}^{-1}$. The error is (statistically) constant for $N \geq 64000$ because for $(\Delta t, N) = (32^{-1}, 64000)$ the two sources of error are again roughly in balance and hence, to the right of this point, the dominant source of error is that due to the time step Δt.

Note that this balance point lies on the same diagonal as the one found earlier, $\Delta t = 4^{-1}$, $N = 1000$. This diagonal represents the optimal choice of computational parameters. On either side of this diagonal we would be doing more work to achieve the same level of error. Of course, in practice it is usually impossible to determine this diagonal. Therefore, the best strategy is to simply refine the parameters at the optimal rate, presumably that given by (9.3). Although this will not necessarily result in the least amount of work for a given level of error, it will result in the error decreasing at the best possible rate.

TABLE 2

Error in the L^1 norm after one run.

First-order solution of the ODE $u_t = u(1-u)$.

	Number of Particles					
Δt	1000	4000	16000	64000	2560000	1024000
1	0.5504	0.3863	0.4548	0.4394	0.4447	0.4439
2^{-1}	0.2287	0.2110	0.2168	0.2192	0.2202	0.2215
4^{-1}	0.0995	0.1148	0.0989	0.1056	0.1125	0.1136
8^{-1}	0.1116	0.0531	0.0575	0.0579	0.0551	0.0576
16^{-1}	0.0976	0.0419	0.0276	0.0359	0.0300	0.0289
32^{-1}	0.1034	0.0453	0.0192	0.0256	0.0116	0.0137

In Table 2 we present the L^1 error at time $T = 1$ after only one trial. In other words, one realization of the random variable whose expected value and standard deviation have been estimated in Table 1. In all cases the error after one trial lies within 2σ of $\overline{\text{err}}$. Furthermore, along the diagonals the errors in Table 2 decrease very nearly at the rate of $\Delta t = O(\sqrt{N}^{-1})$. The important point to note here is that one generally obtains good results with one trial. It is not necessary to average the computed solution over several trials in order to obtain decent results.

This statement can be made rigorous in the following way. Suppose $T/C_0 = 1$. For $\gamma = 1$, inequality (8.1) implies

$$P\left(\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 < e^T \|u^0 - \tilde{u}^0\|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}\right) \geq 1 - \frac{4T}{C_0 N}. \tag{9.4}$$

The right-hand side of (9.4) is an increasing function of N and, when $N = 1000$, we have $1 - 4T/C_0 N = 996/1000$. Thus, inequality (9.4) assures us that if $N \geq 1000$,

then better than 99% of the time

$$\|F_{\Delta t}^k u^0 - (\tilde{D}_{\Delta t}\tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 < e^T \|u^0 - \tilde{u}^0\|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}.$$

We know of no way to improve the accuracy of the random walk. However, there are several ways to obtain a method which is higher-order in time. We begin by considering a second-order ODE solver. Define

$$\tilde{R}_{\Delta t}^{2\text{nd}} \tilde{u}^0(x) = \tilde{u}^0(x) + \frac{\Delta t}{2}[f(\tilde{u}^0(x)) + f(\tilde{u}^0(x) + \Delta t f(\tilde{u}^0(x)))].$$

This is simply Heun's method for solving (3.1a,b) ([11, p. 364]). Table 3 contains the result of a series of runs with $\tilde{R}_{\Delta t}$ replaced by $\tilde{R}_{\Delta t}^{2\text{nd}}$. It is immediately apparent that there has been an overall decrease in the error as compared to Table 1. However, the *rate* of convergence has not changed—even as a function of Δt alone. On the average the errors still decay like $O(\Delta t)$.

TABLE 3

Estimated mean and standard deviation of the error in the L^1 norm.

Second-order solution of the ODE.

	Number of Particles					
Δt	1000	4000	16000	64000	256000	1024000
1	.2588 ± .0516	.2394 ± .0238	.2410 ± .0117	.2414 ± .0055	.2421 ± .0029	.2417 ± .0013
2^{-1}	.1177 ± .0347	.0947 ± .0199	.0788 ± .0111	.0814 ± .0062	.0786 ± .0029	.0788 ± .0020
4^{-1}	.1009 ± .0236	.0584 ± .0132	.0348 ± .0073	.0334 ± .0055	.0297 ± .0027	.0295 ± .0020
8^{-1}	.0997 ± .0175	.0494 ± .0099	.0265 ± .0084	.0171 ± .0048	.0131 ± .0030	.0127 ± .0016
16^{-1}	.0964 ± .0200	.0476 ± .0100	.0234 ± .0057	.0135 ± .0028	.0079 ± .0022	.0061 ± .0013
32^{-1}	.0979 ± .0189	.0449 ± .0119	.0232 ± .0053	.0121 ± .0023	.0073 ± .0018	.0039 ± .0010

We interpret this data in the following way. When $\tilde{R}_{\Delta t}$ is replaced by $\tilde{R}_{\Delta t}^{2\text{nd}}$, the $(\Delta t)^2$ in (3.9) is replaced by $(\Delta t)^3$.† Thus, the dependence of the last term on the right in (3.5) on Δt is now $O((\Delta t)^2)$ rather than $O(\Delta t)$. However, its dependence on N is still $O(\sqrt{N^{-1}})$. From (9.2) we see that the middle term is $O(N^{-1})$ and hence is presumably negligible compared to the last term. However, the first term—the error due to operator splitting—remains $O(\Delta t)$.

To increase the accuracy of this first term, we now employ the following operator splitting algorithm known as *Strang splitting* [36],

$$\tilde{u}_s^{j+1} = \tilde{R}_{\Delta t/2}^{2\text{nd}} \tilde{D}_{\Delta t} \tilde{R}_{\Delta t/2}^{2\text{nd}} \tilde{u}_s^j.$$

†This statement is easily proved. We simply use the well-known fact that the local truncation error for a second-order ODE solver is $O((\Delta t)^3)$ to replace the right-hand side of (6.2) by $(\Delta t)^3$ times the appropriate constant.

Note that Strang splitting costs no more than first-order splitting. One simply takes half a time step at the beginning and another half time step at the end:

$$\tilde{u}_s^j = \tilde{R}^{2\text{nd}}_{\Delta t/2}\tilde{D}_{\Delta t}(\tilde{R}^{2\text{nd}}_{\Delta t}\tilde{D}_{\Delta t})^{j-1}\tilde{R}^{2\text{nd}}_{\Delta t/2}\tilde{u}^0.$$

With Strang splitting we expect the error due to exact operator splitting to be

$$\|F^k_{\Delta t}u^0 - (R_{\Delta t/2}D_{\Delta t}R_{\Delta t/2})^k u^0\|_1 \le C\sqrt{\nu}(\Delta t)^2.$$

Although we do not prove this statement, it should be possible to prove it by applying the argument in the proof of Theorem 5.3 to the function

$$w_s(x,t) = F_t u^0(x) - R_{t/2}D_t R_{t/2}u^0(x).$$

In this regard, we note that Beale and Majda have shown that Strang splitting for the Navier-Stokes equations is second-order accurate [2].

TABLE 4. (L^1 *norm*)

Second-order solution of the ODE with Strang splitting.

	Number of Particles					
Δt	1000	4000	16000	64000	256000	1024000
1	.1124 ± .0192	.0668 ± .0121	.0444 ± .0085	.0396 ± .0052	.0397 ± .0029	.0394 ± .0024
2^{-1}	.1008 ± .0229	.0585 ± .0115	.0324 ± .0056	.0273 ± .0054	.0236 ± .0026	.0239 ± .0015
4^{-1}	.1008 ± .0196	.0517 ± .0091	.0238 ± .0033	.0146 ± .0032	.0087 ± .0019	.0086 ± .0013
8^{-1}	.0994 ± .0184	.0484 ± .0095	.0241 ± .0057	.0129 ± .0024	.0064 ± .0011	.0036 ± .0009
16^{-1}	.0969 ± .0202	.0476 ± .0101	.0225 ± .0051	.0125 ± .0026	.0066 ± .0015	.0029 ± .0007
32^{-1}	.0982 ± .0192	.0447 ± .0118	.0233 ± .0056	.0121 ± .0025	.0067 ± .0016	.0031 ± .0006

In Table 4 we present the results of using this algorithm on the test problem. We note a further decrease in the error as compared to Tables 1 and 3. In particular, for $\Delta t = 1$ and $N \ge 16{,}000$ the error is an order of magnitude smaller than that in Table 1! Also note that for $\Delta t \le 1/8$ the errors that depend on Δt appear to be so small there is little further decrease in the error if one fixes N and lets $\Delta t \to 0$. However, the overall dependence of the error on N has not changed—the error still depends on N like $O(\sqrt{N^{-1}})$.

We conjecture that the choice of parameters which results in the first and last terms on the right in (3.5) decreasing at the same rate is now $\Delta t = O(1/\sqrt[4]{N})$. However, we find it somewhat puzzling that for large, fixed N, say $N = 1024000$, the error does not decrease like $O(\Delta t^2)$ and can offer no explanation.

Table 5 contains the data from columns 2, 4, and 6 of Table 4 organized so that, provided our conjecture is correct, the most efficient way to decrease the errors now lies on the diagonal. Note that on and below the diagonal which begins with $(\Delta t, N) = (2^{-1}, 4000)$ the error decreases at a rate roughly equal to $\Delta t^2 = O(1/\sqrt{N})$, which is consistent with our conjecture. This results in a small savings

TABLE 5

The most efficient use of resources lies along the diagonals.

Second-order solution of the ODE with Strang splitting.

	Number of Particles		
Δt	4000	64000	1024000
1	$.0668 \pm .0121$	$.0396 \pm .0052$	$.0394 \pm .0024$
2^{-1}	$.0585 \pm .0115$	$.0273 \pm .0054$	$.0239 \pm .0015$
4^{-1}	$.0517 \pm .0091$	$.0146 \pm .0032$	$.0086 \pm .0013$
8^{-1}	$.0484 \pm .0095$	$.0129 \pm .0024$	$.0036 \pm .0009$
16^{-1}	$.0476 \pm .0101$	$.0125 \pm .0026$	$.0029 \pm .0007$
32^{-1}	$.0447 \pm .0118$	$.0121 \pm .0025$	$.0031 \pm .0006$

TABLE 6

Estimated mean and standard deviation of the error in the L^1 norm.

Exact solution of the ODE.

	Number of Particles					
Δt	1000	4000	16000	64000	256000	1024000
1	$.1372 \pm .0351$	$.0993 \pm .0217$	$.0877 \pm .0105$	$.0848 \pm .0057$	$.0850 \pm .0030$	$.0843 \pm .0013$
2^{-1}	$.1053 \pm .0290$	$.0697 \pm .0168$	$.0462 \pm .0085$	$.0445 \pm .0061$	$.0413 \pm .0025$	$.0410 \pm .0020$
4^{-1}	$.1004 \pm .0224$	$.0554 \pm .0116$	$.0292 \pm .0063$	$.0251 \pm .0048$	$.0208 \pm .0026$	$.0203 \pm .0020$
8^{-1}	$.0996 \pm .0177$	$.0491 \pm .0097$	$.0259 \pm .0079$	$.0159 \pm .0044$	$.0114 \pm .0029$	$.0106 \pm .0016$
16^{-1}	$.0964 \pm .0201$	$.0476 \pm .0100$	$.0233 \pm .0056$	$.0134 \pm .0028$	$.0077 \pm .0021$	$.0056 \pm .0012$
32^{-1}	$.0979 \pm .0190$	$.0449 \pm .0119$	$.0232 \pm .0053$	$.0121 \pm .0023$	$.0073 \pm .0018$	$.0038 \pm .0010$

in computational effort. To decrease the error by four, the original version requires $N \to 16N$ and $\Delta t \to \Delta t/4$, resulting in 64 times as much work.[††] On the other hand, the higher-order method only requires 32 times as much work to achieve one fourth the error. For methods in which the work required at each time step is $O(N^2)$ the savings is proportionally smaller.

Finally, we replace $\tilde{R}_{\Delta t}$ by $R_{\Delta t}$ (it is easy to compute the exact solution of Eq. (3.1a,b)). In this experiment we do not use Strang splitting. The results are presented in Table 6. For those choices of N and Δt for which one expects the errors due to Δt to be noticeable, we find a moderate improvement over the results displayed in Table 3. On the other hand, when the $O(\sqrt{N^{-1}})$ errors dominate, the errors in Table 6 are quite close to those in Table 3. Upon comparing Table 6 with Table 4, we conclude that if one is going to go to the trouble of using a higher-order

[††]Again, we have neglected the work required to sort the particles at the end of every time step and assumed the work at every time step is $O(N)$.

solution of the ODE (3.1a), then one should also use Strang splitting, especially since it requires no additional computational effort.

9.3. *Conclusions.* The theoretical estimates presented in Theorem 1.1 are most likely an underestimate of the true rate of convergence. In order to prove this theorem, we have found it necessary to assume that $\Delta t = O(\sqrt[4]{N^{-1}})$. However, based on the numerical results presented here, we conclude that this is an inefficient choice of parameters. One can argue that this is a special test problem and that more general problems may converge at a slower rate. However, most solutions of (1.3a,b) converge to traveling wave solutions in time (e.g. [5]) and it seems likely that the method's behavior with this particular traveling wave solution is representative of its general behavior when approximating a traveling wave solution. The failure of our analysis to accurately predict the true rate of convergence is probably due to our use of the triangle inequality in (3.8). In contrast, Hald was able to establish the correct rate of convergence for the method considered in [22], precisely because he could write down the exact and computed solutions at any time t, thereby avoiding the need to apply the triangle inequality.

Lawrence Livermore National Laboratory
P.O. Box 808
Livermore, California 94550
Puckett@crg.llnl.gov

1. C. ANDERSON & C. GREENGARD, "On vortex methods," *SIAM J. Numer. Anal.*, v. 22, 1985, pp. 413–440.

2. J. T. BEALE & A. MAJDA, "Rates of convergence for viscous splitting of the Navier-Stokes equations," *Math. Comp.*, v. 37, 1981, pp. 243–259.

3. J. T. BEALE & A. MAJDA, "Vortex methods I: Convergence in three dimensions," *Math. Comp.*, v. 39, 1982, pp. 1–27.

4. J. T. BEALE & A. MAJDA, "Vortex methods II: Higher order accuracy in two and three dimensions," *Math. Comp.*, v. 39, 1982, pp. 29–52.

5. M. BRAMSON, "Convergence of solutions of the Kolmogorov equation to travelling waves," *Mem. Amer. Math. Soc.*, no. 285, 1983.

6. Y. BRENIER, *A Particle Method for One Dimensional Non-Linear Reaction Advection Diffusion Equations*, Technical Report No. 351, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autonoma de Mexico, 1983.

7. A. J. CHORIN, "Numerical study of slightly viscous flow," *J. Fluid Mech.*, v. 57, 1973, pp. 785–796.

8. A. J. CHORIN, "Vortex sheet approximation of boundary layers," *J. Comput. Phys.*, v. 27, 1978, pp. 428–442.

9. A. J. CHORIN, *Numerical Methods For Use in Combustion Modeling.* Proc. Internat. Conf. Numer. Methods in Science and Engineering, Versailles, 1979.

10. K. L. CHUNG, *A Course in Probability Theory*, Harcourt, Brace & World Inc., 1974.

11. S. D. CONTE & C. DE BOOR, *Elementary Numerical Analysis*, 3rd ed., McGraw-Hill, New York, 1980.

12. G. H. COTTET, *Méthodes Particulaires pour l'Equation d'Euler dans le Plan*, Thèse de 3ème cycle, Université Pierre et Marie Curie, Paris, 1982.

13. G. H. COTTET & S. GALLIC, *A Particle Method to Solve Transport-Diffusion Equations, Part* I: *The Linear Case*, Rapport Interne No: 115, Centre de Mathématiques Appliquées, Ecole Polytechnique, Paris, 1984.

14. G. B. FOLLAND, *Introduction to Partial Differential Equations*, Princeton Univ. Press, 1976.

15. A. F. GHONIEM & F. S. SHERMAN, "Grid free simulation of diffusion using random walk methods," *J. Comput. Phys.*, v. 61, 1985, pp. 1–37.

16. J. GOODMAN, "Convergence of the random vortex method," *Comm. Pure Appl. Math.*, v. 40, 1987, pp. 189–220.

17. C. GREENGARD, "Convergence of the vortex filament method," *Math. Comp.*, v. 47, 1986, pp. 387–398.

18. O. HALD, Private communication.

19. O. HALD & V. M DEL PRETE, "Convergence of vortex methods for Euler's equations," *Math. Comp.*, v. 32, 1978, pp. 791–809.

20. O. H. HALD, "The convergence of vortex methods for Euler's equations II," *SIAM J. Numer. Anal.*, v. 16, 1979, pp. 726–755.

21. O. H. HALD, "Convergence of random methods for a reaction-diffusion equation," *SIAM J. Sci. Statist. Comput.*, v. 2, 1981, pp. 85–94.

22. O. H. HALD, "Convergence of a random method with creation of vorticity," *SIAM J. Sci. Statist. Comput.*, v. 7, 1986, pp. 1373–1386.

23. O. H. HALD, "Convergence of vortex methods for Euler's equations, III," *SIAM J. Numer. Anal.*, v. 24, 1987, pp. 538–582.

24. W. HOEFFDING, "Probability Inequalities for Sums of Bounded Random Variables," *Amer. Statist. Assoc. J.*, v. 58, 1963, pp. 13-30.

25. M. LOÈVE, *Probability Theory*, Springer-Verlag, Berlin and New York, 1977.

26. D. G. LONG, *Convergence of the Random Vortex Method in One and Two Dimensions*, Ph.D. Thesis, Univ. of California, Berkeley, 1986.

27. E. G. PUCKETT, "A study of the vortex sheet method and its rate of convergence," *SIAM J. Sci. Statist. Comput.*, v. 10, 1989, pp. 298–327.

28. P. A. RAVIART, "An analysis of particle methods," *CIME Course on Numerical Methods in Fluid Dynamics*, Publications du Laboratoire d'Analyse numérique, Université Pierre et Marie Curie, Paris, 1983.

29. P. A. RAVIART, *Particle Approximation of Linear Hyperbolic Equations of the First Order*, Publications du Laboratoire d' Analyse numérique, Université Pierre et Marie Curie, Paris, 1983.

30. S. G. ROBERTS, "Accuracy of the random vortex method for a problem with non-smooth initial conditions," *J. Comput. Phys.*, v. 58, 1985, pp. 29–43.

31. S. ROBERTS, "Convergence of a random walk method for the Burgers equation," this issue, pp. 647–673.

32. G. ROSEN, "Brownian-motion correspondence method for obtaining approximate solutions to nonlinear reaction-diffusion equations," *Phys. Rev. Lett.*, v. 53, 1984, pp. 307–310.

33. A. S. SHERMAN & C. S PESKIN, "A Monte Carlo method for scalar reaction diffusion equations," *SIAM J. Sci. Statist. Comput.*, v. 7, 1986, pp. 1360–1372.

34. A. S. SHERMAN & C. S. PESKIN, "Solving the Hodgkin-Huxley equations by a random walk method," *SIAM J. Sci. Statist. Comput.*, v. 9, 1988, pp. 170–190.

35. J. SMOLLER, *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York, 1983.

36. G. STRANG, "On the construction and comparison of difference schemes," *SIAM J. Numer. Anal.*, v. 5, 1968, pp. 506–517.