JOURNAL OF COMPUTATIONAL PHYSICS 109, 155–163 (1993)

Optimal Smoothing in Function-Transport Particle Methods for Diffusion Problems

AARON L. FOGELSON AND ROBERT H. DILLON

Department of Mathematics, University of Utah, Salt Lake City, Utah 84112

Received August 16, 1991; revised July 9, 1992

We discuss a class of particle methods for diffusion problems with small diffusivity, in which diffusion is modeled by random walk update of the particle positions; each particle carries a point-value of the problem's initial data; and the numerical solution is obtained as a discrete convolution of the particle data with an approximate δ-function. While it is widely believed that such a particle method fails to converge, we prove that if the number of particles M and the degree of smoothing, measured by the width ϵ of the approximate δ-function, are *coupled* so that $(M\epsilon)^{-1} \to 0$ as $M \to \infty$ and $\epsilon \to 0$, then the computed solution $u_M^\epsilon(x, t)$ converges to the solution $v(x, t)$ of the diffusion equation in the sense that $E(|u_M^\epsilon(x, t) - v(x, t)|^2) \to 0$. We also present numerical results which illustrate the theory, and, in particular, show that the method performs uniformly well in the limit that the diffusion coefficient vanishes.

1. INTRODUCTION

It is widely believed that particle methods for diffusion problems in which the diffusion particles carry information about the values of the solution function (rather than those of its gradient) are doomed to failure because of statistical error. In this paper, we show that this need not be true. We show that by explicitly smoothing the computed solution, and by suitably coupling the amount of smoothing to the number of particles used, the statistical error can be made arbitrarily small. These results are established by analysis of particle methods for the simple diffusion equation and are illustrated by numerical experiments. While other numerical methods are certainly more efficient at solving the simple diffusion equation, this setting allows us to perform the analysis that corrects a widespread misconception about a class of particle methods, and in so doing to provide a new tool which we believe will be important in the design of efficient particle methods for realistic problems.

Random particle methods in general are useful computational tools for solving equations in which diffusion is, by some measure, small and whose solutions therefore have steep gradients. Examples of such equations include convection-dominated convection-diffusion equations [6] and reaction-diffusion equations with rapid bistable kinetics [5]. The attraction of particle methods in these settings is that, unlike finite-difference or finite-element methods, they do not introduce numerical diffusion which can dominate the physical diffusion actually in the problem. Also, while the performance of spectral methods typically degrades as the diffusion coefficient becomes smaller [2], the performance of particle methods is uniformly good for all small diffusion coefficients (see [11] and Section 4 below).

In this paper, we consider a class of particle methods in the context of the initial value problem for a simple diffusion equation

$$v_t = \nu v_{xx} \tag{1}$$

$$v(x, 0) = f(x) \tag{2}$$

with diffusion coefficient $\nu \ll 1$. Of course, if this were the real problem, the diffusion coefficient could be scaled to one by adjusting the length and time scales. We study Eq. (1), however, as a model for a more difficult equation, such as a convection-dominated convection-diffusion equation, and we imagine that some other part of the problem fixes the length and time scales and that, with respect to those scales, ν is small.

Two major types of random particle methods for solving Eqs. (1)–(2) are function-transport methods, such as those discussed in this paper, and gradient-transport methods. These methods work as follows. A finite number of particles is introduced at time $t = 0$ at locations distributed within the support of the initial data f. During each of a succession of timesteps, each of these particles takes a random step from its current position. This simulates the diffusion. Associated with each particle is a point value of v (for a function-transport method) or a point value of v_x (for a gradient-transport method). These values are obtained from the initial data. At any time $t > 0$, the particle locations and the transported values are used to construct an approximation to $v(\cdot, t)$. For the function-transport methods, this construction is

0021-9991/93 $5.00

done directly, and the approximation is given by an expression of the form

$$u(x, t) = \sum_{p=1}^{M} M^{-1} f_p \phi(x - x_p), \tag{3}$$

where u is the approximation to v, M is the number of particles used, f_p is the function value associated with particle p, x_p is the location of the pth particle at time t, and ϕ is a specified function which will be discussed below. For the gradient-transport methods, an approximation of this form is constructed for $v_x(\cdot, t)$, and this is then integrated with respect to x to obtain an approximation to $v(\cdot, t)$.

Published work on random particle methods has focused almost exclusively on the use and analysis of gradient-transport methods (e.g., see [3, 7, 8, 13, 14] for applications and [10–12] for analysis of such methods). It is widely believed that gradient-transport methods are superior to function-transport methods. This belief is based on the following reasoning: On the one hand, for the gradient-transport method the spatial integration step smoothes the statistical fluctuations which result from the random-walk model of diffusion [3, 10]. On the other hand, for function-transport methods there is no similar averaging of the fluctuations and the following dilemma arises: If smoothing is not explicitly built into the function-transport method through the choice of the function ϕ, that is, if ϕ is chosen to be the delta function centered at the point x_p, then the approximations u do not converge to v as the number of particles is increased. In fact, the variance of the approximating function is infinite. If a smooth function ϕ is used, then the approximating functions u may converge as the number of particles increases, but the limiting function will not be the solution v to the diffusion equation. In either case, the function-transport method fails.

The purpose of this paper is to describe a class of function-transport methods in which the smoothing introduced through the function ϕ is *coupled* to the spacing between particles (roughly $1/M$) in such a way that convergence to the solution of the diffusion equations occurs. We refer to coupled smoothing which leads to convergence as "optimal smoothing."

In Section 2 we introduce the methods. In Section 3, we state and prove a convergence theorem for these methods. In Section 4, we present computational results illustrating our theoretical results.

2. THE NUMERICAL METHOD

We consider the initial value problem Eqs. (1)–(2) and assume that the initial data f has support in the interval $[0, 1]$. We introduce M particles at locations x_p^0 in this interval. We assume that these locations are chosen so that $x_p^0 < x_{p+1}^0$ for all p and we define $h_p = \frac{1}{2}(x_{p+1}^0 - x_{p-1}^0)$ (we take $x_0^0 = 0$ and $x_{M+1}^0 = 1$). With particle p we associate the value $h_p f_p \equiv h_p f(x_p^0)$. We discretize time into intervals of length k. During each timestep, each particle takes a step η chosen from a Gaussian distribution with mean zero and variance $2\nu k$. The random steps are chosen independently for each particle. The random steps that a given particle takes in different timesteps are also chosen independently of one another. At time $t = nk$, we have particles at locations

$$x_p^n = x_p^0 + \eta_p^1 + \eta_p^2 + \cdots + \eta_p^n \tag{4}$$

for $p = 1, \ldots, M$, where $\eta_p^1, \eta_p^2, \ldots, \eta_p^n$ are independent Gaussian random variables with mean zero and variance $2\nu k$. Alternatively, we can write

$$x_p^n = x_p^0 + \eta_p, \tag{5}$$

where

$$\eta_p = \eta_p^1 + \eta_p^2 + \cdots + \eta_p^n \tag{6}$$

is a Gaussian random variable with mean zero and variance $2\nu nk = 2\nu t$. Define the function

$$\phi_\varepsilon(x) = \frac{1}{\sqrt{\pi}} \frac{1}{\varepsilon} e^{-x^2/\varepsilon^2}. \tag{7}$$

Note that ϕ_ε is an approximate δ-function and that $\int_{-\infty}^{\infty} \phi_\varepsilon(x)\, dx = 1$ for all $\varepsilon > 0$. Using the function ϕ_ε, we define an approximate solution to the diffusion equation at time $t = nk$ by the relation

$$u_M^\varepsilon(x, nk) = \sum_{p=1}^{M} h_p f_p \phi_\varepsilon(x - x_p^n). \tag{8}$$

For fixed values of ε and M, this completely specifies the method except for a prescription for choosing the initial particle locations. We will see from the proof presented in the next section that it suffices that these locations be chosen in $[0, 1]$ in such a way that $h_p \leqslant BM^{-1}$ as $M \to \infty$ for some constant $B > 0$, and henceforth, we make the assumption that such an inequality holds.

3. CONVERGENCE OF THE METHOD

In this section we investigate the convergence of the function u_M^ε, defined by Eq. (8), to the solution v of Eqs. (1)–(2). The approximate solution u_M^ε is, through its dependence on x_p^n, $p = 1, \ldots, M$, a function of Mn independent identically distributed random variables η_p^j, $p = 1, \ldots, M$, $j = 1, \ldots, n$. Let $E(w)$ denote the expected value of a random variable w with respect to the Mn random variables η_p^j. We will show that

for any fixed $t = nk > 0$, $E(|u^\varepsilon_M(x, t) - v(x, t)|^2)$ vanishes uniformly in x on any compact interval as $\varepsilon \to 0$ and $M \to \infty$, provided ε and M are coupled in such a way that $(M\varepsilon)^{-1} \to 0$. In particular, this is the case if $\varepsilon = CM^{-q}$ for some $0 < q < 1$ and a constant $C > 0$.

The following easily derived identity is central to our analysis:

$$E(|u^\varepsilon_M(x, t) - v(x, t)|^2) = \mathrm{Var}(u^\varepsilon_M(x, t)) + \{E(u^\varepsilon_M(x, t)) - v(x, t)\}^2. \quad (9)$$

Here, $\mathrm{Var}(u^\varepsilon_M(x, t))$ denotes the variance of $u^\varepsilon_M(x, t)$. We will derive expressions for $E(u^\varepsilon_M(x, t))$ and $\mathrm{Var}(u^\varepsilon_M(x, t))$ and we will use these to show that both terms on the right-hand side of Eq. (9) can be made to vanish.

From the definition of $u^\varepsilon_M(x, t)$ in Eq. (8) and the independence of the random variables on which it depends, we find that

$$E(u^\varepsilon_M(x, t)) = \sum_{p=1}^{M} h_p f_p E(\phi_\varepsilon(x - x^n_p)) \quad (10)$$

and

$$\mathrm{Var}(u^\varepsilon_M(x, t)) = \mathrm{Var}\left(\sum_{p=1}^{M} h_p f(x^0_p)\, \phi_\varepsilon(x - x^n_p)\right) = \sum_{p=1}^{M} h_p^2 f(x^0_p)^2 \,\mathrm{Var}(\phi_\varepsilon(x - x^n_p)). \quad (11)$$

In order to estimate these expressions, we calculate the expected value and variance of the function $\phi_\varepsilon(x, t)$. The calculation of $E(\phi_\varepsilon(x - x^n_p))$ is straightforward. Recall that $x^n_p = x^0_p + \eta_p$, where x^0_p is the initial position of particle p and η_p is a Gaussian random variable with mean zero and variance $2vt$. For the calculation of the expected value, we set $x' = x - x^0_p$ and suppress the subscript p on η. Then

$$E(\phi_\varepsilon(x - x^n_p)) = E(\phi_\varepsilon(x' - \eta)) = \int_{-\infty}^{\infty} \phi_\varepsilon(x' - \eta) \frac{1}{\sqrt{4\pi vt}} e^{-\eta^2/(4vt)}\, d\eta = \frac{1}{\pi\sqrt{4vt}} \int_{-\infty}^{\infty} e^{-(x'-\eta)^2/\varepsilon^2} e^{-\eta^2/(4vt)}\, d\eta. \quad (12)$$

Upon combining the exponents and completing the square, we find that

$$E(\phi_\varepsilon(x - x^n_p)) = \frac{1}{\sqrt{4\pi vt}} e^{-(x')^2/(4vt + \varepsilon^2)} \frac{1}{\sqrt{1 + \varepsilon^2/4vt}}. \quad (13)$$

Define the function

$$G_\varepsilon(x, t) = \frac{1}{\sqrt{4\pi vt}} e^{-x^2/(4vt + \varepsilon^2)} \frac{1}{\sqrt{1 + \varepsilon^2/4vt}}. \quad (14)$$

We have just shown:

LEMMA 1.

$$E(\phi_\varepsilon(x - x^n_p)) = G_\varepsilon(x - x^0_p, nk). \quad (15)$$

Note that $G_\varepsilon(x, t)$ is related to the Green's function for the diffusion equation:

$$G(x, t) = \frac{1}{\sqrt{4\pi vt}} e^{-x^2/(4vt)}. \quad (16)$$

In fact, for any x and $t > 0$,

$$\lim_{\varepsilon \to 0} G_\varepsilon(x, t) = G(x, t). \quad (17)$$

In a calculation similar to that done to obtain Eq. (15), we find that

$$E(\phi_\varepsilon(x - x^n_p)^2) = \varepsilon^{-1} \left\{\frac{1}{\pi} \frac{1}{\sqrt{8vt + \varepsilon^2}} e^{-2(x - x^0_p)^2/(8vt + \varepsilon^2)}\right\}. \quad (18)$$

Using this, we calculate the variance of $\phi_\varepsilon(x - x^n_p)$ to obtain:

LEMMA 2.

$$\mathrm{Var}(\phi_\varepsilon(x - x^n_p)) = \varepsilon^{-1} \left\{\frac{1}{\pi} \frac{1}{\sqrt{8vt + \varepsilon^2}} e^{-2(x - x^0_p)^2/(8vt + \varepsilon^2)}\right\} - \left\{\frac{1}{\pi} \frac{1}{4vt + \varepsilon^2} e^{-2(x - x^0_p)^2/(4vt + \varepsilon^2)}\right\}. \quad (19)$$

Note that $\mathrm{Var}(\phi_\varepsilon(x - x^n_p)) \to \infty$ as $\varepsilon \to 0$. This is the reason that function-transport methods for which $\phi(x) = \delta(x)$ in Eq. (3) have infinite variance. We will see that by coupling the approach of M to ∞ with that of ε to 0 in the way mentioned before, we can avoid that problem with our method.

Returning to our consideration of $E(u^\varepsilon_M(x, t))$, we see from Eq. (10) and Lemma 1 that

$$E(u^\varepsilon_M(x, t)) = \sum_{p=1}^{M} h_p f(x^0_p)\, G_\varepsilon(x - x^0_p, t). \quad (20)$$

We assume that the initial data f is bounded and piecewise continuously differentiable. Then, the right-hand

side of Eq. (20) is a midpoint-rule approximation to $\int_0^1 f(x')\, G_\varepsilon(x-x', t)\, dx'$, and, since $h_p \leqslant BM^{-1}$, it converges to this integral as $M \to \infty$ for any value of $\varepsilon > 0$. Hence,

$$\lim_{M\to\infty} E(u_M^\varepsilon(x, t)) = \int_0^1 f(x')\, G_\varepsilon(x-x', t)\, dx'. \quad (21)$$

From this equation, and the fact that $G_\varepsilon(x-x', t) \to G(x-x', t)$ uniformly on $[0, 1]$ for any $t > 0$, we find also that

LEMMA 3.

$$\lim_{\varepsilon\to 0}\lim_{M\to\infty} E(u_M^\varepsilon(x, t)) = \int_0^1 f(x')\, G(x-x', t)\, dx'. = v(x, t). \quad (22)$$

We have used here our assumption that the support of f is contained in $[0, 1]$. Equation (22) tells us that as the two numerical parameters M and ε go to their respective limits, the expected value of our approximate solution $u_M^\varepsilon(x, t)$ converges to the solution of the initial value problem. We note that there is no restriction imposed here on the order in which the two limits are taken; for $t > 0$, we could equally well have let $\varepsilon \to 0$ first, and then $M \to \infty$.

As we discuss below, the other component of the expected squared error, i.e., $\mathrm{Var}(u_M^\varepsilon(x, t))$, vanishes only with suitable coupling between the limiting processes $M \to \infty$ and $\varepsilon \to 0$. It suffices to take $\varepsilon = CM^{-q}$ for constants $C > 0$ and $0 < q < 1$. We now estimate the *rate* of convergence of $E(u_M^\varepsilon(x, t))$ to $v(x, t)$ under the assumption that $\varepsilon = CM^{-q}$. We write

$$\{E(u_M^\varepsilon(x, t)) - v(x, t)\}^2 = \left\{ \sum_{p=1}^{M} h_p f(x_p^0)\, G_\varepsilon(x-x_p^0, t) - \int_0^1 G_\varepsilon(x-x', t) f(x')\, dx' + \int_0^1 (G_\varepsilon(x-x', t) - G(x-x', t)) f(x')\, dx' \right\}^2. \quad (23)$$

The first pair of terms is the error in the midpoint-rule approximation to the integral of $f(x')\, G_\varepsilon(x-x', t)$ over $[0, 1]$. $G_\varepsilon(x-x', t)$ is smooth and bounded uniformly in ε and $(x-x')$ for any fixed $t > 0$. Hence, provided f is bounded and piecewise continuously differentiable, this error is $O(M^{-1})$. (If $f \in C^2([0, 1])$, then this error is $O(M^{-2})$.) The difference $G_\varepsilon(x-x', t) - G(x-x', t)$ which appears in the second pair of terms in Eq. (23) can be expressed

$$G_\varepsilon(x-x', t) - G(x-x', t) = \frac{1}{\sqrt{\pi(4vt+\varepsilon^2)}} e^{-(x-x')^2/(4vt+\varepsilon^2)} \times \left\{ 1 - \sqrt{1+\varepsilon^2/(4vt)}\, e^{-\varepsilon^2(x-x')^2/[(4vt)(4vt+\varepsilon^2)]} \right\}. \quad (24)$$

For fixed $t > 0$, this difference vanishes exponentially fast for large $|x-x'|$ and for all $\varepsilon > 0$. For $|x-x'|$ bounded, the difference is $O(\varepsilon^2)$. Thus, $|\int_0^1 (G_\varepsilon(x-x', t) - G(x-x', t)) f(x')\, dx'| = O(\varepsilon^2) = O(M^{-2q})$ uniformly in x on any bounded interval. (In fact, this term is $O(\varepsilon^\alpha)$ uniformly for *all* x for any $\alpha < 2$.) From this and our estimate of the first pair of terms in Eq. (23), we conclude that

$$\{E(u_M^\varepsilon(x, t)) - v(x, t)\}^2 = \{O(M^{-1}) + O(M^{-2q})\}^2. \quad (25)$$

Therefore, we have established

LEMMA 4.

$$\{E(u_M^\varepsilon(x, t)) - v(x, t)\}^2 = \begin{cases} O(M^{-2}) & \text{if } q > \frac{1}{2} \\ O(M^{-4q}) & \text{if } q < \frac{1}{2}. \end{cases} \quad (26)$$

We note that if $f \in C^2([0, 1])$, then the $O(M^{-1})$ term in Eq. (25) can be replaced by $O(M^{-2})$, and in Lemma 4 we would have $\{E(u_M^\varepsilon(x, t)) - v(x, t)\}^2 = O(M^{-4q})$ for $0 < q < 1$.

Now we turn to estimating $\mathrm{Var}(u_M^\varepsilon(x, t))$. Rearranging the expression for $\mathrm{Var}(\phi_\varepsilon(x-x_p^n))$ in Lemma 2, we find that

$$\mathrm{Var}(\phi_\varepsilon(x-x_p^n)) = \varepsilon^{-1}\psi(x-x_p^0, vt; \varepsilon), \quad (27)$$

where

$$\psi(x, vt; \varepsilon) \equiv \frac{1}{\pi}\frac{1}{\sqrt{8vt+\varepsilon^2}} e^{-2x^2/(8vt+\varepsilon^2)} \times \left\{ 1 - \varepsilon\frac{\sqrt{8vt+\varepsilon^2}}{4vt+\varepsilon^2} e^{-8vtx^2/[(4vt+\varepsilon^2)(8vt+\varepsilon^2)]} \right\}. \quad (28)$$

Note that for any fixed $vt > 0$ and $\varepsilon > 0$, $\psi(x, vt; \varepsilon)$ has its maximum at $x = 0$. Thus

$$\psi(x, vt; \varepsilon) \leqslant \psi(0, vt; \varepsilon) = \frac{1}{\pi}\left\{ \frac{1}{\sqrt{8vt+\varepsilon^2}} - \frac{\varepsilon}{4vt+\varepsilon^2} \right\}. \quad (29)$$

For any fixed $vt > 0$, this in turn is bounded uniformly in ε by

$$D(vt) \equiv \frac{1}{\pi} \frac{1}{\sqrt{8vt}}. \tag{30}$$

Hence, we can write

$$\mathrm{Var}(\phi_\varepsilon(x - x_p^n)) \leqslant D\varepsilon^{-1}, \tag{31}$$

where D depends only on vt. Combining this with Eq. (11), we find that

$$\mathrm{Var}(u_M^\varepsilon(x, t)) \leqslant D\varepsilon^{-1} \sum_{p=1}^{M} h_p^2 f(x_p^0)^2. \tag{32}$$

Since we assume that the initial locations of the particles are distributed in $[0, 1]$ in a roughly uniform manner so that $h_p \leqslant BM^{-1}$, we can replace one of the factors h_p in Eq. (32) by BM^{-1} and we find that for bounded and piecewise continuously differentiable initial data f,

$$\mathrm{Var}(u_M^\varepsilon(x, t)) \leqslant \frac{1}{M\varepsilon} D \|f\|_2^2 + O\left(\frac{1}{M^2\varepsilon}\right), \tag{33}$$

where $D = D(vt)$ and $\|\cdot\|_2$ is the L_2 norm. We therefore conclude:

LEMMA 5. *If $\varepsilon = CM^{-q}$ for some constant $C > 0$ and $0 < q < 1$, then as $M \to \infty$,*

$$\mathrm{Var}(u_M^\varepsilon(x, t)) \leqslant \frac{1}{M^{1-q}} D \|f\|_2^2 + O\left(\frac{1}{M^{2-q}}\right). \tag{34}$$

Thus, the variance of our approximation vanishes in the limit $M \to \infty$, provided that the amount of smoothing in the function ϕ_ε is appropriately coupled to the initial spacing between particles.

Putting together the results of Lemmas 4 and 5, we have the following convergence theorem:

THEOREM. *Suppose the initial data f for Eqs. (1)–(2) is bounded and piecewise continuously differentiable. If $\varepsilon = CM^{-q}$ for some constants $C > 0$ and $0 < q < 1$, then for any $t > 0$, there exists positive constants C_1 and C_2 such that*

$$E(|u_M^\varepsilon(x, t) - v(x, t)|^2 \leqslant C_1 \frac{1}{M^{1-q}} + C_2 \max\left(\frac{1}{M^2}, \frac{1}{M^{4q}}\right) \tag{35}$$

for M sufficiently large. The coefficients C_1 and C_2 depend on vt, f, and x. The estimate is uniform in x on any compact interval.

We note that the timestep k plays no role in the theorem; all that matters is the time $t = nk$. This follows from Eqs. (4)–(5). The random variable x_p^n has the same distribution whether it is obtained by taking n independent Gaussian steps each with mean zero and variance $2vk$ as in Eq. (4), or one Gaussian step of mean zero and variance $2vt$ as in Eq. (5).

4. COMPUTATIONAL EXAMPLES

In this section, we present results of numerical experiments which illustrate and extend the discussion of the previous section. The function transport method was applied to the initial value problem (1)–(2) with diffusion coefficient $v = 0.001$ for two different sets of initial data:

Problem 1.

$$f(x) = \begin{cases} N_\sigma & \text{for } -5\sigma < x < 5\sigma \\ 0 & \text{otherwise,} \end{cases} \tag{36}$$

where

$$N_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$$

and $\sigma = 0.3$.

Problem 2.

$$f(x) = \begin{cases} 1 & \text{for } -\frac{1}{2} < x < \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{37}$$

For each of these problems, a closed-form analytic solution is available with which to compare the numerical solution.

In Figs. 1 and 2, we illustrate the shortcomings of the "decoupled" function transport method in which either M or ε is fixed while the other parameter goes to its natural limit. Figure 1 shows the result of applying the function-transport method to Problem 1 with the number of particles *M fixed* at 2000 and with ε made progressively smaller. In each frame, we show, for $t = 10.0$, the exact solution (dashed curve), and one or more numerical solutions computed from Eq. (8) (solid curves). The values of ε used to produce these results were $\varepsilon = 0.2$, 0.1, 0.05, and 0.01 for curves (A), (B), (C), and (D), respectively. The sharp increase in the variation of the numerical solution is evident. In Fig. 2, we depict the results of *fixing* $\varepsilon = 0.6$ and allowing M to increase. We show, at $t = 10.0$, the exact solution (dashed curve) and numerical solutions obtained with $M = 2000$ (solid curve (A)) and $M = 32000$ (solid curve (B). There is little difference in the numerical solution obtained with these values M, and even less difference between the numeri-

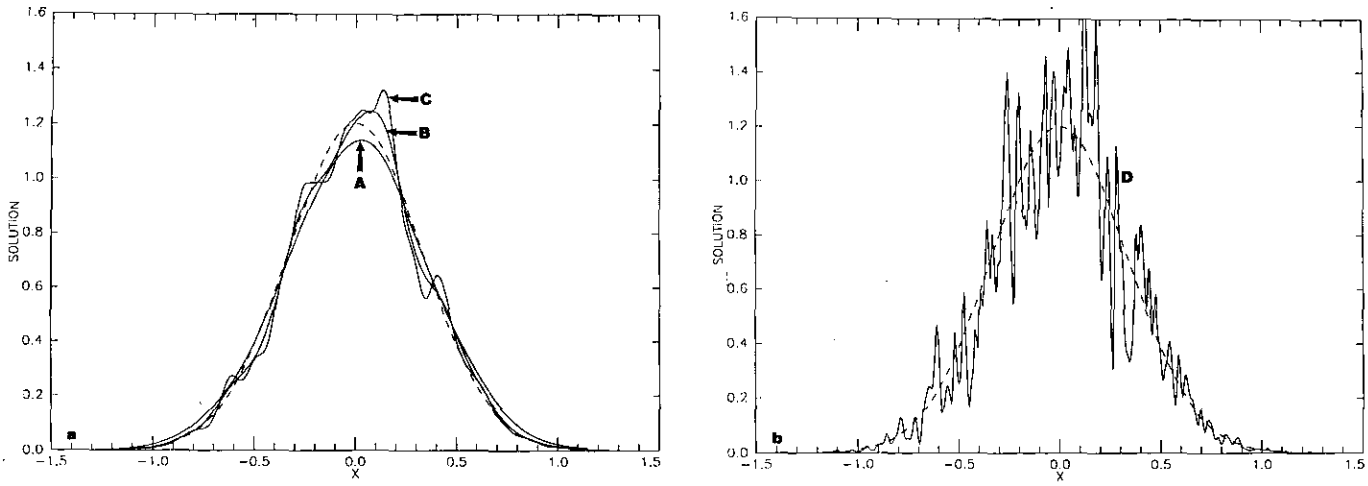

FIG. 1. Behavior of the function-transport method for Problem 1 with M fixed at 2000. Exact solution $v(x, t)$ (dashed curve) and numerical solution $u^{\varepsilon}_{M}(x, t)$ (solid curves) at $t = 10.0$ are shown for (A) $\varepsilon = 0.2$, (B) $\varepsilon = 0.1$, (C) $\varepsilon = 0.05$, and (D) $\varepsilon = 0.01$.

cal solution obtained with $M = 32000$ and $M = 64000$ (not shown), so it is fair to say that the numerical solution shown in curve (B) has converged. However, it has clearly converged to something far from the true solution.

Figures 3 and 4 depict results, again at $t = 10.0$, for the case in which M and ε are coupled through the relation $\varepsilon = CM^{-q}$ with $C = 1$ and $q = 0.2$. (This value of q balances the rates at which the two terms in the bound (35) approach zero.) In Fig. 3, we show numerical solutions (solid curves) obtained from Eq. (8) for $M = 2000$, 8000, 32000, and 128000, and for a particular realization of the particle random walks. Also shown is the exact solution (dashed curve). In Fig. 4, we compare the average behavior of the method with $\varepsilon = M^{-0.2}$ to the true solution. The solid curves show $E(u^{\varepsilon}_{M}(x, t)) \pm \{\mathrm{Var}(u^{\varepsilon}_{M}(x, t))\}^{1/2}$ for the above values of M, and the dashed curve shows the exact solution. Both the approach of the expected value to the true solution, and the decrease in the variance are evident.

It is not surprising, in view of the bound (35), that the rates at which $E(u^{\varepsilon}_{M}(x, t)) \to v(x, t)$ and $\mathrm{Var}(u^{\varepsilon}_{M}(x, t)) \to 0$ depend on the choice of C and q in the relation $\varepsilon = CM^{-q}$. In Fig. 5, we illustrate this by showing the average behavior of the method with $C = 1$, as in Fig. 4, but with $q = 0.4$. In this case $u^{\varepsilon}_{M}(x, t)$ is much closer to $v(x, t)$ for the same value of M, but $\mathrm{Var}(u^{\varepsilon}_{M}(x, t))$ is larger. The dashed curve again shows the exact solution $v(x, t)$. Each of the three pairs of solid curves shows $E(u^{\varepsilon}_{M}(x, t)) \pm \{\mathrm{Var}(u^{\varepsilon}_{M}(x, t))\}^{1/2}$ for a

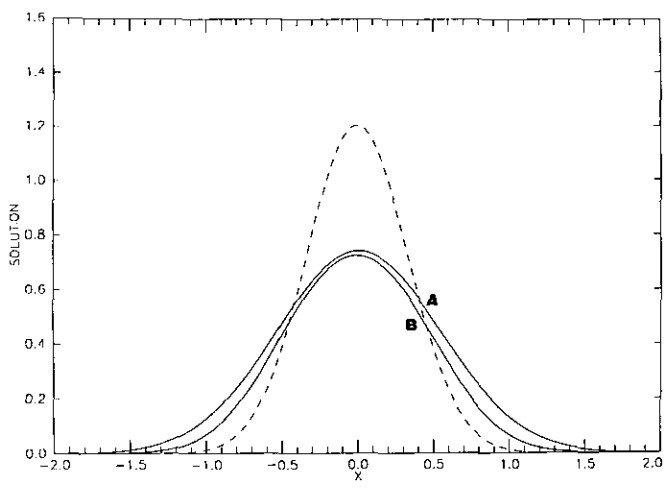

FIG. 2. Behavior of the function-transport method for Problem 1 with ε fixed at 0.6. Exact solution $v(x, t)$ (dashed curve) and numerical solution $u^{\varepsilon}_{M}(x, t)$ (solid curves) at $t = 10.0$ are shown for (A) $M = 2000$ and (B) $M = 32000$.

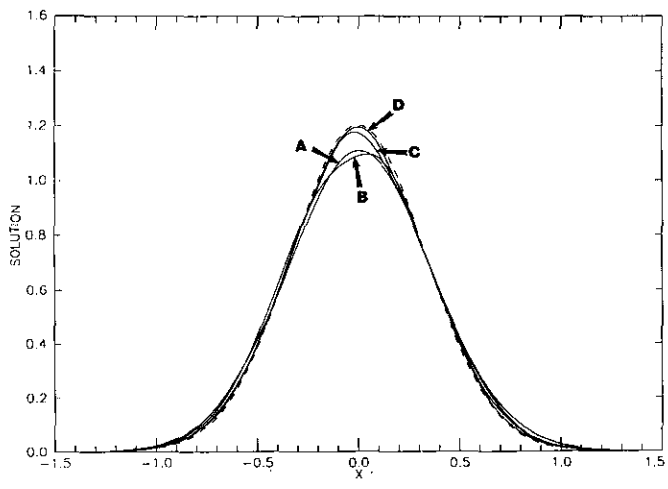

FIG. 3. Behavior of the function-transport method for Problem 1 with $\varepsilon = CM^{-q}$, with $C = 1$ and $q = 0.2$. Exact solution $v(x, t)$ (dashed curve) and numerical solution $u^{\varepsilon}_{M}(x, t)$ (solid curves) are shown at $t = 10.0$ for (A) $M = 2000$, $\varepsilon = 0.2187$; (B) $M = 8000$, $\varepsilon = 0.1657$; (C) $M = 32000$, $\varepsilon = 0.1256$; and (D) $M = 128000$, $\varepsilon = 0.0952$.

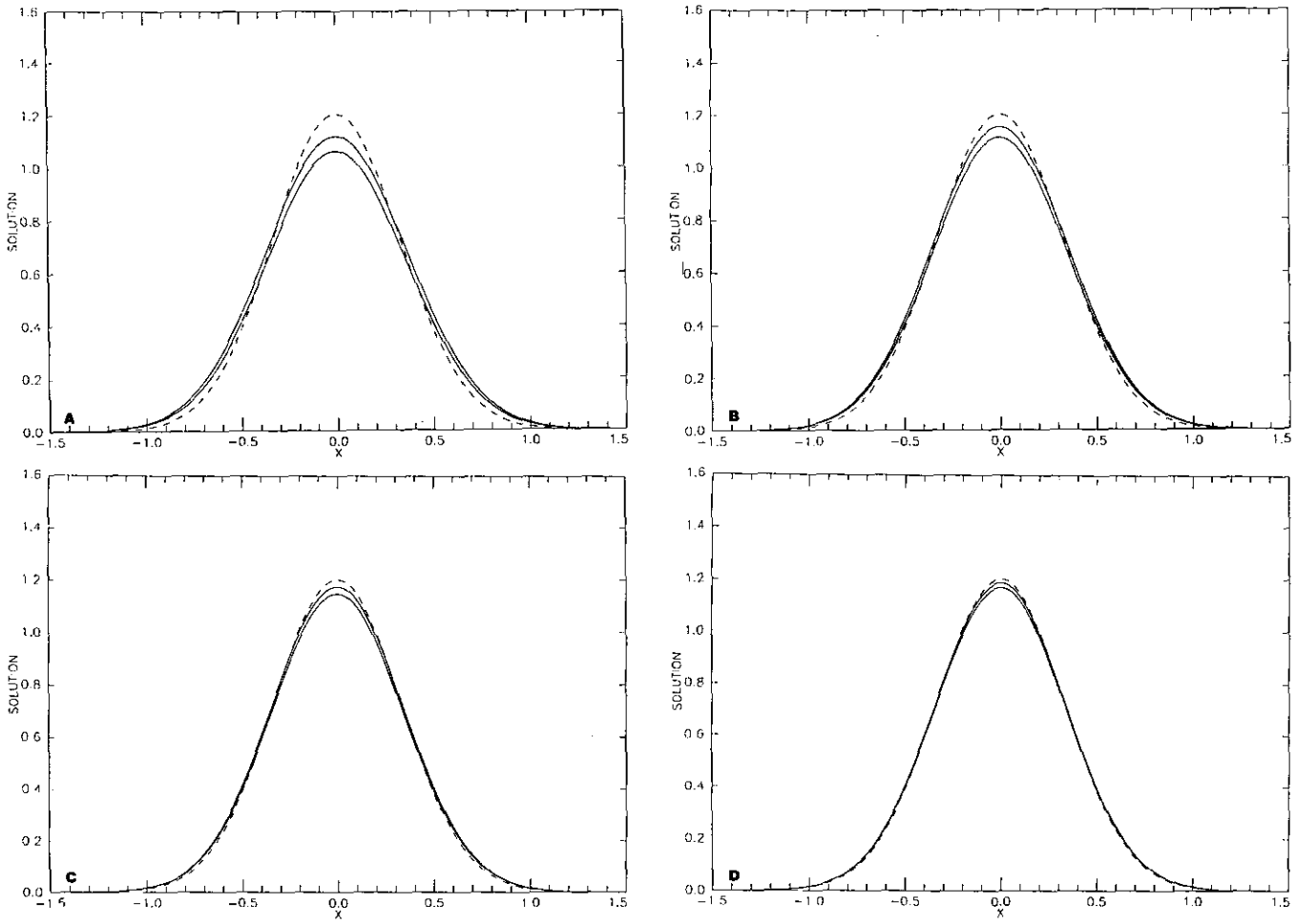

FIG. 4. Average behavior of the function-transport method for Problem 1 with $\varepsilon = CM^{-q}$, with $C = 1$ and $q = 0.2$. Exact solution $v(x, t)$ (dashed curve) and $E(u^{\varepsilon}_M(x, t)) \pm \{\mathrm{Var}(u^{\varepsilon}_M(x, t))\}^{1/2}$ (solid curves) are shown at $t = 10.0$ for (A) $M = 2000$, $\varepsilon = 0.2187$; (B) $M = 8000$, $\varepsilon = 0.1657$; (C) $M = 32000$, $\varepsilon = 0.1256$; and (D) $M = 128000$, $\varepsilon = 0.0952$.

given value of $M = 2000$, 8000, or 32000. We note that for $M \geqslant 2000$ the curves for $v(x, t)$ and $E(u^{\varepsilon}_M(x, t))$ (not shown) are indistinguishable. Because of the larger variance for a given value of M, numerical solutions obtained with $q = 0.4$ are significantly more oscillatory than those obtained with $q = 0.2$. The effect of using different values of q is illustrated also in Fig. 6 for Problem 2. In this figure, we show the exact solution (dashed curve) and the numerical solution (solid curve) obtained using $M = 32000$ at $t = 0.1$. For curve (A), $C = 1$ and $q = 0.2$, while for curve (B), $C = 1$ and $q = 0.4$. As for Problem 1, the expected value is closer to the true solution and the variance is larger with $q = 0.4$ than with $q = 0.2$. For this problem, with its very steep gradients, the solution obtained with $q = 0.4$ is preferable.

We next discuss the performance of the method as $vt \to 0$. Consider formulas (24) and (28) which influence the bounds on $|E(u^{\varepsilon}_M(x, t)) - v(x, t)|$ and $\mathrm{Var}(u^{\varepsilon}_M(x, t))$, respectively. From these formulas, we see that the coefficients in these bounds (e.g., D given by (30)) are very large when vt and ε are simultaneously small. One might expect, therefore,

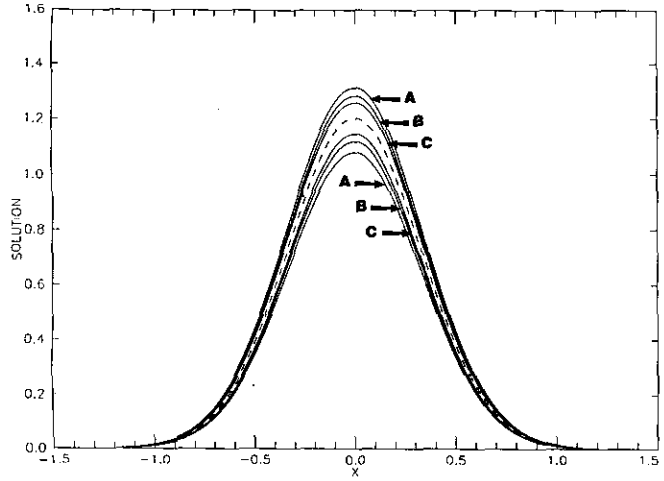

FIG. 5. Average behavior of the function-transport method for Problem 1 with $\varepsilon = CM^{-q}$ with $C = 1$ and $q = 0.4$. Exact solution $v(x, t)$ (dashed curve) and $E(u^{\varepsilon}_M(x, t)) \pm \{\mathrm{Var}(u^{\varepsilon}_M(x, t))\}^{1/2}$ (solid curves) are shown at $t = 10.0$ for (A) $M = 2000$, $\varepsilon = 0.0478$; (B) $M = 8000$, $\varepsilon = 0.0275$; and (C) $M = 32000$, $\varepsilon = 0.0158$.

that increasingly large values of M might be needed to obtain reasonable results from the method as $vt \to 0$. That this is *not* the case is illustrated in Fig. 7 which shows data for times $t = 0.01$, 0.1, 1.0, 10.0 from the application of the method to Problem 1 with $v = 0.001$. In each frame are four curves, one corresponding to each of these times. Since only the combination vt, and not v or t separately, matters in both the exact and numerical solutions, we can interpret these curves as corresponding to a fixed time $t = 1.0$ and to values of v ranging from 10^{-2} down to 10^{-5}. In Figs. 7a and 7c, we plot $\max_x \{(E(u^\varepsilon_M(x, t)) - v(x, t))^2\}$ as a function of M for each of these values of v. Similarly, in Figs. 7b and 7d, we plot $\max_x \{\mathrm{Var}(u^\varepsilon_M(x, t))\}$ as a function of M for each value of v. The data used for Figs. 7a–b were obtained with $q = 0.2$ and $C = 1$, those for Figs. 7c–d are based on $q = 0.4$ and $C = 1$, and M ranged from 500 to 256000 in both cases. From Figs. 7b and d, we see that the variance decreases as v decreases. From Figs. 7a and c, we see that $\max_x \{(E(u^\varepsilon_M(x, t)) - v(x, t))^2\}$ increases when v decreases

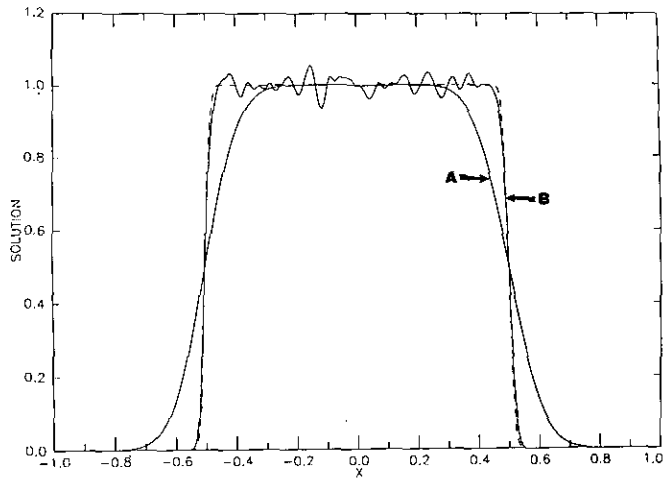

FIG. 6. Behavior of the function-transport method for Problem 2 with $\varepsilon = CM^{-q}$, with $C = 1$ and $M = 32000$. Exact solution $v(x, t)$ (dashed curve) and numerical solution $u^\varepsilon_M(x, t)$ (solid curves) are shown at $t = 0.1$ for (A) $q = 0.2$ ($\varepsilon = 0.1256$) and (B) $q = 0.4$ ($\varepsilon = 0.0158$).

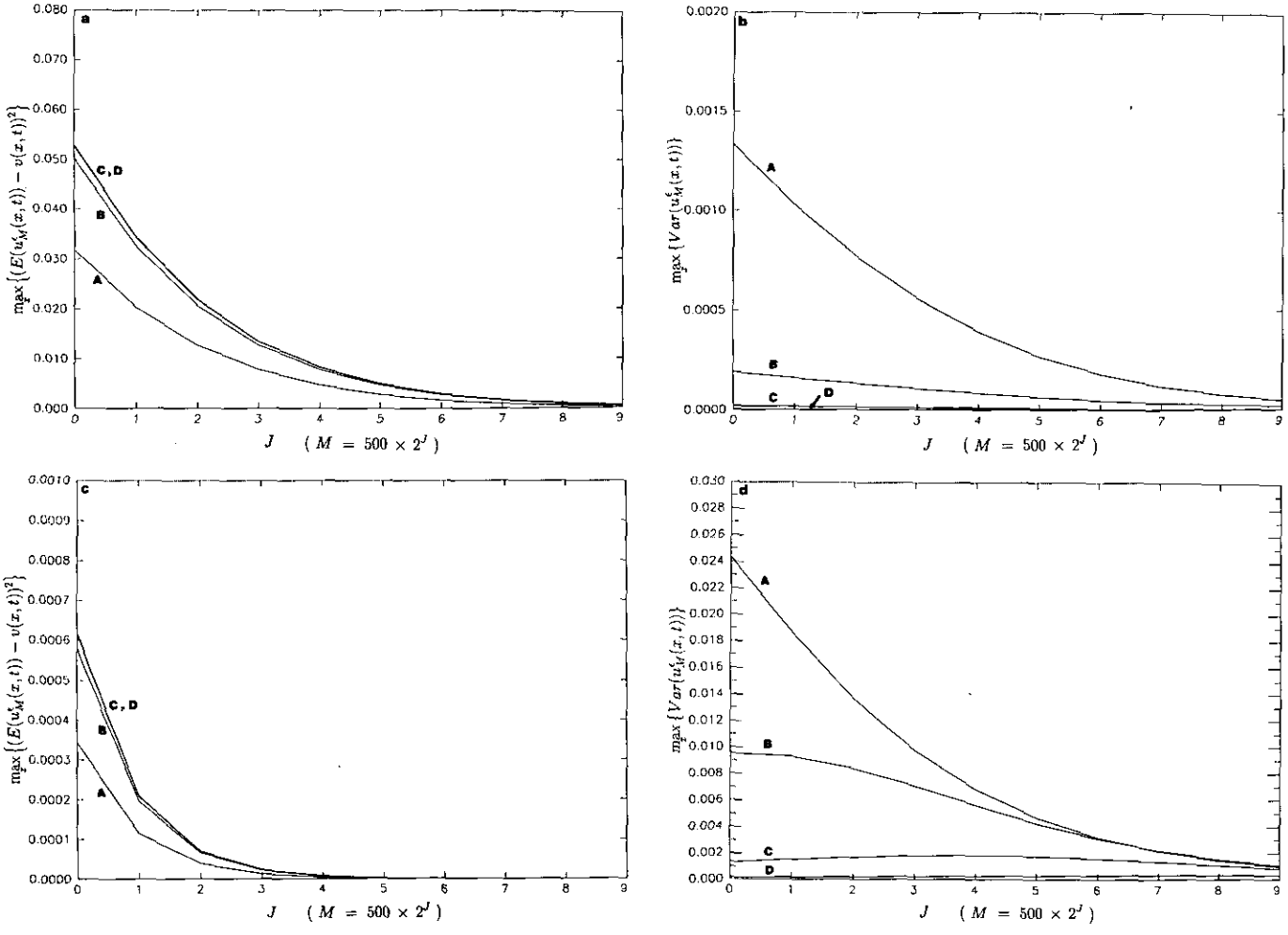

FIG. 7. Performance of function-transport method for Problem 1 as diffusivity $v \to 0$. The curves in each frame correspond to $t = 1.0$ and (A) $v = 10^{-2}$, (B) $v = 10^{-3}$, (C) $v = 10^{-4}$; (D) $v = 10^{-5}$. We plot (a) $\max_x \{(E(u^\varepsilon_M(x, t)) - v(x, t))^2\}$ for $q = 0.2$, $C = 1$; (b) $\max_x \{\mathrm{Var}(u^\varepsilon_M(x, t))\}$ for $q = 0.2$, $C = 1$; (c) $\max_x \{(E(u^\varepsilon_M(x, t)) - v(x, t))^2\}$ for $q = 0.4$, $C = 1$; and (d) $\max_x \{\mathrm{Var}(u^\varepsilon_M(x, t))\}$ for $q = 0.4$, $C = 1$, as functions of M for $M = 500$, 1000, 2000, ..., 256000.

from 10^{-2} to 10^{-4}, but that the curves for $\nu = 10^{-4}$ and $\nu = 10^{-5}$ are virtually indistinguishable. The quantity $\max_x \{(E(u_M^\varepsilon(x, t)) - v(x, t))^2\}$ changes most with ν for $500 \leqslant M \leqslant 1000$ and shows much less sensitivity to ν for values of $M \geqslant 4000$. Thus, the computational effort needed to obtain a given level of accuracy is roughly the same for all values of ν as $\nu \to 0$.

5. CONCLUSION

We have presented a new function-transport particle method for diffusion problems with small diffusivity ν. We have resolved the dilemma of how much smoothing to include in this type of method by proving that the method converges as the number of particles M increases, and the degree of smoothing, as measured by ε, decreases, provided that these limits are coupled in such a way that $(M\varepsilon)^{-1} \to 0$. We have illustrated the performance of the method for the special coupling $\varepsilon = CM^{-q}$, $C > 0$, $0 < q < 1$, and showed that the method performs uniformly well for all ν small.

We have not made a systematic comparison of the method with gradient-transport methods. Preliminary results suggest the gradient method is superior for problems with very steep transitions (such as Problem 2 above), but that the function-transport method works better for problems that are somewhat smoother (such as Problem 1). The former is not surprising because the gradient method concentrates its effort at the steep gradient. What is surprising is that function-transport methods, which according to conventional wisdom do not work at all, can outperform the gradient mthod on some problems.

We are exploring the effect of incorporating the explicit smoothing of the current method into gradient-transport methods. Our hope is that the explicit smoothing will increase the rate of convergence of gradient-transport methods. In the combined method, a formula like Eq. (8) would be used to approximate $v_x(x, t)$, and an approximation to $v(x, t)$ would then be obtained by integration. Since the indefinite integral of our smoothing function ϕ_ε is expressible in terms of the complimentary error function, this is straightforward to implement. If one-dimensional studies of the smoothed gradient-transport method are encouraging, we will explore the use of similar smoothing in the multi-dimensional methods for convection-diffusion problems introduced by Anderson [1] and Fogelson [7].

ACKNOWLEDGMENTS

The authors are grateful to Victor Moll and Robert Palais for helpful discussions and to Alexandre Chorin for encouragement. The work of Aaron Fogelson was supported in part by NSF Grants DMS88-03482 and DMS91-04410 and by an Alfred P. Sloan Foundation Fellowship. The work of Robert Dillon was supported in part by NIH Grant GM 29123. Computations were performed at the Center for Scientific Computation of the University of Utah.

REFERENCES

1. C. R. Anderson, *J. Comput. Phys.* **61**, 417 (1985).
2. N. Bressan and A. Quarteroni, *SIAM J. Numer. Anal.* **23**, 1138 (1986).
3. A. J. Chorin, "Numerical Methods for Use in Combustion Modeling," in *Computing Methods in Applied Sciences and Engineering*, edited by R. Glowinski and J. L. Lions (North-Holland, Amsterdam, 1980), p. 229.
4. G. Dahlquist and A. Bjorck, *Numerical Methods* (Prentice-Hall, Englewood Cliffs, NJ, 1974), p. 453.
5. P. C. Fife, *Mathematical Aspects of Reacting and Diffusing Systems* (Springer-Verlag, New York, 1979), p. 84.
6. A. L. Fogelson, *J. Comput. Phys.* **56**, 111 (1984).
7. A. L. Fogelson, *J. Comput. Phys.* **100**, 1 (1992).
8. A. F. Ghoniem and F. S. Sherman, *J. Comput. Phys.* **61**, 1 (1985).
9. L. Greengard and V. Rohklin, *J. Comput. Phys.* **73**, 325 (1987).
10. O. H. Hald, *SIAM J. Sci. Stat. Comput.* **2**, 85 (1981).
11. E. G. Puckett, *Math. Comput.* **52**, 615 (1989).
12. S. Roberts, *Math. Comput.* **52**, 647 (1989).
13. A. S. Sherman and C. S. Peskin, *SIAM J. Sci. Stat. Comput.* **7**, 1360 (1986).
14. A. S. Sherman and C. S. Peskin, *SIAM J. Sci. Stat. Comput.* **9**, 170 (1988).