

**$L^1$ –FRAMEWORK FOR CONTINUOUS DEPENDENCE  
AND ERROR ESTIMATES FOR QUASILINEAR ANISOTROPIC  
DEGENERATE PARABOLIC EQUATIONS**

GUI-QIANG CHEN AND KENNETH H. KARLSEN

ABSTRACT. We develop a general  $L^1$ –framework for deriving continuous dependence and error estimates for quasilinear anisotropic degenerate parabolic equations with the aid of the Chen-Perthame kinetic approach. We apply our  $L^1$ –framework to establish an explicit estimate for continuous dependence on the nonlinearities and an optimal error estimate for the vanishing anisotropic viscosity method, without imposition of bounded variation of the approximate solutions. Finally, as an example of a direct application of this framework to numerical methods, we focus on a linear convection-diffusion model equation and derive an  $L^1$  error estimate for an upwind-central finite difference scheme.

1. INTRODUCTION

We are concerned with the Cauchy problem for quasilinear anisotropic degenerate parabolic equations of second order with the form

$$(1.1) \quad \partial_t u + \operatorname{div} f(u) = \nabla \cdot (A(u) \nabla u), \quad u(0, x) = u_0(x),$$

where  $(t, x) \in \mathbf{R}_+ \times \mathbf{R}^d$ ,  $\operatorname{div}$  and  $\nabla$  are with respect to  $x \in \mathbf{R}^d$ ,  $u = u(t, x)$  is the scalar unknown function that is sought,

$$(1.2) \quad u_0 \in L^1(\mathbf{R}^d) \cap L^\infty(\mathbf{R}^d)$$

is the initial function,

$$(1.3) \quad f = (f_1, \dots, f_d) \in (\operatorname{Lip}_{\operatorname{loc}}(\mathbf{R}))^d$$

is the vector-valued flux function, and

$$(1.4) \quad A(u) = \sigma^A(u) \sigma^A(u)^\top \geq 0, \quad \text{with } \sigma^A \in (L^\infty_{\operatorname{loc}}(\mathbf{R}))^{d \times K}, \quad 1 \leq K \leq d,$$

is the matrix-valued diffusion function. The symmetric  $d \times d$  matrix

$$A(u) = (a_{ij}(u))$$

---

Received by the editors January 11, 2004.

2000 *Mathematics Subject Classification*. Primary 35K65, 35B35, 35G25, 35D99.

*Key words and phrases*.  $L^1$ –framework, degenerate parabolic equations, quasilinear, anisotropic, entropy solutions, kinetic formulation, continuous dependence, error estimates, vanishing viscosity, difference schemes.

©2004 American Mathematical Society  
Reverts to public domain 28 years from publication

has entries of the form

$$a_{ij}(u) = \sum_{k=1}^K \sigma_{ik}^A(u) \sigma_{jk}^A(u), \quad i, j = 1, \dots, d.$$

On the space of symmetric matrices, we employ the usual ordering in the sense of quadratic forms. Note that the scalar hyperbolic conservation law ( $A \equiv 0$ ) is a special case of (1.1).

Nonlinear partial differential equations of type (1.1) model convection-diffusion motions in nature and occur in a variety of applications. Being very selective, we mention here only flow in porous media (see, e.g., [12] and the references therein) and sedimentation-consolidation processes [6]. It is well known that (1.1) possesses discontinuous solutions, and weak solutions are not uniquely determined by their initial data; hence (1.1) must be interpreted in the sense of entropy solutions [19], [30], [31]. The uniqueness of entropy solutions was proved in the one-dimensional context by Wu-Yin [32] and Bénilan-Touré [2]. In the multi-dimensional context with isotropic diffusion (that is,  $A(\cdot) \geq 0$  is a scalar function), a general uniqueness result is much more recent and was proved by Carrillo [7] by using Kružkov's doubling of variables technique. Various extensions of his result can be found in Bürger-Evje-Karlsen [5], Eymard-Gallouët-Herbin-Michel [15], Karlsen-Risebro [18], Karlsen-Ohlberger [16], Mascia-Porretta-Terracina [24], Michel-Vovelle [25], and Rouvre-Gagneux [28]. Chen-DiBenedetto [8] proved the uniqueness of unbounded entropy solutions by using the doubling of variables technique. Chen-Perthame [9] finally introduced the notion of kinetic solutions and established an  $L^1$  well-posedness theory for the general anisotropic diffusion case by developing a kinetic approach for (1.1). Let us also mention the earlier work by Tassa [29], who proved the uniqueness for piecewise smooth weak solutions. There are also several recent studies concerned with the convergence of various numerical schemes; see [12] for operator splitting methods, [13], [17] for monotone finite difference schemes, [15], [26], [25] for monotone finite volume schemes, and [1], [4] for BGK schemes. All these papers provide the  $L^1$  convergence of approximate solutions without a rate of convergence (an error estimate). As is well known,  $L^1$  error estimates are more desirable for robust scientific computation and prediction, which have been an open problem for the general anisotropic case in numerical analysis.

In the hyperbolic context (i.e.,  $A \equiv 0$ ), error estimates for the vanishing isotropic viscosity method were derived first in Kuznetsov [20] and more recently in Cockburn-Gremaud [10] and Bouchut-Perthame [3], [27], while various estimates for continuous dependence on the nonlinearity (i.e., the flux function  $f$ ) were obtained first in Lucier [22] and later in Bouchut-Perthame [3]. Regarding degenerate parabolic problems with isotropic diffusion (that is,  $A(\cdot)$  is a scalar function), continuous dependence estimates for semigroup solutions, and hence also error estimates for the vanishing isotropic viscosity method, were obtained by Cockburn-Gripenberg [11]; see also [14], [18] for a different approach for the case that the flux function  $f$  also depends on  $(t, x)$ .

We are concerned with explicit estimates for continuous dependence on the nonlinearities and error estimates for the vanishing anisotropic viscosity method for (1.1). We mention that, even in the isotropic case, continuous dependence estimates have never been derived directly for entropy solutions. The purpose of this

paper is to use the Chen-Perthame kinetic approach [9] to develop an abstract  $L^1$ -framework for continuous dependence and error estimates for (1.1) and to present several applications of this framework.

More precisely, we are interested in comparing an entropy solution  $u = u(t, x)$  of (1.1) with an entropy solution  $v = v(t, x)$  of

$$(1.5) \quad \partial_t v + \operatorname{div} g(v) = \nabla \cdot (B(v)\nabla v) + \text{error terms}, \quad v(0, x) = v_0(x),$$

where

$$(1.6) \quad v_0 \in L^1(\mathbf{R}^d) \cap L^\infty(\mathbf{R}^d),$$

$$(1.7) \quad g = (g_1, \dots, g_d) \in (\operatorname{Lip}_{\text{loc}}(\mathbf{R}))^d,$$

and

$$(1.8) \quad B = \sigma^B(v)\sigma^B(v)^\top \geq 0, \quad \sigma^B \in (L^\infty_{\text{loc}}(\mathbf{R}))^{d \times \tilde{K}}, \quad 1 \leq K \leq d.$$

The symmetric  $d \times d$  matrix  $B = B(v) = (b_{ij}(v))$  has entries

$$b_{ij}(v) = \sum_{k=1}^{\tilde{K}} \sigma_{ik}^B(v)\sigma_{jk}^B(v), \quad i, j = 1, \dots, d.$$

Similar to the treatment of hyperbolic problems [3], [27], the error terms will take the form of “partial derivatives” for applications, which will be specified later in Section 3.

The first application of our general  $L^1$ -framework is an explicit estimate for continuous dependence on the nonlinearities in (1.1). If  $g \equiv f$  (see Section 4 for the general case),  $u_0 \in BV(\mathbf{R}^d)$ , and the error terms in (1.5) are zero, we obtain that, for any  $t > 0$ ,

$$(1.9) \quad \begin{aligned} & \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ & \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + C\sqrt{t} \sqrt{\left\| \left( \sqrt{A} - \sqrt{B} \right) \left( \sqrt{A} - \sqrt{B} \right)^\top \right\|_\infty}, \end{aligned}$$

where the  $\infty$ -norm is taken componentwise (see Section 3 for the precise definition). We must emphasize that the proof of a result like (1.9) depends in a fundamental way on using the parabolic dissipation/defect measure identified in Chen-Perthame [9], which is also the cornerstone of the uniqueness proof in [9].

The second application of our  $L^1$ -framework is an error estimate for the vanishing anisotropic viscosity method for (1.1):

$$(1.10) \quad \partial_t v + \operatorname{div} f(v) = \nabla \cdot (A(v)\nabla v) + \mu \nabla \cdot (B(v)\nabla v), \quad v(0, x) = v_0(x),$$

where the matrix  $B(v) > 0$  is of the same type as in (1.8). If  $u_0 \in BV(\mathbf{R}^d)$ , we prove that, for any  $t > 0$ ,

$$(1.11) \quad \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + C\sqrt{t\mu},$$

where  $C$  depends only on the  $L^\infty$  norms of the matrices  $A$  and  $B$ .

Within our  $L^1$ -framework, there are two ways to obtain an  $L^1$  estimate for  $u - v$ . A traditional way is to view the equation for the anisotropic viscous approximate solutions as the original equation perturbed by the error terms taking the form of partial derivatives. If  $v$  is uniformly  $BV$  bounded in space variables, one obtains

the optimal  $\frac{1}{2}$  rate of convergence. However, if  $v$  is not  $BV$  bounded, only a sub-optimal rate of convergence can be obtained in this way. The more efficient way is to derive the optimal rate of convergence from an estimate like (1.9) for continuous dependence with  $B$  properly chosen, without imposition of bounded variation of  $v$ . Indeed, in this paper we apply the second way to establish the optimal rate of convergence for the vanishing anisotropic viscosity method for (1.1).

While the vanishing anisotropic viscosity method has received almost no attention in the literature, the vanishing isotropic viscosity method for the purely hyperbolic case ( $A \equiv 0$ ) is well studied [3], [10], [19], [20], [27]. One motivation for studying the vanishing anisotropic viscosity method is that anisotropic viscosity approximations are closely related to finite volume numerical schemes on unstructured grids, for which uniform  $BV$  bounds are not available for finite volume schemes, and the standard error estimate theory for hyperbolic problems provides only a sub-optimal rate of convergence (see Makridakis and Perthame [23]).

Although the significant applications of our  $L^1$ -framework are the estimate for continuous dependence on the nonlinearities and the error estimate for the vanishing anisotropic viscosity method, as an example of direct applications of this framework to numerical methods, we focus in Section 6 on a linear convection-diffusion model equation and derive an  $L^1$  error estimate for a upwind-central difference scheme. We will present further applications of our  $L^1$ -framework to numerical methods for nonlinear degenerate parabolic-hyperbolic equations elsewhere. Also we remark that the results in this paper can be extended to more general equations with  $(t, x)$ -dependent coefficients; the details will be presented elsewhere.

This paper is organized as follows. We first establish the  $L^1$ -framework for continuous dependence and error estimates in Sections 2 and 3. Then we apply our general  $L^1$ -framework to obtain the following results: (i) an explicit estimate for continuous dependence on the nonlinearities in Section 4; (ii) an optimal error estimate for the anisotropic vanishing viscosity method in Section 5; (iii) an error estimate for an upwind-central finite difference scheme for a linear convection-diffusion equation in Section 6.

## 2. ENTROPY SOLUTIONS AND KINETIC FORMULATION

For any entropy function  $\eta : \mathbf{R} \rightarrow \mathbf{R}$ , the corresponding entropy fluxes

$$q = (q_1, \dots, q_d) : \mathbf{R} \rightarrow \mathbf{R}^d \quad \text{and} \quad R = (r_{ij}) : \mathbf{R} \rightarrow \mathbf{R}^{d \times d}$$

are defined by

$$q'(u) = \eta'(u)f'(u), \quad R'(u) = \eta'(u)A(u).$$

We will refer to  $(\eta, q, R)$  as an *entropy-entropy flux triple*.

For  $i = 1, \dots, d$  and  $k = 1, \dots, K$ , we let

$$\zeta_{ik}^A(u) = \int_0^u \sigma_{ik}^A(w) dw$$

and

$$\zeta_{ik}^{A,\psi}(u) = \int_0^u \sqrt{\psi(w)} \sigma_{ik}^A(w) dw, \quad \text{for } \psi \in C_0(\mathbf{R}).$$

According to Chen-Perthame [9], entropy solutions can now be defined as follows.

**Definition 2.1** (entropy solutions). A function

$$u \in L^\infty(\mathbf{R}_+; L^1(\mathbf{R}^d)) \cap L^\infty(\mathbf{R}_+ \times \mathbf{R}^d)$$

is an entropy solution of the Cauchy problem (1.1) if the following conditions are satisfied:

(D.1) For any  $k = 1, \dots, K$ ,

$$\sum_{i=1}^d \partial_{x_i} \zeta_{ik}^A(u) \in L^2(\mathbf{R}_+ \times \mathbf{R}^d).$$

(D.2) For any  $k = 1, \dots, K$  and  $\psi \in C_0(\mathbf{R})$  with  $\psi \geq 0$ ,

$$\sum_{i=1}^d \partial_{x_i} \zeta_{ik}^{A,\psi}(u) = \sqrt{\psi(u)} \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^A(u) \in L^2(\mathbf{R}_+ \times \mathbf{R}^d),$$

and the parabolic dissipation measure  $n^{u,\psi}(t, x)$ , defined by

$$n^{u,\psi}(t, x) = \psi(u(t, x)) \sum_{k=1}^K \left( \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^A(u(t, x)) \right)^2,$$

satisfies

$$n^{u,\psi}(t, x) = \sum_{k=1}^K \left( \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^{A,\psi}(u(t, x)) \right)^2 \quad \text{a.e. in } \mathbf{R}_+ \times \mathbf{R}^d.$$

(D.3) There exists an entropy dissipation measure  $m^{u,\psi}(t, x)$  of the form

$$m^{u,\psi}(t, x) = \int_{\mathbf{R}} m^u(\xi, t, x) \psi(\xi) d\xi, \quad \text{for any } \psi \in C_0(\mathbf{R}),$$

for some nonnegative entropy defect measure  $m^u(\xi, t, x)$  such that, for any  $C^2$  entropy-entropy flux triple  $(\eta, q, R)$  with  $\eta'' \in C_0(\mathbf{R})$ , there holds

$$(2.1) \quad \partial_t \eta(u) + \operatorname{div} q(u) - \nabla \cdot (R'(u) \nabla u) = - \left( m^{u,\eta''} + n^{u,\eta''} \right) \quad \text{in } \mathcal{D}'(\mathbf{R}_+ \times \mathbf{R}^d),$$

with initial data  $\eta(u)|_{t=0} = \eta(u_0)$ . That is, for any test function  $\phi \in \mathcal{D}(\mathbf{R}_+ \times \mathbf{R}^d)$ ,

$$\begin{aligned} \int_{\mathbf{R}_+ \times \mathbf{R}^d} \left( \eta(u) \partial_t \phi + \sum_{i=1}^d q_i(u) \partial_{x_i} \phi + \sum_{i,j=1}^d r_{ij}(u) \partial_{x_i x_j}^2 \phi \right) dx dt \\ + \int_{\mathbf{R}^d} \eta(u_0(x)) \phi(0, x) dx = \int_{\mathbf{R}_+ \times \mathbf{R}^d} \left( m^{u,\eta''} + n^{u,\eta''} \right) \phi dx dt. \end{aligned}$$

*Remark 2.1.* The nonnegative parabolic defect measure  $n^u(\xi, t, x)$  can be defined as

$$(2.2) \quad n^u(\xi, t, x) = \delta(\xi - u(t, x)) \sum_{k=1}^K \left( \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^A(u(t, x)) \right)^2.$$

Using the duality  $(C_0(\mathbf{R}); \mathcal{M}(\mathbf{R}))$ , the parabolic dissipation measure  $n^{u,\psi}$  then takes the form

$$n^{u,\psi}(t, x) = \int_{\mathbf{R}} n^u(\xi, t, x) \psi(\xi) d\xi, \quad \psi \in C_0(\mathbf{R}).$$

In the “diagonal case”  $a_{ij} \equiv 0$  for all  $i \neq j$ , the chain rule (D.2) is automatically satisfied.

We also follow Chen-Perthame [9] to give the equivalent kinetic formulation of entropy solutions for (1.1) which can be derived essentially from duality and the representation formula

$$(2.3) \quad h(u) - h(0) = \int_{\mathbf{R}} h'(\xi) \chi(\xi; u) d\xi, \quad \text{for any } h \in C^1,$$

where the indicator function  $\chi(\xi; u)$  is defined by

$$\chi(\xi; u) = \begin{cases} \mathbf{1}_{0 < \xi < u}, & \text{when } u > 0, \\ 0, & \text{when } u = 0, \\ -\mathbf{1}_{u < \xi < 0}, & \text{when } u < 0; \end{cases}$$

see also Lions-Perthame-Tadmor [21].

For later use, we note that the following formulas are valid:

$$(2.4) \quad \partial_u \chi(\xi; u) = \delta(\xi - u), \quad \partial_\xi \chi(\xi; u) = \delta(\xi) - \delta(\xi - u).$$

**Definition 2.2** (kinetic formulation). Let  $u$  be an entropy solution of (1.1) in the sense of Definition 2.1. Then the kinetic formulation of (1.1) reads

$$(2.5) \quad \begin{aligned} & \partial_t \chi(\xi; u) + f'(\xi) \cdot \nabla_x \chi(\xi; u) \\ &= \sum_{i,j=1}^d a_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; u) + \partial_\xi (m^u + n^u)(\xi, t, x) \quad \text{in } \mathcal{D}'_{\xi, t, x}, \\ & \chi(\xi; u)|_{t=0} = \chi(\xi; u_0), \end{aligned}$$

for some nonnegative entropy defect measure  $m^u$ , which measures “hyperbolicity” in the solution, and some nonnegative parabolic defect measure  $n^u$  with the form (2.2), which measures “parabolicity” in the solution.

### 3. GENERAL $L^1$ -FRAMEWORK

Let  $u$  be an entropy solution of the original problem (1.1). Let  $g$  be the flux function defined in (1.7) and  $B = (b_{ij})$  be the  $d \times d$  symmetric matrix defined in (1.8). We then let  $v$  solve the “approximate” kinetic problem

$$(3.1) \quad \begin{aligned} & \partial_t \chi(\xi; v) + g'(\xi) \cdot \nabla_x \chi(\xi; v) \\ &= \sum_{i,j=1}^d b_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; v) + \partial_\xi (m^v + n^v + E)(\xi, t, x) \quad \text{in } \mathcal{D}'_{\xi, t, x}, \\ & \chi(\xi; v)|_{t=0} = \chi(\xi; v_0), \end{aligned}$$

for some nonnegative entropy defect measure  $m^v$  and nonnegative parabolic defect measure  $n^v$  taking the particular form

$$n^v(\xi, t, x) = \delta(\xi - v(t, x)) \sum_{k=1}^K \left( \sum_{i=1}^d \zeta_{ik}^B(v(t, x)) \right)^2$$

with

$$\zeta_{ik}^B(v) = \int_0^v \sigma_{ik}^B(w) dw.$$

Correspondingly, for  $\psi \in C_0(\mathbf{R})$ , define the function

$$\zeta_{ik}^{B,\psi}(v) = \int_0^v \sqrt{\psi(w)} \sigma_{ik}^B(w) dw.$$

Motivated by Bouchut-Perthame [3] and Perthame [27] in their treatment of the hyperbolic problem, we assume that the error term  $E(\xi, t, x)$  takes the form of “partial derivatives”:

$$E(\xi, t, x) = \partial_t^{J_0} e_0(\xi, t, x) + \sum_{\substack{J=(J_1, \dots, J_d) \\ 0 \leq |J| \leq J_*}} D_x^J e_1^J(\xi, t, x)$$

for some error terms  $e_0$  and  $e_1^J$  with  $J_0, J_* \geq 0$  integers and  $J$  multi-indices. We assume that the error terms  $e_0$  and  $e_1^J$  satisfy

$$\left\| \left( \sup_{\xi} |e_0(\xi, \cdot, \cdot)|, \sup_{\xi} |e_1^J(\xi, \cdot, \cdot)| \right) \right\|_{L^1_{loc}(\mathbf{R}_+; L^1(\mathbf{R}^d))} < \infty, \quad 0 \leq |J| \leq J_*,$$

where sup is taken over all

$$\xi \in I(v) := [\inf v, \sup v].$$

Define the  $d \times d$  symmetric matrix

$$\begin{aligned} S(\xi) &= \left( \sqrt{A(\xi)} - \sqrt{B(\xi)} \right) \left( \sqrt{A(\xi)} - \sqrt{B(\xi)} \right)^\top \\ &= (\sigma^A(\xi) - \sigma^B(\xi)) (\sigma^A(\xi) - \sigma^B(\xi))^\top. \end{aligned}$$

Then the entries of  $S(\xi) = (s_{ij}(\xi))$  take the form

$$\begin{aligned} (3.2) \quad s_{ij}(\xi) &= \sum_{k=1}^K (\sigma_{ik}^A(\xi) - \sigma_{ik}^B(\xi)) (\sigma_{jk}^A(\xi) - \sigma_{jk}^B(\xi)) \\ &= \sum_{k=1}^K \left\{ \sigma_{ik}^A(\xi) \sigma_{jk}^A(\xi) - \sigma_{ik}^A(\xi) \sigma_{jk}^B(\xi) - \sigma_{ik}^B(\xi) \sigma_{jk}^A(\xi) + \sigma_{ik}^B(\xi) \sigma_{jk}^B(\xi) \right\}. \end{aligned}$$

To state the following theorem, we use the notations

$$S_\infty := \|S\|_\infty = \sup_{\xi \in I(u_0)} \|S(\xi)\|_\infty = \sup_{\substack{\xi \in I(u_0) \\ i,j=1, \dots, d}} |s_{ij}(\xi)|$$

and

$$\|f' - g'\|_\infty := \sup_{\xi \in I(u_0)} \|f'(\xi) - g'(\xi)\|_\infty = \sup_{\substack{\xi \in I(u_0) \\ i=1, \dots, d}} |f'_i(\xi) - g'_i(\xi)|.$$

Hereafter,  $C$  will denote positive constants, not necessarily the same at different occurrences, which are independent of the small parameters and time variable  $t$ .

The main result of this section is the following abstract  $L^1$ -framework for error estimates.

**Theorem 3.1** (general  $L^1$ -framework). *Let  $u \in C(\mathbf{R}_+; L^1(\mathbf{R}^d))$  be an entropy solution of (1.1), and suppose  $v \in L^\infty(\mathbf{R}_+; L^1(\mathbf{R}^d)) \cap L^\infty(\mathbf{R}_+ \times \mathbf{R}^d) \cap C(\mathbf{R}_+; L^1(\mathbf{R}^d))$  solves the “approximate” kinetic problem (3.1). Then, for any  $t > 0$  and any  $\varepsilon_1, \tilde{\varepsilon}_0, \tilde{\varepsilon}_1 > 0$ ,*

$$(3.3) \quad \begin{aligned} & \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} \\ & + C \left( \mathcal{E}_{u,t}^x(\varepsilon_1) + \mathcal{E}_{v,t}^t(\tilde{\varepsilon}_0) + \mathcal{E}_{v,t}^x(\tilde{\varepsilon}_1) + \mathcal{E}_t^{f-g}(\varepsilon_1) + \mathcal{E}_t^{A-B}(\varepsilon_1) + \mathcal{E}_{v,t}(\tilde{\varepsilon}_0, \tilde{\varepsilon}_1) \right), \end{aligned}$$

where

$$\begin{aligned} \mathcal{E}_{u,t}^x(\varepsilon_1) &= \sup_{\substack{|y| < \varepsilon_1 \\ \tau=0,t}} \|u(\tau, \cdot + y) - u(\tau, \cdot)\|_{L^1(\mathbf{R}^d)}, \\ \mathcal{E}_{v,t}^t(\tilde{\varepsilon}_0) &= \sup_{\substack{0 < s-\tau < \tilde{\varepsilon}_0 \\ \tau=0,t}} \|v(s, \cdot) - v(\tau, \cdot)\|_{L^1(\mathbf{R}^d)}, \\ \mathcal{E}_{v,t}^x(\tilde{\varepsilon}_1) &= \sup_{\substack{|y| < \tilde{\varepsilon}_1 \\ \tau=0,t}} \|v(\tau, \cdot + y) - v(\tau, \cdot)\|_{L^1(\mathbf{R}^d)}, \end{aligned}$$

$$\begin{aligned} \mathcal{E}_t^{f-g}(\varepsilon_1) &= \begin{cases} \|u_0\|_{L^1(\mathbf{R}^d)} \frac{t\|f'-g'\|_\infty}{\varepsilon_1}, & u_0 \notin BV(\mathbf{R}^d), \\ \|u_0\|_{BV(\mathbf{R}^d)} t\|f'-g'\|_\infty, & u_0 \in BV(\mathbf{R}^d), \end{cases} \\ \mathcal{E}_t^{A-B}(\varepsilon_1) &= \begin{cases} \|u_0\|_{L^1(\mathbf{R}^d)} \frac{tS_\infty}{\varepsilon_1^2}, & u_0 \notin BV(\mathbf{R}^d), \\ \|u_0\|_{BV(\mathbf{R}^d)} \frac{tS_\infty}{\varepsilon_1}, & u_0 \in BV(\mathbf{R}^d) \end{cases} \end{aligned}$$

and

$$\begin{aligned} \mathcal{E}_{v,t}(\tilde{\varepsilon}_0, \tilde{\varepsilon}_1) &= \frac{1}{\tilde{\varepsilon}_0^{J_0}} \left\| \sup_{\xi} |e_0(\xi, \cdot, \cdot)| \right\|_{L^1(0,t+\tilde{\varepsilon}_0; L^1(\mathbf{R}^d))} \\ &+ \sum_{\substack{J=(J_1, \dots, J_d) \\ 0 \leq |J| \leq J_*}} \frac{1}{\tilde{\varepsilon}_1^{|J|}} \left\| \sup_{\xi} |e_1^J(\xi, \cdot, \cdot)| \right\|_{L^1(0,t+\tilde{\varepsilon}_0; L^1(\mathbf{R}^d))}. \end{aligned}$$

If  $g \equiv f$  and  $B \equiv A$ , then the terms  $\mathcal{E}_{u,t}^x(\varepsilon_1)$ ,  $\mathcal{E}_t^{f-g}(\varepsilon_1)$ , and  $\mathcal{E}_t^{A-B}(\varepsilon_1)$  in (3.3) can be dropped, that is, there holds

$$(3.4) \quad \begin{aligned} & \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} \\ & + C \left( \mathcal{E}_{v,t}^t(\tilde{\varepsilon}_0) + \mathcal{E}_{v,t}^x(\tilde{\varepsilon}_1) + \mathcal{E}_{v,t}(\tilde{\varepsilon}_0, \tilde{\varepsilon}_1) \right) \end{aligned}$$

for any  $t > 0$  and  $\tilde{\varepsilon}_0, \tilde{\varepsilon}_1 > 0$ .

*Proof.* Some arguments in this proof follow Chen-Perthame [9] closely, for which we are very concise here and refer instead to [9] for more details.

We set  $\varepsilon = (\varepsilon_0, \varepsilon_1)$ ,  $\varepsilon_0 > 0$  for the forward time regularization and  $\varepsilon_1 > 0$  for the space regularization. We then define

$$\omega_\varepsilon(t, x) := \omega_{\varepsilon_0}(t)\omega_{\varepsilon_1}(x),$$

where

$$\omega_{\varepsilon_0}(t) := \frac{1}{\varepsilon_0} \omega_0\left(\frac{t}{\varepsilon_0}\right), \quad \omega_{\varepsilon_1}(x) := \frac{1}{\varepsilon_1^d} \omega_1\left(\frac{x_1}{\varepsilon_1}\right) \cdots \omega_1\left(\frac{x_d}{\varepsilon_1}\right),$$

and  $\omega_\ell \geq 0$ ,  $\ell = 0, 1$ , denote the normalized regularization kernels with

$$\int_{\mathbf{R}} \omega_\ell(\tau) d\tau = 1, \quad \text{supp}(\omega_0) \subset (-1, 0), \quad \text{supp}(\omega_1) \subset (-1, 1).$$

We use the notations

$$\begin{aligned} \chi &:= \chi(\xi, t, x) = \chi(\xi; u(t, x)), & \tilde{\chi} &:= \tilde{\chi}(\xi, t, x) = \chi(\xi; v(t, x)), \\ \chi_\varepsilon &:= \chi_\varepsilon(\xi, t, x) = \left(\chi \star_{(t,x)} \omega_\varepsilon\right)(\xi, t, x), & \tilde{\chi}_\varepsilon &:= \tilde{\chi}_\varepsilon(\xi, t, x) = \left(\tilde{\chi} \star_{(t,x)} \omega_\varepsilon\right)(\xi, t, x), \end{aligned}$$

where  $\tilde{\varepsilon} = (\tilde{\varepsilon}_0, \tilde{\varepsilon}_1) > 0$  is another pair of time-space regularization parameters. Moreover, we use the notations

$$\begin{aligned} m_\varepsilon^u &:= m_\varepsilon^u(\xi, t, x) = \left(m^u \star_{(t,x)} \omega_\varepsilon\right)(\xi, t, x), \\ n_\varepsilon^u &:= n_\varepsilon^u(\xi, t, x) = \left(n^u \star_{(t,x)} \omega_\varepsilon\right)(\xi, t, x), \\ m_{\tilde{\varepsilon}}^v &:= m_{\tilde{\varepsilon}}^v(\xi, t, x) = \left(m^v \star_{(t,x)} \omega_{\tilde{\varepsilon}}\right)(\xi, t, x), \\ n_{\tilde{\varepsilon}}^v &:= n_{\tilde{\varepsilon}}^v(\xi, t, x) = \left(n^v \star_{(t,x)} \omega_{\tilde{\varepsilon}}\right)(\xi, t, x), \end{aligned}$$

and  $E_{\tilde{\varepsilon}} = E_{\tilde{\varepsilon}}(\xi, t, x)$ , which is similarly defined.

We intend to study the microscopic functional

$$0 \leq Q_{\varepsilon, \tilde{\varepsilon}}(\xi, t, x) = |\chi_\varepsilon| + |\tilde{\chi}_{\tilde{\varepsilon}}| - 2\chi_\varepsilon \tilde{\chi}_{\tilde{\varepsilon}}.$$

More precisely, we will calculate

$$(3.5) \quad \frac{d}{dt} \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, t, x) dx d\xi.$$

Note that  $\chi_\varepsilon(\xi, t, x)$  satisfies

$$(3.6) \quad \partial_t \chi_\varepsilon + f'(\xi) \cdot \nabla_x \chi_\varepsilon = \sum_{i,j=1}^d a_{ij}(\xi) \partial_{x_i x_j}^2 \chi_\varepsilon + \partial_\xi (m_\varepsilon^u + n_\varepsilon^u),$$

and that  $\tilde{\chi}_{\tilde{\varepsilon}}(\xi, t, x)$  satisfies

$$(3.7) \quad \partial_t \tilde{\chi}_{\tilde{\varepsilon}} + g'(\xi) \cdot \nabla_x \tilde{\chi}_{\tilde{\varepsilon}} = \sum_{i,j=1}^d b_{ij}(\xi) \partial_{x_i x_j}^2 \tilde{\chi}_{\tilde{\varepsilon}} + \partial_\xi (m_{\tilde{\varepsilon}}^v + n_{\tilde{\varepsilon}}^v + E_{\tilde{\varepsilon}}).$$

Multiplying (3.6) by  $\text{sign}(\xi)$ , using  $\text{sign}(\xi)\chi_\varepsilon = |\chi_\varepsilon|$ , and then integrating in  $(\xi, x) \in \mathbf{R}_\xi \times \mathbf{R}_x^d$  yields

$$(3.8) \quad \frac{d}{dt} \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} |\chi_\varepsilon| dx d\xi = -2 \int_{\mathbf{R}_x^d} (m_\varepsilon^u + n_\varepsilon^u)(0, t, x) dx.$$

Similarly,

$$(3.9) \quad \frac{d}{dt} \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} |\tilde{\chi}_{\varepsilon}| \, dx \, d\xi = -2 \int_{\mathbf{R}_x^d} (m_\varepsilon^v + n_\varepsilon^v + E_\varepsilon)(0, t, x) \, dx.$$

We now consider the quadratic term. To this end, we need an additional regularization in the kinetic/velocity variable  $\xi$ :

$$\begin{aligned} \chi_{\varepsilon, \delta}(\xi, t, x) &:= \left( \chi_\varepsilon \star_\xi \psi_\delta \right) (\xi, t, x), \\ \tilde{\chi}_{\varepsilon, \delta}(\xi, t, x) &:= \left( \tilde{\chi}_\varepsilon \star_\xi \psi_\delta \right) (\xi, t, x) \end{aligned}$$

for a standard regularization kernel  $\psi_\delta$ . We also need a  $\xi$ -truncation  $\mathcal{T}_L(\xi)$ , which is a smooth nonnegative function with bounded support. That is,  $\mathcal{T}_L(\xi) = \mathcal{T}(\xi/L) \rightarrow 1$  as  $L \rightarrow \infty$  with

$$\begin{aligned} 0 \leq \mathcal{T}(\xi) \leq 1, & \quad \text{for } \xi \in (-\infty, \infty), \\ \mathcal{T}(\xi) = 1, & \quad \text{for } |\xi| \leq 1/2, \\ \mathcal{T}(\xi) = 0, & \quad \text{for } |\xi| \geq 1. \end{aligned}$$

The destiny of these additional parameters is that  $\delta \downarrow 0$  first and  $L \uparrow \infty$  second. Then  $\chi_{\varepsilon, \delta}$  satisfies

$$(3.10) \quad \begin{aligned} &\partial_t \chi_{\varepsilon, \delta} + f'(\xi) \cdot \nabla_x \chi_{\varepsilon, \delta} \\ &= \sum_{i, j=1}^d \partial_{x_i x_j}^2 \left( (a_{ij} \chi_\varepsilon) \star_\xi \psi_\delta \right) + \partial_\xi \left( (m_\varepsilon^u + n_\varepsilon^u) \star_\xi \psi_\delta \right) + R_{\varepsilon, \delta}^u, \end{aligned}$$

and  $\tilde{\chi}_{\varepsilon, \delta}$  satisfies

$$(3.11) \quad \begin{aligned} &\partial_t \tilde{\chi}_{\varepsilon, \delta} + g'(\xi) \cdot \nabla_x \tilde{\chi}_{\varepsilon, \delta} \\ &= \sum_{i, j=1}^d \partial_{x_i x_j}^2 \left( (b_{ij} \tilde{\chi}_\varepsilon) \star_\xi \psi_\delta \right) + \partial_\xi \left( (m_\varepsilon^v + n_\varepsilon^v + E_\varepsilon) \star_\xi \psi_\delta \right) + R_{\varepsilon, \delta}^v. \end{aligned}$$

In (3.10) and (3.11),

$$\begin{aligned} R_{\varepsilon, \delta}^u &= \operatorname{div}_x \left( f'(\xi) \chi_{\varepsilon, \delta} - (f' \chi_\varepsilon) \star_\xi \psi_\delta \right), \\ R_{\varepsilon, \delta}^v &= \operatorname{div}_x \left( g'(\xi) \tilde{\chi}_{\varepsilon, \delta} - (g' \tilde{\chi}_\varepsilon) \star_\xi \psi_\delta \right). \end{aligned}$$

A simple calculation reveals

$$\begin{aligned}
 (3.12) \quad & \frac{d}{dt} \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \chi_{\varepsilon, \delta} \tilde{\chi}_{\tilde{\varepsilon}, \delta} dx d\xi \\
 &= - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \tilde{\chi}_{\tilde{\varepsilon}, \delta} (f'(\xi) - g'(\xi)) \cdot \nabla_x \chi_\varepsilon dx d\xi \\
 & \quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \tilde{\chi}_{\tilde{\varepsilon}, \delta} \sum_{i,j=1}^d \partial_{x_i x_j}^2 \left( (a_{ij} \chi_\varepsilon) \star_\xi \psi_\delta \right) dx d\xi \\
 & \quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \tilde{\chi}_{\tilde{\varepsilon}, \delta} \partial_\xi \left( (m_\varepsilon^u + n_\varepsilon^u) \star_\xi \psi_\delta \right) dx d\xi \\
 & \quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \chi_{\varepsilon, \delta} \sum_{i,j=1}^d \partial_{x_i x_j}^2 \left( (b_{ij} \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) dx d\xi \\
 & \quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \chi_{\varepsilon, \delta} \partial_\xi \left( (m_\varepsilon^v + n_\varepsilon^v + E_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) dx d\xi \\
 & \quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) (\tilde{\chi}_{\tilde{\varepsilon}, \delta} R_{\varepsilon, \delta}^u + \chi_{\varepsilon, \delta} R_{\varepsilon, \delta}^v) dx d\xi \\
 & =: \sum_{\ell=1}^6 I_\ell(t; \varepsilon, \tilde{\varepsilon}, \delta, L).
 \end{aligned}$$

As in Chen-Perthame [9], we have

$$(3.13) \quad \lim_{\delta \downarrow 0} I_6(t; \varepsilon, \delta, L) = 0 \quad \text{in } L^p(0, T) \text{ for any } 1 \leq p < \infty.$$

Writing out the convolution products explicitly, we have

$$\begin{aligned}
 & I_1(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\
 &= - \int \mathcal{T}_L(\xi) (f'(\xi) - g'(\xi)) \cdot \nabla_x \omega_\varepsilon(t - s, x - y) \omega_{\tilde{\varepsilon}}(t - s', x - y') \\
 & \quad \times \psi_\delta(\xi - \eta) \psi_\delta(\xi - \eta') \chi(\eta; u(s, y)) \chi(\eta'; v(s', y')) ds dy d\eta ds' dy' d\eta' dx d\xi.
 \end{aligned}$$

Sending first  $\delta \downarrow 0$  and second  $L \uparrow \infty$ , we get

$$\begin{aligned}
 (3.14) \quad & \lim_{L \uparrow \infty} \lim_{\delta \downarrow 0} I_1(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\
 &= - \int (f'(\xi) - g'(\xi)) \cdot \nabla_x \omega_\varepsilon(t - s, x - y) \omega_{\tilde{\varepsilon}}(t - s', x - y') \\
 & \quad \times \chi(\xi; u(s, y)) \chi(\xi; v(s', y')) ds dy ds' dy' dx d\xi \\
 & =: -\mathcal{E}^{f-g}(t; \varepsilon, \tilde{\varepsilon}).
 \end{aligned}$$

Integrating by parts yields

$$\begin{aligned}
 (3.15) \quad I_3(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &= - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}'_L(\xi) \tilde{\chi}_{\tilde{\varepsilon}, \delta} \left( (m_\varepsilon^u + n_\varepsilon^u) \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \psi_\delta(\xi) \left( (m_\varepsilon^u + n_\varepsilon^u) \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \left( \delta(\xi - u) \star_{(\xi, t, x)} (\omega_\varepsilon \psi_\delta) \right) \left( m_\varepsilon^u \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \left( \delta(\xi - u) \star_{(\xi, t, x)} (\omega_\varepsilon \psi_\delta) \right) \left( n_\varepsilon^u \star_\xi \psi_\delta \right) dx d\xi \\
 &=: \sum_{\ell=1}^4 I_{3,\ell}(t; \varepsilon, \tilde{\varepsilon}, \delta, L)
 \end{aligned}$$

and

$$\begin{aligned}
 (3.16) \quad I_5(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &= - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}'_L(\xi) \chi_{\varepsilon, \delta} \left( (m_\varepsilon^v + n_\varepsilon^v + E_\varepsilon) \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \psi_\delta(\xi) \left( (m_\varepsilon^v + n_\varepsilon^v + E_\varepsilon) \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \left( \delta(\xi - v) \star_{(\xi, t, x)} (\omega_\varepsilon \psi_\delta) \right) \left( m_\varepsilon^v \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \left( \delta(\xi - v) \star_{(\xi, t, x)} (\omega_\varepsilon \psi_\delta) \right) \left( n_\varepsilon^v \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \left( \delta(\xi - v) \star_{(\xi, t, x)} (\omega_\varepsilon \psi_\delta) \right) \left( E_\varepsilon \star_\xi \psi_\delta \right) dx d\xi \\
 &=: \sum_{\ell=1}^5 I_{5,\ell}(t; \varepsilon, \tilde{\varepsilon}, \delta, L).
 \end{aligned}$$

As in [9], we have

$$(3.17) \quad \lim_{L \uparrow \infty} \lim_{\delta \downarrow 0} I_{3,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) = \lim_{L \uparrow \infty} \lim_{\delta \downarrow 0} I_{5,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) = 0$$

in  $L^p(0, T)$  for any  $1 \leq p < \infty$ .

Moreover,

$$\begin{aligned}
 (3.18) \quad &\lim_{\delta \downarrow 0} \left( I_{3,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, R) + I_{5,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, R) \right) \\
 &= - \int_{\mathbf{R}_x^d} (m_\varepsilon^u + n_\varepsilon^u + m_\varepsilon^v + n_\varepsilon^v + E_\varepsilon)(0, t, x) dx
 \end{aligned}$$

in  $L^p(0, T)$  for any  $1 \leq p < \infty$ .

Clearly,

$$I_{3,3}(t; \varepsilon, \tilde{\varepsilon}, \delta, L), I_{5,3}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \geq 0 \quad \text{for any } t > 0.$$

Integrating by parts also yields

$$\begin{aligned}
 (3.19) \quad I_2(\varepsilon, \delta, L) &= - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \sum_{k=1}^K \sum_{i,j=1}^d \partial_{x_i} \left( (\sigma_{ik}^A \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) \partial_{x_j} \left( (\sigma_{jk}^A \chi_\varepsilon) \star_\xi \psi_\delta \right) dx d\xi \\
 &\quad + \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \sum_{k=1}^K \sum_{i,j=1}^d \left\{ \partial_{x_i} \left( (\sigma_{ik}^A \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) \partial_{x_j} \left( (\sigma_{jk}^A \chi_\varepsilon) \star_\xi \psi_\delta \right) \right. \\
 &\quad \quad \quad \left. - \partial_{x_i} \tilde{\chi}_{\tilde{\varepsilon},\delta} \partial_{x_j} \left( (a_{ij} \chi_\varepsilon) \star_\xi \psi_\delta \right) \right\} dx d\xi \\
 &=: I_{2,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{2,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, L).
 \end{aligned}$$

Similarly,

$$(3.20) \quad I_4(t; \varepsilon, \tilde{\varepsilon}, \delta, L) = I_{4,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{4,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, L),$$

where

$$\begin{aligned}
 (3.21) \quad I_{4,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &= - \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \sum_{k=1}^K \sum_{i,j=1}^d \partial_{x_i} \left( (\sigma_{ik}^B \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) \partial_{x_j} \left( (\sigma_{jk}^B \chi_\varepsilon) \star_\xi \psi_\delta \right) dx d\xi
 \end{aligned}$$

and

$$\begin{aligned}
 (3.22) \quad I_{4,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &= \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} \mathcal{T}_L(\xi) \sum_{k=1}^K \sum_{i,j=1}^d \left\{ \partial_{x_i} \left( (\sigma_{ik}^B \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) \partial_{x_j} \left( (\sigma_{jk}^B \chi_\varepsilon) \star_\xi \psi_\delta \right) \right. \\
 &\quad \quad \quad \left. - \partial_{x_i} \left( (b_{ij} \tilde{\chi}_{\tilde{\varepsilon}}) \star_\xi \psi_\delta \right) \partial_{x_j} \chi_{\varepsilon,\delta} \right\} dx d\xi.
 \end{aligned}$$

As in Chen-Perthame [9], we have

$$(3.23) \quad \lim_{\delta \downarrow 0} I_{2,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) = \lim_{\delta \downarrow 0} I_{4,2}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) = 0$$

in  $L^p(0, T)$  for any  $1 \leq p < \infty$ .

We now study the new term

$$\begin{aligned}
 (3.24) \quad \mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &:= I_{2,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{3,4}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{4,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{5,4}(t; \varepsilon, \tilde{\varepsilon}, \delta, L).
 \end{aligned}$$

From now on in this proof, for notational simplicity, we drop writing the domains of integration. Writing out explicitly the convolution products, we have

$$\begin{aligned}
 I_{2,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) &= - \sum_{k=1}^K \sum_{i,j=1}^d \int \mathcal{T}_L(\xi) \partial_{x_i} \omega_\varepsilon(t-s, x-y) \partial_{x_j} \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi-\eta) \psi_\delta(\xi-\eta') \\
 &\quad \times \sigma_{ik}^A(\eta) \chi(\eta; u(s, y)) \sigma_{jk}^A(\eta') \chi(\eta'; v(s', y')) ds dy d\eta ds' dy' d\eta' dx d\xi.
 \end{aligned}$$

Similarly,

$$\begin{aligned} & I_{4,1}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &= - \sum_{k=1}^K \sum_{i,j=1}^d \int \mathcal{T}_L(\xi) \partial_{x_i} \omega_\varepsilon(t-s, x-y) \partial_{x_j} \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi - \eta) \\ & \quad \times \psi_\delta(\xi - \eta') \sigma_{ik}^B(\eta) \chi(\eta; u(s, y)) \sigma_{jk}^B(\eta') \chi(\eta'; v(s', y')) ds dy d\eta ds' dy' d\eta' dx d\xi. \end{aligned}$$

Note that

$$\begin{aligned} & I_{3,4}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) + I_{5,4}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &= \sum_{k=1}^K \int \mathcal{T}_L(\xi) \omega_\varepsilon(t-s, x-y) \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi - u(s, y)) \psi_\delta(\xi - v(s', y')) \\ & \quad \times \left\{ \left( \sum_{i=1}^d \partial_{y_i} \zeta_{ik}^A(u(s, y)) \right)^2 + \left( \sum_{j=1}^d \partial_{y'_j} \zeta_{jk}^B(v(s', y')) \right)^2 \right\} ds dy ds' dy' dx d\xi \\ &\geq 2 \sum_{k=1}^K \int \mathcal{T}_L(\xi) \omega_\varepsilon(t-s, x-y) \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi - u(s, y)) \psi_\delta(\xi - v(s', y')) \\ & \quad \times \sum_{i=1}^d \partial_{y_i} \zeta_{ik}^A(u(s, y)) \sum_{j=1}^d \partial_{y'_j} \zeta_{jk}^B(v(s', y')) ds dy ds' dy' dx d\xi \\ &= 2 \sum_{k=1}^K \int \mathcal{T}_L(\xi) \omega_\varepsilon(t-s, x-y) \omega_{\tilde{\varepsilon}}(t-s', x-y') \\ & \quad \times \sum_{i=1}^d \partial_{y_i} \zeta_{ik}^{A, \psi_\delta(\xi-\cdot)^2}(u(s, y)) \sum_{j=1}^d \partial_{y'_j} \zeta_{jk}^{B, \psi_\delta(\xi-\cdot)^2}(v(s', y')) ds dy ds' dy' dx d\xi \\ &= 2 \sum_{k=1}^K \sum_{i,j=1}^d \int \mathcal{T}_L(\xi) \partial_{x_i} \omega_\varepsilon(t-s, x-y) \partial_{x_j} \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi - \eta) \psi_\delta(\xi - \eta') \\ & \quad \times \sigma_{ik}^A(\eta) \chi(\eta; u(s, y)) \sigma_{jk}^B(\eta') \chi(\eta'; v(s', y')) ds dy d\eta ds' dy' d\eta' dx d\xi, \end{aligned}$$

where we have used the chain rule (D.2), integration by parts, and (2.3). From this and the previous calculations, we find

$$\begin{aligned} & \mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &\geq - \sum_{k=1}^K \sum_{i,j=1}^d \int \mathcal{T}_L(\xi) \partial_{x_i} \omega_\varepsilon(t-s, x-y) \partial_{x_j} \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(\xi - \eta) \psi_\delta(\xi - \eta') \\ & \quad \times \left\{ \sigma_{ik}^A(\eta) \sigma_{jk}^A(\eta') - 2\sigma_{ik}^A(\eta) \sigma_{jk}^B(\eta') + \sigma_{ik}^B(\eta) \sigma_{jk}^B(\eta') \right\} \\ & \quad \times \chi(\eta; u(s, y)) \chi(\eta'; v(s', y')) ds dy d\eta ds' dy' d\eta' dx d\xi. \end{aligned}$$

After performing the changes of variables,

$$z = \xi - \eta, \quad z' = \xi - \eta', \quad dz dz' = d\eta d\eta',$$

we get

$$\begin{aligned} &\mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &\geq - \sum_{k=1}^K \sum_{i,j=1}^d \int \mathcal{T}_L(\xi) \partial_{x_i} \omega_\varepsilon(t-s, x-y) \partial_{x_j} \omega_{\tilde{\varepsilon}}(t-s', x-y') \psi_\delta(z) \psi_\delta(z') \\ &\quad \times \left\{ \sigma_{ik}^A(\xi-z) \sigma_{jk}^A(\xi-z') - 2\sigma_{ik}^A(\xi-z) \sigma_{jk}^B(\xi-z') + \sigma_{ik}^B(\xi-z) \sigma_{jk}^B(\xi-z') \right\} \\ &\quad \times \chi(\xi-z; u(s, y)) \chi(\xi-z'; v(s', y')) ds dy dz ds' dy' dz' dx d\xi. \end{aligned}$$

From this, it follows easily that

$$\begin{aligned} (3.25) \quad &\lim_{L \uparrow \infty} \lim_{\delta \downarrow 0} \mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &\geq \sum_{i,j=1}^d \int \partial_{x_i x_j}^2 \omega_\varepsilon(t-s, x-y) \omega_{\tilde{\varepsilon}}(t-s', x-y') \\ &\quad \times \left\{ \sigma_{ik}^A(\xi) \sigma_{jk}^A(\xi) - 2\sigma_{ik}^A(\xi) \sigma_{jk}^B(\xi) + \sigma_{ik}^B(\xi) \sigma_{jk}^B(\xi) \right\} \\ &\quad \times \chi(\xi; u(s, y)) \chi(\xi; v(s', y')) ds dy ds' dy' dx d\xi, \end{aligned}$$

where we have also performed integration by parts.

Now it is crucial to exploit the symmetry property of  $(i, j) \mapsto \partial_{x_i x_j}^2 \omega_\varepsilon(t-s, x-y)$  to obtain a favorable quadratic form:

$$\begin{aligned} (3.26) \quad &\sum_{i,j=1}^d \partial_{x_i x_j}^2 \omega_\varepsilon(t-s, x-y) \left\{ \sigma_{ik}^A(\xi) \sigma_{jk}^A(\xi) - 2\sigma_{ik}^A(\xi) \sigma_{jk}^B(\xi) + \sigma_{ik}^B(\xi) \sigma_{jk}^B(\xi) \right\} \\ &= \sum_{i,j=1}^d \partial_{x_i x_j}^2 \omega_\varepsilon(t-s, x-y) (\sigma_{ik}^A(\xi) - \sigma_{ik}^B(\xi)) (\sigma_{jk}^A(\xi) - \sigma_{jk}^B(\xi)). \end{aligned}$$

In view of (3.25), (3.26), and the definition of  $s_{ij}(\xi)$  in (3.2), we finally get

$$\begin{aligned} (3.27) \quad &\lim_{L \uparrow \infty} \lim_{\delta \downarrow 0} \mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\ &\geq \sum_{i,j=1}^d \int \partial_{x_i x_j}^2 \omega_\varepsilon(t-s, x-y) \omega_{\tilde{\varepsilon}}(t-s', x-y') \\ &\quad \times s_{ij}(\xi) \chi(\xi; u(s, y)) \chi(\xi; v(s', y')) ds dy ds' dy' dx d\xi \\ &=: -\mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}). \end{aligned}$$

Writing out the convolution products, we have

(3.28)

$$\begin{aligned}
 & I_{5,5}(t; \varepsilon, \tilde{\varepsilon}, \delta, L) \\
 &= \int \mathcal{T}_L(\xi) \psi_\delta(\xi - v(s, y)) \psi_\delta(\xi - \eta') \omega_{\tilde{\varepsilon}}(t - s, x - y) \\
 &\quad \times \left( \partial_t^{J_0} \omega_{\tilde{\varepsilon}}(t - s', x - y') e_0(\eta', s', y') \right. \\
 &\quad \left. + \sum_{\substack{J=(J_1, \dots, J_d) \geq 0 \\ |J| \leq J_*}} D_x^J \omega_{\tilde{\varepsilon}}(t - s', x - y') e_1^J(\eta', s', y') \right) ds dy d\eta ds' dy' d\eta' dx d\xi \\
 &\longrightarrow \int \omega_{\tilde{\varepsilon}}(t - s, x - y) \left( \partial_t^{J_0} \omega_{\tilde{\varepsilon}}(t - s', x - y') e_0(v(s, y), s', y') \right. \\
 &\quad \left. + \sum_{\substack{J=(J_1, \dots, J_d) \geq 0 \\ |J| \leq J_*}} D_x^J \omega_{\tilde{\varepsilon}}(t - s', x - y') e_1^J(v(s, y), s', y') \right) ds dy ds' dy' dx \\
 &=: -\mathcal{E}_v(t; \tilde{\varepsilon}), \quad \text{when } L \uparrow \infty \text{ and } \delta \downarrow 0.
 \end{aligned}$$

Summarizing our calculations from (3.12) to (3.28), we obtain that, for any  $\varepsilon, \tilde{\varepsilon} > 0$ ,

$$\begin{aligned}
 (3.29) \quad \frac{d}{dt} \int_{\mathbf{R}_\varepsilon \times \mathbf{R}_x^d} \chi_\varepsilon \tilde{\chi}_{\tilde{\varepsilon}} dx d\xi &\geq - \int_{\mathbf{R}_x^d} (m_\varepsilon^u + n_\varepsilon^u + m_\varepsilon^v + n_\varepsilon^v + E_{\tilde{\varepsilon}})(0, t, x) dx \\
 &\quad - \mathcal{E}^{f-g}(t; \varepsilon, \tilde{\varepsilon}) - \mathcal{E}^{A-B}(t; \varepsilon, \tilde{\varepsilon}) - \mathcal{E}_v(t; \tilde{\varepsilon}).
 \end{aligned}$$

Then the estimates (3.5), (3.8), (3.9), and (3.29) yield that, for any  $\varepsilon, \tilde{\varepsilon} > 0$ ,

$$\begin{aligned}
 & \int_{\mathbf{R}_\varepsilon \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, t, x) dx d\xi \\
 &\leq \int_{\mathbf{R}_\varepsilon \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, 0, x) dx d\xi + 2 \int_0^t \mathcal{E}^{f-g}(\tau; \varepsilon, \tilde{\varepsilon}) d\tau \\
 &\quad + 2 \int_0^t \mathcal{E}^{A-B}(\tau; \varepsilon, \tilde{\varepsilon}) d\tau + 2 \int_0^t \mathcal{E}_v(\tau; \varepsilon, \tilde{\varepsilon}) d\tau.
 \end{aligned}$$

Similarly, we have

$$\begin{aligned}
 & \int_{\mathbf{R}_\varepsilon \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, 0, x) dx d\xi \\
 &= \int \left( |u(s, y)| + |v(s', y')| - 2 \min(|u(s, y)|, |v(s', y')|) \mathbf{1}_{\{\text{sign}(u(s, y)v(s', y')) > 0\}} \right) \\
 &\quad \times \omega_\varepsilon(-s, x - y) \omega_{\tilde{\varepsilon}}(-s', x - y') ds dy ds' dy' dx \\
 &= \int |u(s, y) - v(s', y')| \omega_\varepsilon(-s, x - y) \omega_{\tilde{\varepsilon}}(-s', x - y') ds dy ds' dy' dx.
 \end{aligned}$$

A standard calculation reveals

$$\begin{aligned} & \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, 0, x) \, dx \, d\xi \\ & \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + \sup_{0 < s < \varepsilon_0} \|u(s, \cdot) - u_0(\cdot)\|_{L^1(\mathbf{R}^d)} \\ & \quad + \sup_{0 < s < \tilde{\varepsilon}_0} \|v(s, \cdot) - v_0(\cdot)\|_{L^1(\mathbf{R}^d)} + \sup_{|y| < \varepsilon_1} \|u_0(\cdot + y) - u_0(\cdot)\|_{L^1(\mathbf{R}^d)} \\ & \quad + \sup_{|y| < \tilde{\varepsilon}_1} \|v_0(\cdot + y) - v_0(\cdot)\|_{L^1(\mathbf{R}^d)}. \end{aligned}$$

Similarly, we find

$$\begin{aligned} & \int_{\mathbf{R}_\xi \times \mathbf{R}_x^d} Q_{\varepsilon, \tilde{\varepsilon}}(\xi, t, x) \, dx \, d\xi \\ & \geq \|u(t) - v(t)\|_{L^1(\mathbf{R}^d)} - \sup_{0 < s - t < \varepsilon_0} \|u(s, \cdot) - u(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ & \quad - \sup_{0 < s - t < \tilde{\varepsilon}_0} \|v(s, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} - \sup_{|y| \leq \varepsilon_1} \|u(t, \cdot + y) - u(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ & \quad - \sup_{|y| \leq \tilde{\varepsilon}_1} \|v(t, \cdot + y) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)}. \end{aligned}$$

Sending  $\varepsilon_0 \downarrow 0$ , we conclude that

$$\begin{aligned} & \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ & \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + 2\mathcal{E}_{u,t}^x(\varepsilon_1) + 2\mathcal{E}_{v,t}^t(\tilde{\varepsilon}_0) + 2\mathcal{E}_{v,t}^x(\tilde{\varepsilon}_1) \\ & \quad + 2 \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{f-g}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau + 2 \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{A-B}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau \\ & \quad + 2 \int_0^t \mathcal{E}_v(\tau; \tilde{\varepsilon}) \, d\tau, \quad 0 \leq t \leq T. \end{aligned}$$

It remains to estimate the last three terms.

Recall that  $u(t, x) \in I(u_0) = [\inf u_0, \sup u_0]$  for a.e.  $(t, x) \in \mathbf{R}_+ \times \mathbf{R}^d$ . It is easy to see that

$$\begin{aligned} & \left| \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{f-g}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau \right| \\ & \leq \int \|f' - g'\|_\infty \sum_{i=1}^d |\partial_{x_i} \omega_{\varepsilon_1}(x - y)| \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \\ & \quad \times \chi(\xi; u(\tau, y)) \, dy \, ds' \, dy' \, d\tau \, dx \, d\xi \\ & \leq C \|u\|_{L^\infty(0,T;L^1(\mathbf{R}^d))} \frac{t \|f' - g'\|_\infty}{\varepsilon_1}. \end{aligned}$$

Note that, if  $u_0 \in L^1(\mathbf{R}^d)$ , then  $\|u\|_{L^\infty(0,T;L^1(\mathbf{R}^d))} \leq \|u\|_{L^1(\mathbf{R}^d)} < \infty$  (see [31]).

On the other hand, if  $u \in L^\infty(\mathbf{R}_+; BV(\mathbf{R}^d))$ , then we can first integrate by parts in (3.14) and utilize (2.4) to obtain

$$\begin{aligned} & \left| \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{f-g}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau \right| \\ &= \left| \int \sum_{i=1}^d (f'_i(u(s, y)) - g'_i(u(s, y))) \omega_{\varepsilon_1}(x - y) \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \right. \\ & \quad \left. \times \partial_{y_i} u(\tau, y) \chi(u(\tau, y); v(s', y')) \, dy \, ds' \, dy' \, d\tau \, dx \right| \\ &\leq C \|u\|_{L^\infty(0, T; BV(\mathbf{R}^d))} t \|f' - g'\|_\infty. \end{aligned}$$

Note that, if  $u_0 \in BV(\mathbf{R}^d)$ , then  $\|u\|_{L^\infty(\mathbf{R}_+; BV(\mathbf{R}^d))} \leq \|u_0\|_{BV(\mathbf{R}^d)} < \infty$  (see [31]).

Again using the standard properties of regularization kernels, we find

$$\begin{aligned} & \left| \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{A-B}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau \right| \\ &\leq S_\infty \sum_{i, j=1}^d \int \left| \partial_{x_i x_j}^2 \omega_{\varepsilon_1}(x - y) \right| \omega_{\tilde{\varepsilon}}(\tau - s', x - y') |\chi(\xi; u(\tau, y))| \, dy \, ds' \, dy' \, d\tau \, dx \, d\xi \\ &= S_\infty \sum_{i, j=1}^d \int \left| \partial_{x_i x_j}^2 \omega_{\varepsilon_1}(x - y) \right| |u(\tau, y)| \, dy \, d\tau \, dx \\ &\leq C \frac{t S_\infty}{\varepsilon_1^2} \|u\|_{L^\infty(0, T; L^1(\mathbf{R}^d))}. \end{aligned}$$

For  $u \in L^\infty(\mathbf{R}_+; BV(\mathbf{R}^d))$ , we can then improve this estimate. In this case, we may integrate by parts in (3.27) and employ (2.4) to obtain

$$\begin{aligned} & \left| \lim_{\varepsilon_0 \downarrow 0} \int_0^t \mathcal{E}^{A-B}(\tau; \varepsilon, \tilde{\varepsilon}) \, d\tau \right| \\ &= \left| \sum_{i, j=1}^d \int \partial_{x_j} \omega_{\varepsilon_1}(x - y) \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \right. \\ & \quad \left. \times s_{ij}(u(\tau, y)) \partial_{y_i} u(\tau, y) \chi(u(\tau, y); v(s', y')) \, dy \, ds' \, dy' \, d\tau \, dx \, d\xi \right| \\ &\leq S_\infty \sum_{i, j=1}^d \int \left| \partial_{x_j} \omega_{\varepsilon_1}(x - y) \right| \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \\ & \quad \times |\partial_{y_i} u(\tau, y)| |\chi(u(\tau, y); v(s', y'))| \, dy \, ds' \, dy' \, d\tau \, dx \\ &\leq S_\infty \sum_{i, j=1}^d \int \left| \partial_{x_j} \omega_{\varepsilon_1}(x - y) \right| |\partial_{y_i} u(\tau, y)| \, dy \, d\tau \, dx \\ &\leq C \|u\|_{L^\infty(0, T; BV(\mathbf{R}^d))} \frac{t S_\infty}{\varepsilon_1}. \end{aligned}$$

Finally, we estimate

$$\begin{aligned} \left| \int_0^t \mathcal{E}_v(\tau; \tilde{\varepsilon}) \, d\tau \right| &\leq \int \omega_{\tilde{\varepsilon}}(\tau - s, x - y) \\ &\times \left( \left| \partial_\tau^{J_0} \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \right| \sup_\xi |e_0(\xi, s', y')| \right. \\ &\quad \left. + \sum_{\substack{J=(J_1, \dots, J_d) \geq 0 \\ |J| \leq J_*}} \left| D_x^J \omega_{\tilde{\varepsilon}}(\tau - s', x - y') \right| \sup_\xi |e_1^J(\xi, s', y')| \right) \, ds \, dy \, ds' \, dy' \, d\tau \, dx \\ &\leq \frac{C}{\tilde{\varepsilon}_0^{J_0}} \left\| \sup_\xi |e_0(\xi, \cdot, \cdot)| \right\|_{L^1(0, t + \tilde{\varepsilon}_0; L^1(\mathbf{R}^d))} \\ &\quad + C \sum_{\substack{J=(J_1, \dots, J_d) \geq 0 \\ |J| \leq J_*}} \frac{1}{\tilde{\varepsilon}_1^{|J|}} \left\| \sup_\xi |e_1^J(\xi, \cdot, \cdot)| \right\|_{L^1(0, t + \tilde{\varepsilon}_0; L^1(\mathbf{R}^d))}. \end{aligned}$$

This concludes the proof of (3.3). If  $g \equiv f$  and  $B \equiv A$ , then  $\mathcal{E}_t^{f-g}, \mathcal{E}_t^{A-B} \equiv 0$ . Consequently, we can let  $\varepsilon_1 \downarrow 0$ , and hence  $\mathcal{E}_{u,t}^x \rightarrow 0$  in (3.3). This completes the proof of (3.4).  $\square$

4. ESTIMATES FOR CONTINUOUS DEPENDENCE ON THE NONLINEARITIES

We now apply the  $L^1$ -framework developed in Section 3 to derive an explicit estimate for continuous dependence on the nonlinearities in quasilinear degenerate parabolic equations with anisotropic diffusion. Consider the problem

$$(4.1) \quad \partial_t v + \operatorname{div} g(v) = \nabla \cdot (B(v) \nabla v), \quad v(0, x) = v_0(x),$$

where  $g, B$ , and  $v_0$  satisfy the conditions stated in Section 1. The kinetic formulation of (4.1) is (3.1) with  $E \equiv 0$ . For simplicity of presentation, we assume that  $u_0 \in BV(\mathbf{R}^d)$ .

We can now apply Theorem 3.1 with  $\mathcal{E}_{v,t} \equiv 0, \tilde{\varepsilon}_0, \tilde{\varepsilon}_1 \downarrow 0$  (so that  $\mathcal{E}_{v,t}^t, \mathcal{E}_{v,t}^x \rightarrow 0$ ) and  $\mathcal{E}_{u,t}^x(\varepsilon_1) \leq \varepsilon_1 \|u_0\|_{BV(\mathbf{R}^d)}$ . Hence, for any  $t > 0$  and any  $\varepsilon_1 > 0$ ,

$$\begin{aligned} \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} &\leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} \\ &\quad + C \|u_0\|_{BV(\mathbf{R}^d)} \left( \varepsilon_1 + t \|f' - g'\|_\infty + \frac{tS_\infty}{\varepsilon_1} \right). \end{aligned}$$

Choosing the optimal  $\varepsilon_1$ , we end up with the following theorem.

**Theorem 4.1** (continuous dependence estimate). *Suppose  $u_0 \in BV(\mathbf{R}^d)$ . Let  $u$  be an entropy solution of (1.1) with (1.2)–(1.4). Let  $v \in C(\mathbf{R}_+; L^1(\mathbf{R}^d))$  be an entropy solution of (4.1) with (1.6)–(1.8). Then, for any  $t > 0$ ,*

$$(4.2) \quad \begin{aligned} &\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ &\leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} \\ &\quad + C \|u_0\|_{BV(\mathbf{R}^d)} \left( t \|f' - g'\|_\infty + \sqrt{t} \sqrt{\left\| \left( \sqrt{A} - \sqrt{B} \right) \left( \sqrt{A} - \sqrt{B} \right)^\top \right\|_\infty} \right). \end{aligned}$$

*Remark 4.1.* Note that Theorem 4.1 holds without the assumption that  $v \in L^\infty(\mathbf{R}_+; BV(\mathbf{R}^d))$ . Also, observe that, in the isotropic case:  $A = \text{diag}(a_1, \dots, a_d)$  and  $B = \text{diag}(b_1, \dots, b_d)$ ,

$$\sqrt{\left\| \left( \sqrt{A} - \sqrt{B} \right) \left( \sqrt{A} - \sqrt{B} \right)^\top \right\|_\infty} = \sup_{\substack{\xi \in I(u_0) \\ i=1, \dots, d}} \left| \sqrt{a_i(\xi)} - \sqrt{b_i(\xi)} \right|.$$

We remark that, in the estimate (4.2) for continuous dependence, the strong norms for  $f - g$  can be replaced by weaker norms in the spirit of [3], although we do not pursue this here.

5. AN ERROR ESTIMATE FOR THE VANISHING ANISOTROPIC VISCOSITY APPROXIMATION

We now consider the anisotropic viscous problem (1.10) with  $f, A, B$ , and  $v_0$  satisfying (1.3), (1.4), (1.6), and (1.8), respectively. Suppose  $B(v) > 0$  (i.e., (1.10) is uniformly parabolic which admits a unique classical solution) and

$$B_\infty := \|B\|_\infty = \sup_{\substack{\xi \in I(u_0) \\ i,j=1, \dots, d}} b_{ij}(\xi) < \infty.$$

We are interested in applying the  $L^1$ -framework established in Section 3 to derive an explicit error estimate for  $u - v$  as  $\mu \downarrow 0$ , where  $u$  is an entropy solution of the original problem (1.1). As in the previous section, we assume that  $u_0 \in BV(\mathbf{R}^d)$ .

Let  $(\eta, q, R)$  be an entropy-entropy flux triple. Multiplying (1.10) by  $\eta'(v)$ , we recover the usual dissipation structure

$$\begin{aligned} \partial_t \eta(v) + \text{div} q(v) &= \nabla \cdot (R'(v) \nabla v) - \eta''(v) \sum_{i,j=1}^d a_{ij}(v) \partial_{x_i} v \partial_{x_j} v \\ &+ \mu \nabla \cdot (\eta'(v) B(v) \nabla v) - \mu \eta''(v) \sum_{i,j=1}^d b_{ij}(v) \partial_{x_i} v \partial_{x_j} v. \end{aligned} \tag{5.1}$$

We identify the entropy defect measure  $m^v(\xi, t, x)$  as

$$m^v(\xi, t, x) = \delta(\xi - v(t, x)) \sum_{k=1}^K \left( \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^B(v) \right)^2,$$

and also the entropy dissipation measure  $m^{v,\psi}(t, x)$  as

$$m^{v,\psi}(t, x) = \int_{\mathbf{R}} m^v(\xi, t, x) \psi(\xi) d\xi, \quad \psi \in C_0(\mathbf{R}),$$

via the duality  $(C_0(\mathbf{R}); \mathcal{M}(\mathbf{R}))$ .

The parabolic defect measure  $n^v(\xi, t, x)$  is identified as

$$n^v(\xi, t, x) = \delta(\xi - v(t, x)) \sum_{k=1}^K \left( \sum_{i=1}^d \partial_{x_i} \zeta_{ik}^A(v) \right)^2 \geq 0,$$

and again, via the duality, the parabolic dissipation measure  $n^{v,\psi}(t, x)$  as

$$n^{v,\psi}(t, x) = \int_{\mathbf{R}} n^v(\xi, t, x) \psi(\xi) d\xi, \quad \psi \in C_0(\mathbf{R}).$$

Hence we can write (5.1) as

$$(5.2) \quad \begin{aligned} & \partial_t \eta(v) + \operatorname{div} q(v) - \nabla \cdot (R'(u) \nabla v) \\ &= - \left( \mu m^{v, \eta''} + n^{v, \eta''} \right) + \mu \nabla \cdot (\eta'(v) B(v) \nabla v). \end{aligned}$$

We can now transform the dissipation structure (5.2) via the duality into the kinetic structure [9]:

$$(5.3) \quad \begin{aligned} & \partial_t \chi(\xi; v) + f'(\xi) \cdot \nabla_x \chi(\xi; v) \\ &= \sum_{i,j=1}^d a_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; v) + \partial_\xi (\mu m^v + n^v) + \mu \sum_{i,j=1}^d b_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; v). \end{aligned}$$

We first assume that  $v_0 \in BV(\mathbf{R}^d)$ . Then we can write

$$\begin{aligned} & \mu \sum_{i,j=1}^d b_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; v) \\ &= \partial_\xi \left( \mu \sum_{j=1}^d \partial_{x_j} \left( \sum_{i=1}^d \partial_{x_i} \left( \int^\xi b_{ij}(\eta) \chi(\eta; v) d\eta \right) \right) \right) =: \partial_\xi E(\xi, t, x), \end{aligned}$$

where

$$E(\xi, t, x) = \sum_{j=1}^d \partial_{x_j} e_1^j(\xi, t, x), \quad e_1^j(\xi, t, x) = \mu \sum_{i=1}^d \partial_{x_i} \left( \int^\xi b_{ij}(\eta) \chi(\eta; v) d\eta \right).$$

From (2.4), it is clear that, for any  $\xi$ ,

$$\left| e_1^j(\xi, t, x) \right| \leq \mu \sum_{i=1}^d b_{ij}(v) |\partial_{x_i} v| \leq \mu B_\infty \sum_{i=1}^d |\partial_{x_i} v|, \quad j = 1, \dots, d.$$

Hence

$$\left\| \sup_\xi \left| e_1^j(\xi, \cdot, \cdot) \right| \right\|_{L^1(0,t;L^1(\mathbf{R}^d))} \leq \mu t B_\infty \|v_0\|_{BV(\mathbf{R}^d)}.$$

Applying Theorem 3.1 with  $\varepsilon_0, \varepsilon_1, \tilde{\varepsilon}_0 \downarrow 0$  (so that  $\mathcal{E}_{u,t}^t, \mathcal{E}_{u,t}^x, \mathcal{E}_{v,t}^t \rightarrow 0$ ), and

$$\mathcal{E}_{v,t}^x(\tilde{\varepsilon}_1) \leq \tilde{\varepsilon}_1 \|v_0\|_{BV(\mathbf{R}^d)},$$

we obtain that, for any  $t > 0$  and  $\tilde{\varepsilon}_1 > 0$ ,

$$\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + C \|v_0\|_{BV(\mathbf{R}^d)} \left( \tilde{\varepsilon}_1 + \frac{t\mu}{\tilde{\varepsilon}_1} \right).$$

Choosing the optimal  $\tilde{\varepsilon}_1$ , we get a rate of convergence in  $\mu \downarrow 0$  that is of order  $\sqrt{t\mu}$ .

If  $v_0 \notin BV(\mathbf{R}^d)$ , then we must write the error term as

$$\mu \sum_{i,j=1}^d b_{ij}(\xi) \partial_{x_i x_j}^2 \chi(\xi; v) = \partial_\xi E(\xi, t, x),$$

where

$$E(\xi, t, x) = \sum_{i,j=1}^d \partial_{x_i x_j}^2 e_1^{ij}(\xi, t, x), \quad e_1^{ij}(\xi, t, x) = \mu \int^\xi b_{ij}(\eta) \chi(\eta; v) d\eta.$$

Clearly, for any  $\xi$ ,

$$\left| e_1^{ij}(\xi, t, x) \right| \leq B_\infty \mu |v(t, x)|, \quad i, j = 1, \dots, d.$$

Hence

$$\left\| \sup_{\xi} \left| e_1^{ij}(\xi, \cdot, \cdot) \right| \right\|_{L^1(0, t; L^1(\mathbf{R}^d))} \leq t B_\infty \mu \|v_0\|_{L^1(\mathbf{R}^d)},$$

so that, for any  $t > 0$  and  $\tilde{\varepsilon}_1 > 0$ ,

$$\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + C \|v_0\|_{L^1(\mathbf{R}^d)} \left( \tilde{\varepsilon}_1 + \frac{t\mu}{\tilde{\varepsilon}_1^2} \right).$$

Choosing the optimal  $\varepsilon_1$ , we get a rate of convergence in  $\mu \downarrow 0$  of order  $(t\mu)^{\frac{1}{3}}$ .

Observe that we do not get an optimal convergence rate when  $v_0 \notin BV(\mathbf{R}^d)$  in this way. However, as already mentioned in Section 1, one of our observations is that, by interpreting the desired error estimate as a continuous dependence estimate, we can obtain the optimal result. Indeed, in the present context, Theorem 4.1 gives

$$\begin{aligned} & \|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \\ & \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} \\ & \quad + C \|u_0\|_{BV(\mathbf{R}^d)} \sqrt{t} \sqrt{\left\| \left( \sqrt{A} - \sqrt{A + \mu B} \right) \left( \sqrt{A} - \sqrt{A + \mu B} \right)^\top \right\|_\infty}, \end{aligned}$$

for any  $t > 0$ . Now a simple calculation reveals

$$\left\| \left( \sqrt{A} - \sqrt{A + \mu B} \right) \left( \sqrt{A} - \sqrt{A + \mu B} \right)^\top \right\|_\infty \leq C\mu.$$

We summarize the discussion in this section in the following theorem.

**Theorem 5.1** (vanishing anisotropic viscosity). *Suppose  $u_0 \in BV(\mathbf{R}^d)$ , and let  $u$  be an entropy solution of (1.1) with (1.2)–(1.4). Let  $v \in C(\mathbf{R}_+; L^1(\mathbf{R}^d))$  be an entropy solution of (1.10) with (1.3), (1.6), and (1.8). Then, for any  $t > 0$ ,*

$$\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbf{R}^d)} + C \|u_0\|_{BV(\mathbf{R}^d)} \sqrt{t\mu}.$$

### 6. ERROR ESTIMATES FOR A FINITE DIFFERENCE APPROXIMATION

In this section, as an example of a direct application of the  $L^1$ -framework to make error estimates for numerical methods, we focus on a linear convection-diffusion model equation

$$(6.1) \quad \partial_t u + \operatorname{div}(Vu) = \nabla \cdot (A \nabla u), \quad u(0, x) = u_0(x),$$

for some constant velocity vector  $V = (V_1, \dots, V_d) > 0$  and some small constant diffusion matrix  $A = \operatorname{diag}(a_1, \dots, a_d) \geq 0$ . We assume that  $u_0 \in BV(\mathbf{R}^d)$ .

Fix a time step size  $\Delta t > 0$  and a spatial step size  $\Delta x > 0$ . We use  $\partial_{t, \Delta t}$  for the temporal difference operator

$$\partial_{t, \Delta t} v(t, x) = \frac{v(t + \Delta t, x) - v(t, x)}{\Delta t};$$

$\partial_{x_i, \Delta x}$  for the first order spatial difference operator in the direction  $x_i$ :

$$\partial_{x_i, \Delta x} v(t, x) = \frac{v(t, x) - v(t, x - \Delta x e_i)}{\Delta x}, \quad i = 1, \dots, d,$$

where  $e_i$  denotes the  $i$ th unit vector in  $\mathbf{R}^d$ ; and  $\partial_{x_i, \Delta x}^2$  for the second order spatial central difference operator in the direction  $x_i$ :

$$\partial_{x_i, \Delta x}^2 v(t, x) = \frac{v(t, x - \Delta x e_i) - 2v(t, x) + v(t, x + \Delta x e_i)}{(\Delta x)^2}, \quad i = 1, \dots, d.$$

We consider the explicit upwind-central finite difference scheme:

$$(6.2) \quad \partial_{t, \Delta t} v(t, x) + \sum_{i=1}^d V_i \partial_{x_i, \Delta x} v(t, x) = \sum_{i=1}^d a_i \partial_{x_i, \Delta x}^2 v(t, x), \quad v(0, x) = u_0(x).$$

As usual, to ensure the stability, it is necessary to require the CFL condition:

$$(6.3) \quad V_\infty d \frac{\Delta t}{\Delta x} + 2A_\infty d \frac{\Delta t}{(\Delta x)^2} \leq 1, \quad V_\infty := \max_{i=1, \dots, d} V_i, \quad A_\infty := \max_{i=1, \dots, d} a_i.$$

From the CFL condition (6.3) and  $u_0 \in BV$ , it follows in a standard fashion (see, e.g., [13], [17]) that  $\|v(t, \cdot)\|_{L^\infty(\mathbf{R}^d)} \leq \|u_0\|_{L^\infty(\mathbf{R}^d)}$  and  $\|v(t, \cdot)\|_{BV(\mathbf{R}^d)} \leq \|u_0\|_{BV(\mathbf{R}^d)}$  for all  $t > 0$ , and  $\|v(t_2, \cdot) - v(t_1, \cdot)\|_{L^1(\mathbf{R}^d)} \leq C\sqrt{|t_2 - t_1|}$  for all  $t_1, t_2 > 0$ .

Our goal is to derive an  $L^1$  error estimate that is uniform with respect to the small diffusion matrix  $A$ . For technical reasons, we are not going to work directly with  $v$ , but instead with a regularized version  $v_\rho$  defined by

$$v_\rho(t, x) = \left( \omega_\rho \star_{(t,x)} v \right) (t, x), \quad \rho = (\rho_0, \rho_1) > 0,$$

where  $\omega_\rho = \omega_\rho(t, x)$  is a standard regularization kernel of the type used in the proof of Theorem 3.1 with the smoothing radius  $\rho_0$  in  $t$  and  $\rho_1$  in  $x$ .

Evaluating (6.2) at  $(t - s, x - y)$ , then multiplying by  $\omega_\rho(s, y)$ , and finally integrating the result over  $(s, y)$ , we obtain that the “smooth” function  $v_\rho$  satisfies the finite difference equation

$$(6.4) \quad \partial_{t, \Delta t} v_\rho(t, x) + \sum_{i=1}^d V_i \partial_{x_i, \Delta x} v_\rho(t, x) = \sum_{i=1}^d a_i \partial_{x_i, \Delta x}^2 v_\rho(t, x)$$

with initial data

$$v_\rho(0, x) = \left( \omega_\rho \star_x u_0 \right) (x).$$

Clearly, the approximate solution  $v_\rho$  satisfies the following apriori estimates, uniform in  $\Delta t, \Delta x, \rho$ :

$$(6.5) \quad \begin{cases} \|v_\rho(t, \cdot)\|_{L^\infty(\mathbf{R}^d)} \leq \|u_0\|_{L^\infty(\mathbf{R}^d)}, \\ \|v_\rho(t, \cdot)\|_{BV(\mathbf{R}^d)} \leq \|u_0\|_{BV(\mathbf{R}^d)}, \\ \|v_\rho(t_2, \cdot) - v_\rho(t_1, \cdot)\|_{L^1(\mathbf{R}^d)} \leq C\sqrt{|t_2 - t_1|}, \quad \text{for any } t_1, t_2 > 0. \end{cases}$$

A Taylor expansion yields

$$\partial_{t, \Delta t} v_\rho(t, x) = \partial_t v_\rho(t, x) + \bar{E}_0(t, x),$$

with

$$\bar{E}_0(t, x) = \frac{1}{2} \int_0^{\Delta t} \partial_t^2 v_\rho(t + z, x) dz.$$

Observe that, for any  $t > 0$ ,  $\|\bar{E}_0\|_{L^1(0, t; L^1(\mathbf{R}^d))} \leq C \frac{t \Delta t}{\rho_0^{3/2}}$ , where the third part of (6.5) was used.

Similarly,

$$\sum_{i=1}^d V_i \partial_{x_i, \Delta x} v_\rho(t, x) = \sum_{i=1}^d V_i \partial_{x_i} v_\rho(t, x) + \bar{E}_1(t, x),$$

with

$$\bar{E}_1(t, x) = \sum_{i=1}^d \frac{V_i}{2} \int_{-\Delta x}^0 \partial_{x_i}^2 v_\rho(t, x + ze_i) dz,$$

and, for any  $t > 0$ ,  $\|\bar{E}_1\|_{L^1(0,t;L^1(\mathbf{R}^d))} \leq C \frac{t\Delta x}{\rho_1}$ , in which the second part of (6.5) has been used.

Finally,

$$\sum_{i=1}^d a_i \partial_{x_i, \Delta x}^2 v_\rho(t, x) = \sum_{i=1}^d a_i \partial_{x_i}^2 v_\rho(t, x) + \bar{E}_2(t, x),$$

with

$$\bar{E}_2(t, x) = \sum_{i=1}^d \frac{a_i \Delta x}{24} \int_{-\Delta x}^{\Delta x} \partial_{x_i}^4 v_\rho(t, x + ze_i) dz,$$

and, for any  $t > 0$ ,  $\|\bar{E}_2\|_{L^1(0,t;L^1(\mathbf{R}^d))} \leq C \frac{t\Delta x^2}{\rho_1^2}$ , in which we have used (6.5) again.

Hence, from (6.4), it follows that  $v_\rho$  satisfies the “approximate” convection-diffusion equation

$$(6.6) \quad \partial_t v_\rho + \operatorname{div}(V v_\rho) = \nabla \cdot (A \nabla v_\rho) + \bar{E}(t, x),$$

where  $\bar{E}(t, x) := \bar{E}_0(t, x) + \bar{E}_1(t, x) + \bar{E}_2(t, x)$ , which suggests that we may apply Theorem 3.1 with  $J_0 = J_1 = 0$  to estimate  $u - v_\rho$ .

Let  $\eta : \mathbf{R} \rightarrow \mathbf{R}$  be an entropy function. Multiplying (6.6) by  $\eta'(v_\rho)$ , we obtain the usual dissipation structure

$$\partial_t \eta(v_\rho) + \operatorname{div}(V \eta(v_\rho)) - \nabla \cdot (A \nabla \eta(v_\rho)) = -\eta''(v_\rho) \left( \sum_{i=1}^d \sqrt{a_i} \partial_{x_i} v_\rho \right)^2 + \eta'(v_\rho) \bar{E}(t, x).$$

As usual, we can transform this dissipation structure via the duality into the kinetic structure:

$$\partial_t \chi(\xi; v_\rho) + V \cdot \nabla_x \chi(\xi; v_\rho) = \sum_{i=1}^d a_i \partial_{x_i}^2 \chi(\xi; v_\rho) + \partial_\xi (m^v + n^v + E)(\xi, t, x),$$

where  $m^v \equiv 0$  and

$$n^v = \delta(\xi - v_\rho) \left( \sum_{i=1}^d \sqrt{a_i} \partial_{x_i} v_\rho \right)^2, \quad E(\xi, t, x) = \mathbf{1}_{\xi > v_\rho} \bar{E}(t, x),$$

so that  $\partial_\xi E = \delta(\xi - v_\rho) \bar{E}$ .

Observe that, for any  $t > 0$ ,

$$\left\| \sup_{\xi} |E(\xi, \cdot, \cdot)| \right\|_{L^1(0,t;L^1(\mathbf{R}^d))} \leq Ct \left( \frac{\Delta t}{\rho^{3/2}} + \frac{\Delta x}{\rho} + \frac{\Delta x^2}{\rho^3} \right).$$

Hence, using (3.4) in Theorem 3.1 with  $J_0 = J_1 = 0$  (after having sent  $\tilde{\varepsilon}_0, \tilde{\varepsilon}_1 \downarrow 0$ ), we get

$$\|u(t, \cdot) - v_\rho(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq \|u_0 - v_\rho(0, \cdot)\|_{L^1(\mathbf{R}^d)} + Ct \left( \frac{\Delta t}{\rho_0^{3/2}} + \frac{\Delta x}{\rho_1} + \frac{\Delta x^2}{\rho_1^3} \right),$$

for any  $t > 0$  and  $\rho > 0$ . From (6.5), we have

$$\|u_0 - v_\rho(0, \cdot)\|_{L^1(\mathbf{R}^d)}, \|v(t, \cdot) - v_\rho(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq C(\sqrt{\rho_0} + \rho_1).$$

Hence

$$\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq C \left( \sqrt{\rho_0} + \rho_1 + \frac{t\Delta t}{\rho_0^{3/2}} + \frac{t\Delta x}{\rho_1} + \frac{t\Delta x^2}{\rho_1^3} \right), \quad t > 0.$$

Choosing the optimal  $\rho$ , we get the following theorem regarding the convergence rate for the upwind-central finite difference scheme.

**Theorem 6.1.** *Let  $u$  be an entropy solution of (6.1) with (1.2) and  $u_0 \in BV(\mathbf{R}^d)$ . Let  $v = v(t, x)$  be the upwind-central finite difference solution generated by (6.2) with (6.3). Then, for any  $t > 0$ ,*

$$\|u(t, \cdot) - v(t, \cdot)\|_{L^1(\mathbf{R}^d)} \leq C \left( (t\Delta t)^{\frac{1}{4}} + \sqrt{t\Delta x} \right) \leq C\sqrt{t\Delta x}.$$

*Remark 6.1.* Note that the  $L^1$  error estimate in Theorem 6.1 is robust with respect to sending the diffusion matrix  $A$  to zero. We emphasize that, although the problem (6.1) under consideration is linear, our  $L^1$  method of analysis is still very much nonlinear! We will further develop our approach to analyze and derive  $L^1$  error estimates for monotone finite difference schemes for nonlinear degenerate parabolic equations elsewhere.

#### ACKNOWLEDGMENTS

The research of Gui-Qiang Chen was supported in part by the National Science Foundation under Grants DMS-0204225, DMS-0204455, and INT-9987378, and by an Alexander von Humboldt Foundation Fellowship. The research of Kenneth H. Karlsen was supported in part by the BeMatA program of the Research Council of Norway and the European network HYKE, funded by the EC as contract HPRN-CT-2002-00282. Finally, we are grateful to B. Perthame for his comments regarding [23].

#### REFERENCES

- [1] D. Aregba-Driollet, R. Natalini, and S. Tang, Explicit diffusive kinetic schemes for nonlinear degenerate parabolic systems, *Math. Comp.* **73** (2004), 63–94. MR2034111
- [2] P. Bénilan and H. Touré, Sur l'équation générale  $u_t = a(\cdot, u, \phi(\cdot, u)_x)_x + v$  dans  $L^1$ . II. Le problème d'évolution, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **12** (1995), 727–761. MR1360542 (96m:35135)
- [3] F. Bouchut and B. Perthame, Kružkov's estimates for scalar conservation laws revisited, *Trans. Amer. Math. Soc.* **350** (1998), 2847–2870. MR1475677 (98m:65156)
- [4] F. Bouchut, F. R. Guarguaglini, and R. Natalini, Diffusive BGK approximations for nonlinear multidimensional parabolic equations, *Indiana Univ. Math. J.* **49** (2000), 723–749. MR1793689 (2001k:35162)
- [5] R. Bürger, S. Evje, and K. H. Karlsen, On strongly degenerate convection-diffusion problems modeling sedimentation-consolidation processes, *J. Math. Anal. Appl.* **247** (2000), 517–556. MR1769093 (2001d:35110)

- [6] M. C. Bustos, F. Concha, R. Bürger, and E. M. Tory, *Sedimentation and Thickening: Phenomenological Foundation and Mathematical Theory*, Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999. MR1747460 (2002b:76101)
- [7] J. Carrillo, Entropy solutions for nonlinear degenerate problems, *Arch. Rational Mech. Anal.* **147** (1999), 269–361. MR1709116 (2000m:35132)
- [8] G.-Q. Chen and E. DiBenedetto, Stability of entropy solutions to the Cauchy problem for a class of hyperbolic-parabolic equations, *SIAM J. Math. Anal.* **33** (2001), 751–762. MR1884720 (2002m:35125)
- [9] G.-Q. Chen and B. Perthame, Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **20** (2003), 645–668. MR1981403 (2004c:35235)
- [10] B. Cockburn and P.-A. Gremaud, A priori estimates for numerical methods for scalar conservation laws, Part I: The general approach, *Math. Comp.* **65** (1996), 533–573. MR1333308 (96g:65089)
- [11] B. Cockburn and G. Gripenberg, Continuous dependence on the nonlinearities of solutions of degenerate parabolic equations, *J. Diff. Eqs.* **151** (1999), 231–251. MR1669570 (2000a:35101)
- [12] M. S. Espedal and K. H. Karlsen, Numerical solution of reservoir flow models based on large time step operator splitting algorithms, In: *Filtration in Porous Media and Industrial Applications (Cetraro, Italy, 1998), Lecture Notes in Mathematics*, **1734** (2000), pp. 9–77, Springer-Verlag: Berlin. MR1816143 (2002a:76115)
- [13] S. Evje and K. H. Karlsen, Monotone difference approximations of  $BV$  solutions to degenerate convection-diffusion equations, *SIAM J. Numer. Anal.* **37** (2000), 1838–1860. MR1766850 (2001g:65110)
- [14] S. Evje, K. H. Karlsen, and N. H. Risebro, A continuous dependence result for nonlinear degenerate parabolic equations with spatially dependent flux function, In: *Int. Series of Numerical Mathematics* **140** (2001), 337–346, Birkhäuser-Verlag. MR1882934
- [15] R. Eymard, T. Gallouët, R. Herbin, and A. Michel, Convergence of a finite volume scheme for nonlinear degenerate parabolic equations, *Numer. Math.* **92** (2002), 41–82. MR1917365 (2003e:65159)
- [16] K. H. Karlsen and M. Ohlberger, A note on the uniqueness of entropy solutions of nonlinear degenerate parabolic equations, *J. Math. Anal. Appl.* **275** (2002), 439–458. MR1941794 (2004a:35137)
- [17] K. H. Karlsen and N. H. Risebro, Convergence of finite difference schemes for viscous and inviscid conservation laws with rough coefficients, *M2AN Math. Model. Numer. Anal.* **35** (2001), 239–269. MR1825698 (2002b:35138)
- [18] K. H. Karlsen and N. H. Risebro, On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients, *Discrete and Continuous Dynamical Systems* **9** (2003), 1081–1104. MR1974417
- [19] S. N. Kružkov, First order quasi-linear equations in several independent variables, *Math. USSR Sbornik* **10** (1970), 217–243.
- [20] N. N. Kuznetsov, Accuracy of some approximative methods for computing the weak solutions of a first-order quasi-linear equation, *USSR Comput. Math. and Math. Phys. Dokl.* **16** (1976), 105–119.
- [21] P.-L. Lions, B. Perthame, and E. Tadmor, A kinetic formulation of multidimensional scalar conservation laws and related equations, *J. Amer. Math. Soc.* **7** (1994), 169–191. MR1201239 (94d:35100)
- [22] B. J. Lucier, A moving mesh numerical method for hyperbolic conservation laws, *Math. Comp.* **46** (1986), 59–69. MR0815831 (87m:65141)
- [23] C. Makridakis and B. Perthame, Optimal rate of convergence for anisotropic vanishing viscosity limit of a scalar balance law, *SIAM J. Math. Anal.* **34** (2003), 1300–1307. MR2000971
- [24] C. Mascia, A. Porretta, and A. Terracina, Nonhomogeneous Dirichlet problems for degenerate parabolic-hyperbolic equations, *Arch. Rational Mech. Anal.* **163** (2002), 87–124. MR1911095 (2003e:35217)
- [25] A. Michel and J. Vovelle, A finite volume method for parabolic degenerate problems with general Dirichlet boundary conditions, Preprint, 2002.

- [26] M. Ohlberger, A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations, *M2AN Math. Model. Numer. Anal.* **35** (2001), 355–387. MR1825703 (2002a:65142)
- [27] B. Perthame, Uniqueness and error estimates in first order quasilinear conservation laws via the kinetic entropy defect measure, *J. Math. Pures Appl.* **77** (1998), 1055–1064. MR1661021 (2000e:35141)
- [28] É. Rouvre and G. Gagneux, Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées, *C. R. Acad. Sci. Paris Sér. I Math.* **329** (1999), 599–602. MR1717116 (2000g:35160)
- [29] T. Tassa. Uniqueness of piecewise smooth weak solutions of multidimensional degenerate parabolic equations, *J. Math. Anal. Appl.* **210** (1987), 598–608. MR1453194 (98c:35080)
- [30] A. I. Vol’pert, The spaces BV and quasi-linear equations. *Math. USSR Sbornik* **2** (1967), 225–267. MR0216338 (35:7172)
- [31] A. I. Vol’pert and S. I. Hudjaev, Cauchy’s problem for degenerate second order quasilinear parabolic equations, *Math. USSR Sbornik* **7** (1969), 365–387. MR0264232 (41:8828)
- [32] Z. Wu and J. Yin, Some properties of functions in  $BV_x$  and their applications to the uniqueness of solutions for degenerate quasilinear parabolic equations, *Chinese Northeastern Math. J.* **5** (1989), 395–422. MR1053519 (91j:35153)

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, 2033 SHERIDAN ROAD,  
EVANSTON, ILLINOIS 60208-2730

*E-mail address:* [gqchen@math.northwestern.edu](mailto:gqchen@math.northwestern.edu)

*URL:* <http://www.math.northwestern.edu/~gqchen/>

CENTRE OF MATHEMATICS FOR APPLICATIONS, DEPARTMENT OF MATHEMATICS, UNIVERSITY  
OF OSLO, P.O. BOX 1053, BLINDERN, N-0316 OSLO, NORWAY

*E-mail address:* [kennethk@math.uio.no](mailto:kennethk@math.uio.no)

*URL:* <http://www.math.uio.no/~kennethk/>