

NUMERICAL APPROXIMATION OF
YOUNG-MEASURE SOLUTIONS TO PARABOLIC
SYSTEMS OF FORWARD-BACKWARD TYPE

Miles Caddick and Endre Süli

This paper is concerned with the proof of existence and numerical approximation of large-data global-in-time Young measure solutions to initial-boundary-value problems for multidimensional nonlinear parabolic systems of forward-backward type of the form $\partial_t u - \operatorname{div}(a(Du)) + Bu = F$, where $B \in \mathbb{R}^{m \times m}$, $Bv \cdot v \geq 0$ for all $v \in \mathbb{R}^m$, F is an m -component vector-function defined on a bounded open Lipschitz domain $\Omega \subset \mathbb{R}^n$, and a is a locally Lipschitz mapping of the form $a(A) = K(A)A$, where $K : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$. The function a may have unequal lower and upper growth rates; it is not assumed to be monotone, nor is it assumed to be the gradient of a potential. We construct a numerical method for the approximate solution of problems in this class, and we prove its convergence to a Young measure solution of the system.

*Dedicated to Academician Professor Gradimir Milovanović
on the occasion of his 70th birthday.*

1. INTRODUCTION

The paper is concerned with the numerical approximation of Young measure solutions to initial-boundary-value problems for nonlinear multidimensional parabolic systems of forward-backward type, where the existence of a weak solution cannot in general be guaranteed, because the nonlinearity in the equation is neither monotone, nor globally Lipschitz, nor indeed is it the gradient of a potential. Nonlinear parabolic systems of this type arise in certain mathematical models of the atmospheric boundary layer and, to date, there have been no attempts at

2010 Mathematics Subject Classification. 35K51, 35K55, 76M10.

Keywords and Phrases. Young-measure, forward-backward parabolic PDE, finite element method.

the rigorous mathematical analysis of numerical methods for their approximate solution.

The systems of nonlinear parabolic partial differential equations considered here are of the following form:

$$(1) \quad \partial_t u - \operatorname{div}(a(Du)) + Bu = F,$$

with $u : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}^m$, where $\Omega \subset \mathbb{R}^n$ is a bounded open Lipschitz domain. Here Du denotes the spatial gradient of u , with $(Du)_{ik} = \partial u_i / \partial x_k$, $i = 1, \dots, m$, $k = 1, \dots, n$. Throughout we will take real numbers p_i , for $i = 1, \dots, m$, that satisfy

$$p_i > \max \left\{ 1, \frac{2n}{n+2} \right\}.$$

We define

$$p := \min_{i=1, \dots, m} p_i \quad \text{and} \quad q := \max_{i=1, \dots, m} p_i,$$

and require that

$$(2) \quad q - p < 1.$$

We shall also assume that

$$(3) \quad F \in L^{p'}(Q_T; \mathbb{R}^m) \cap L^2(Q_T; \mathbb{R}^m),$$

with $\frac{1}{p} + \frac{1}{p'} = 1$, where $Q_T := (0, T) \times \Omega$, and that

$$(4) \quad Bv \cdot v \geq 0 \quad \forall v \in \mathbb{R}^m,$$

where $B \in \mathbb{R}^{m \times m}$ is a constant matrix. All of our results extend directly to the case where $B \in L^\infty(Q_T; \mathbb{R}^{m \times m})$, without assuming (4), however a more general supposition of this kind would contribute no new insight, so for the sake of clarity of the exposition we shall continue to assume that $B \in \mathbb{R}^{m \times m}$ is a constant matrix satisfying (4). The system (1) is supplemented with the initial condition

$$(5) \quad u(0, \cdot) = u_0(\cdot) \in L^2(\Omega; \mathbb{R}^m),$$

and the homogeneous Dirichlet boundary condition

$$(6) \quad u|_{(0, T] \times \partial\Omega} = 0.$$

The function $a : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ is assumed to be a locally Lipschitz mapping of the form

$$(7) \quad a(A) = K(A)A,$$

where $K : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is continuous, and there exist positive constants c_0 and c_1 , with $c_0 \leq c_1$, such that

$$(8) \quad c_0 \sum_{i=1}^m (\mu_i^2 + |A_i|^2)^{\frac{p_i-2}{2}} \leq K(A) \leq c_1 \sum_{i=1}^m (\mu_i^2 + |A_i|^2)^{\frac{p_i-2}{2}} \quad \forall A \in \mathbb{R}^{m \times n},$$

with $p_i > \max\{1, \frac{2n}{n+2}\}$ for all $i = 1, \dots, m$, and μ_i is a constant defined by

$$(9) \quad \mu_i \begin{cases} \in \mathbb{R} & \text{if } p_i \geq 2, \\ \neq 0 & \text{if } 1 < p_i < 2, \end{cases}$$

and where $|A_i|$ denotes the Euclidean norm of the i -th row of the matrix A . More generally, $|\cdot|$ will signify the absolute value of a real number, the Euclidean norm of a vector, or the Frobenius norm of a matrix: its precise meaning will be clear from the context. We emphasize that we are making no assumptions here about the monotonicity of the mapping a , nor will we assume that a is the gradient of a certain C^1 -function. The results of the paper can be easily adapted to systems in which the form of a is slightly more general: for example, instead of (7)–(9) one can assume that $|a(x)| \leq c_0(1 + |x|^{q-1})$ and $a(x) \cdot x \geq c_1|x|^p$, with $p > \max\{1, \frac{2n}{n+2}\}$ and $p \leq q < p + 1$.

In order to motivate the weak structural assumptions on the function a stated above we shall now present some examples.

Example 1. In [2] the authors consider a system of equations modelling the behaviour of the atmospheric boundary layer, which served as our original motivation for the present study. In that model, $n = 1$, $m = 3$, $\Omega = (0, 1)$, and the function K , referred to as a stability function, is defined as follows:

$$(10) \quad K(\partial_x u) = \begin{cases} \frac{(|\partial_x u_1|^2 + |\partial_x u_2|^2)^{\frac{3}{2}}}{|\partial_x u_1|^2 + |\partial_x u_2|^2 + \partial_x u_3} & \text{if } \partial_x u_3 > 0, \\ \sqrt{|\partial_x u_1|^2 + |\partial_x u_2|^2} - \partial_x u_3 & \text{if } \partial_x u_3 \leq 0. \end{cases}$$

The resulting system of partial differential equations for the vector of unknowns $u(t, x) = (u_1(t, x), u_2(t, x), u_3(t, x))^T$ is

$$\begin{aligned} \partial_t u - \partial_x(K(\partial_x u)\partial_x u) + Bu &= F & \text{for } (t, x) \in (0, T] \times (0, 1), \\ u(t, 1) &= f(t) & \text{for } t \in (0, T], \\ u_1(t, 0) &= g_1(t) & \text{for } t \in (0, T], \\ u_2(t, 0) &= g_2(t) & \text{for } t \in (0, T], \\ \partial_x u_3(t, 0) &= h(t) & \text{for } t \in (0, T], \end{aligned}$$

subject to the initial condition

$$u(0, x) = u_0(x) \quad \text{for } x \in (0, 1).$$

Here, $u : [0, T] \times [0, 1] \rightarrow \mathbb{R}^3$ is a vector with components $u = (u_1, u_2, u_3)^T$, B is the skew-symmetric matrix

$$B = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and F is a forcing term given by the constant vector $F(t, x) = (-V, U, 0)^T$.

What is immediately seen is that, given the above definition of the function K , the product $K(\partial_x u)\partial_x u$ lacks a coercivity estimate in a Sobolev space on the range $\partial_x u_3 > 0$, because the mapping $\xi \in \mathbb{R}^3 \mapsto a(\xi) := K(\xi)\xi \in \mathbb{R}^3$ is not coercive in the range $\xi_3 > 0$; indeed, for $\xi = (1, 0, \xi_3)^\top$ with $\xi_3 > 0$, $a(\xi) \cdot \xi/|\xi| \rightarrow 1$ as $\xi_3 \rightarrow +\infty$. Another difficulty arises from the fact that the function $\xi \mapsto a(\xi) := K(\xi)\xi$ is not monotone on \mathbb{R}^3 , i.e., it is not true that $(a(\xi) - a(\eta)) \cdot (\xi - \eta) \geq 0$ for all $\xi, \eta \in \mathbb{R}^3$, and therefore monotone operator theory cannot be applied in this setting. To see that the mapping $\xi \mapsto a(\xi) := K(\xi)\xi$ is not monotone on \mathbb{R}^3 , take $\xi = (0.035, 0, -0.01)^\top$ and $\eta = (0.05, 0, 0)^\top$; then, computing $(a(\xi) - a(\eta)) \cdot (\xi - \eta)$ will give a negative number. Similarly taking $\xi = (-0.2, -0.1, 0.2)^\top$ and $\eta = (-0.1, 0, 0.5)^\top$ will result in a negative number. Finally to show that the vector field a is not a potential, consider the contour integral of a around the circle C parametrized by the function $r(\theta) = (1 + \cos(\theta), 0, \sin(\theta))$, $\theta \in [0, 2\pi)$. One can show that this is nonzero, and so a is not a conservative vector-field, and thus it is not a potential.

Example 2. For $n = 1$ and $m = 3$ consider the function K given by $\xi \in \mathbb{R}^3 \mapsto K(\xi) = \sqrt{\xi_1^2 + \xi_2^2 + |\xi_3|}$. Then we are in the regime $\partial_x u_3 \leq 0$ from (10). One has the existence of constants c_0 and c_1 such that

$$c_0(|\xi_1| + |\xi_2| + \sqrt{|\xi_3|}) \leq K(\xi) \leq c_1(|\xi_1| + |\xi_2| + \sqrt{|\xi_3|})$$

through bounding and using the equivalence of norms on \mathbb{R}^3 . This function K is therefore covered by our assumptions. We are unable to say anything about the region from (10) where $\partial_x u_3 > 0$, as our analysis does not cover the case $p = 1$.

As will transpire from the discussion that follows, under the stated hypotheses on the function a we are unable to prove the existence of a weak solution to the problem (1)–(9) under consideration and have to weaken the notion of solution to be able to show its existence. We shall therefore, instead, show the existence of a Young measure solution in the sense of the next definition, and will then consider the numerical approximation of such Young measure solutions.

Definition 1. We say that a pair (u, ν) , where u is a function with

$$\begin{aligned} u &\in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)) \cap L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m)) \\ Du_i &\in L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^{m \times n})), \quad i = 1, \dots, m, \\ \partial_t u &\in L^{\hat{q}}(0, T; W^{-1, \hat{q}}(\Omega; \mathbb{R}^m)), \end{aligned}$$

and $(t, x) \in Q_T \mapsto \nu_{t,x}$ is an $\mathbb{R}^{m \times n}$ -valued Young measure, such that

$$\langle \nu, a \rangle \in L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n}),$$

is called a Young measure solution to the problem described by (1), (5) and (6) if

$$\int_0^T \langle \partial_t u, \phi \rangle_{(W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m), W^{-1, \hat{q}}(\Omega; \mathbb{R}^m))} + \int_\Omega \langle \nu_{t,x}, a \rangle : D\phi + Bu \cdot \phi \, dx \, dt = \int_0^T \int_\Omega F \cdot \phi \, dx \, dt$$

for all $\phi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$, where $\hat{q} := \max\{q, 2\}$, $\hat{q}' = \frac{\hat{q}}{\hat{q}-1}$, and the initial condition is satisfied in the sense that $u(0, x) = u_0(x)$ for almost every $x \in \Omega$.

Measure-valued solutions and Young measure solutions to parabolic partial differential equations have been studied by a number of authors. Frehse & Specovius-Neugebauer, in [12], showed the existence of Hölder continuous Young measure solutions to two-dimensional nonmonotone quasilinear problems exhibiting quadratic growth (that is, $p = q = n = 2$), under the assumption that their nonlinearity a was expressible as

$$(11) \quad a = \nabla\Phi$$

for some C^1 function Φ . Our terminology *Young measure solution* follows that of Frehse & Specovius-Neugebauer in [12]; in particular, we have consciously avoided referring to the solutions considered here as *measure-valued* solutions, as the function u , whose existence we prove, is still a real-valued function with spatial Sobolev regularity; however, instead of being a standard weak solution, the function u satisfies the PDE in the sense of gradient Young-measures. Young measure solutions to forward-backward parabolic problems were studied by Demoulini in [7]: under the assumptions that $m = 1$, $p = q = 2$ and that (11) is satisfied, a sequence of approximating solutions was constructed by means of minimizing an integral, which involves the convexification of Φ . Using this construction, a family of Young measures was generated displaying an independence property, from which it was possible to deduce a uniqueness result regarding the function u (not regarding the family of Young measures). From this, a weak-strong uniqueness result follows if the classical solution, which is assumed to exist, satisfies additional constraints on its gradient. More recently, in Thanh [20] and some of the references contained therein, in particular [21], possible regularizing terms are discussed for forward-backward parabolic equations, the inclusion of which allows for solutions with higher regularity to be obtained than those considered here; Thanh also obtained results concerning the long-time behaviour of certain Young measure solutions, and (in the case of an equation in one space dimension) the support of the Young measure appearing in the Young measure solutions. However, once again, that work makes the assumption that (11) holds, and assumes more restrictive growth conditions on a than those considered here. Even more recently, Kim & Yan in [16] have expanded their earlier work in [15] to show the existence of infinitely many Lipschitz solutions (and under further assumptions, the existence of a unique classical solution) for particular classes of forward-backward parabolic equations. Although their work allows one to consider stronger notions of solution than the Young measure solutions that we shall discuss here, many of their structural assumptions regarding the nonlinearities are incompatible with our setting. Furthermore, their focus is solely on scalar equations, whereas we are concerned with systems here.

From the computational point of view, Carstensen & Roubíček [5] considered the numerical approximation of Young measures arising in certain minimization problems in the calculus of variations. As we lack this variational structure ourselves (in this work we are not assuming (11)), neither of the methods in [7] or [5] will be suitable in our setting. In [3] the authors consider numerical experiments for particular forward-backward equations that have a regularizing term added, but only for small values of the regularizing parameter and do not consider numerical

schemes for the limiting case. We mention in passing that for the Keller–Segel system, which is a coupled parabolic-elliptic system of partial differential equations, a stochastic interacting particle approximation of global-in-time measure-valued solutions in two space dimensions was considered by Haškovec & Schmeiser in [14].

The aim of this work is to develop and analyze a numerical scheme for the approximation of Young measure solutions of parabolic systems of equations under minimal assumptions on the form taken by the nonlinearity a , and under no restriction on the dimension of the domain Ω . We make the point here that the scheme discussed in this article will not allow us to compute the Young measures appearing in the references [7, 12, 5, 20, 3] cited above as taking different approximating sequences to the same function will potentially result in different families of Young measures being generated. Furthermore, without a concept of uniqueness, there is no guarantee that the functions appearing in the Young measure solutions will even be the same.

The paper is structured as follows. In the next section we shall collect miscellaneous results that are used in the subsequent analysis. In Section 2.3 we will show the existence of a large-data global-in-time Young measure solution to the system (1)–(9) under consideration by performing a spatial finite element approximation and passing to the limit in the spatial discretization parameter. In Section 2.3, we shall also discretize with respect to the temporal variable and show the convergence of the fully discrete scheme to a Young measure solution of the problem. We note here, however, that the results from Section 2.3 and Section 2.3 only yield to us indirect information about the measure ν through its action on functions belonging to a particular function space. This is the motivation behind Section 2.3, in which, stimulated by recent contributions by Fjordholm et al. [11] and [10], in particular, we shall develop a numerical algorithm for the computation of Young measure solutions which allows for a more direct approximation of the Young measure ν , and will prove that the algorithm converges to a Young measure solution of the problem under consideration.

2. PRELIMINARY DEFINITIONS AND RESULTS

2.1 Function Spaces

We begin by defining, for $r > 0$, the function space

$$E_r := \left\{ g \in C(\mathbb{R}^{m \times n}; \mathbb{R}^m) : \lim_{|A| \rightarrow \infty} \frac{g(A)}{1 + |A|^r} \text{ exists} \right\},$$

which will be of use to us later. On this space we consider the norm

$$\|g\|_{E_r} = \sup_{A \in \mathbb{R}^{m \times n}} \frac{|g(A)|}{1 + |A|^r}.$$

Lemma 1. *The following statements hold true for $r > 0$:*

1. *The space E_r is separable;*
2. *The space E_r is metrizable.*

Proof. For the first property we refer to Kinderlehrer & Pedregal [17]. The second follows by taking the metric induced by the norm $\|\cdot\|_{E_r}$. \square

Remark 1. *We remark here that E_r is a subspace of the larger space*

$$\tilde{E}_r := \left\{ g \in C(\mathbb{R}^{m \times n}; \mathbb{R}^m) : \sup_{A \in \mathbb{R}^{m \times n}} \frac{|g(A)|}{1 + |A|^r} < \infty \right\},$$

which is not separable. For example, under the assumption (8), $a \in \tilde{E}_{q-1}$, and for any $\alpha \in (1, \frac{p}{q-1})$ we have that $0 < q - 1 < \frac{p}{\alpha} < p$, and therefore $a \in E_{\frac{p}{\alpha}}$.

As $C_0(\mathbb{R}^{m \times n}; \mathbb{R}^m) \subset E_r$ for all $r > 0$, we have the following nesting of dual spaces for all $r > 0$:

$$E'_r \subset \mathcal{M}(\mathbb{R}^{m \times n}; \mathbb{R}^m) = C_0(\mathbb{R}^{m \times n}; \mathbb{R}^m)'.$$

Since the dual of a normed space is a Banach space, we have that E'_r is a Banach space for all $r > 0$, as we can equip E_r with the norm induced by its metric. In particular, thanks to the Banach–Alaoglu Theorem, any bounded sequence in E'_r has a weakly-* convergent subsequence.

2.2 Young Measures

We recall here from [1] the so-called Fundamental Theorem for Young Measures, together with some related remarks.

Theorem 1. *Let $\Omega \subset \mathbb{R}^n$ be Lebesgue measurable, let $\mathfrak{K} \subset \mathbb{R}^m$ be closed and let $z_j : \Omega \rightarrow \mathbb{R}^m$, $j = 1, 2, \dots$, be a sequence of Lebesgue measurable functions satisfying*

$$\lim_{j \rightarrow \infty} |\{x \in \Omega : z_j(x) \notin U\}| = 0$$

for any open neighbourhood U of \mathfrak{K} in \mathbb{R}^m . Then, there exists a subsequence $\{z_{j_k}\}$ of $\{z_j\}$ and a family of positive measures on \mathbb{R}^m , $\{\nu_x\}_{x \in \Omega}$, depending measurably on $x \in \Omega$, such that:

1. $\|\nu_x\|_M := \int_{\mathbb{R}^m} d\nu_x \leq 1$ for almost every $x \in \Omega$;
2. $\text{supp } \nu_x \subset \mathfrak{K}$ for almost every $x \in \Omega$; and
3. $f(z_{j_k}) \xrightarrow{*} \langle \nu_x, f \rangle = \int_{\mathbb{R}^m} f(\lambda) d\nu_x(\lambda)$ in $L^\infty(\Omega)$ for each continuous function $f : \mathbb{R}^m \rightarrow \mathbb{R}$ satisfying $\lim_{|\lambda| \rightarrow \infty} f(\lambda) = 0$.

Suppose further that $\{z_{j_k}\}$ satisfies the following tightness condition:

$$\limsup_{s \rightarrow \infty} \sup_k |\{x \in \Omega \cap B_R(0) : |z_{j_k}(x)| \geq s\}| = 0,$$

for every $R > 0$. Then, we have that $\|\nu_x\|_M = 1$ for almost every $x \in \Omega$, and, for any measurable subset A of the set Ω ,

$$f(z_{j_k}) \rightharpoonup \langle \nu_x, f \rangle$$

in $L^1(A)$ for any continuous function $f : \mathbb{R}^m \rightarrow \mathbb{R}$ such that $\{f(z_{j_k})\}$ is sequentially weakly relatively compact in $L^1(A)$.

Remark 2. If Ω is bounded and the sequence $\{z_j\}$ is bounded in $L^p(\Omega; \mathbb{R}^m)$ for some $p \in (1, \infty)$, then we obtain from the theorem the existence of a family of probability measures $\{\nu_x\}_{x \in \Omega}$ and a subsequence z_{j_k} such that

$$f(z_{j_k}) \rightharpoonup \langle \nu, f \rangle$$

in $L^r(\Omega)$, where $f : \mathbb{R}^m \rightarrow \mathbb{R}$ is a continuous function satisfying

$$|f(\lambda)| \leq c(1 + |\lambda|^s)$$

with $1 < r < \frac{p}{s}$ and $s > 0$.

Remark 3. The assumption that

$$\lim_{j \rightarrow \infty} |\{x \in \Omega : z_j(x) \notin U\}| = 0$$

for any open neighbourhood U of \mathfrak{K} is only used to prove that (2) in Theorem 1 holds, and is not needed to obtain any of its other conclusions. The proof of this theorem from [1] is still valid if we take $\mathfrak{K} = \mathbb{R}^m$, only it then provides no information about the support of the Young measure ν . Therefore, if we are not concerned with the support of the Young measure ν we may always take $\mathfrak{K} = \mathbb{R}^m$ in order to apply this theorem.

2.3 Miscellaneous Results

We recall the following result from Strauss [19] (cf. also Lions & Magenes [18], Lemma 8.1, Ch. 3, Sec. 8.4).

Lemma 2. Suppose that X and Y are Banach spaces. Assume that the space X is reflexive and is continuously embedded in the space Y ; then,

$$L^\infty(0, T; X) \cap C_w([0, T]; Y) = C_w([0, T]; X),$$

where $C_w([0, T]; X)$ and $C_w([0, T]; Y)$ denote the spaces of weakly continuous functions from $[0, T]$ into X and Y , respectively.

We shall also require the following consequence of the Arzelà–Ascoli theorem.

Lemma 3. *Let $\{u^j\}$ be a sequence of functions with the following properties:*

1. $\{u^j\}$ is bounded in $L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$ and

$$u^j \overset{*}{\rightharpoonup} u \text{ in } L^\infty(0, T; L^2(\Omega; \mathbb{R}^m));$$

2. $\{\partial_t u^j\}$ is bounded in $L^{s'}(0, T; W^{-1, s'}(\Omega; \mathbb{R}^m))$ for some $s \in (1, \infty)$, and

$$\partial_t u^j \rightharpoonup \partial_t u \text{ in } L^{s'}(0, T; W^{-1, s'}(\Omega; \mathbb{R}^m)).$$

Then, $u^{j_k}, u \in C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$, and there is a subsequence $\{u^{j_k}\}$ such that

$$\int_{\Omega} u^{j_k}(t, x) \cdot w(x) \, dx \rightarrow \int_{\Omega} u(t, x) \cdot w(x) \, dx$$

uniformly in $C([0, T])$ for all $w \in L^2(\Omega; \mathbb{R}^m)$ as $k \rightarrow \infty$.

Proof. We shall apply the Arzelà–Ascoli theorem for sequences of uniformly bounded and equicontinuous functions. For $w \in L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m)$, we let $f^j : [0, T] \rightarrow \mathbb{R}$ be the family of functions defined by

$$f^j(t) = \int_{\Omega} u^j(t, x) \cdot w(x) \, dx.$$

Since $\partial_t u^j \in L^{s'}(0, T; W^{-1, s'}(\Omega; \mathbb{R}^m))$, also $u^j \in C([0, T]; W^{-1, s'}(\Omega; \mathbb{R}^m))$; whereby also $u^j \in C_w([0, T]; W^{-1, s'}(\Omega; \mathbb{R}^m))$. As $u^j \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$, it then follows from Lemma 2 with $X = L^2(\Omega; \mathbb{R}^m)$ and $Y = L^2(\Omega; \mathbb{R}^m) + W^{-1, s'}(\Omega; \mathbb{R}^m) = (L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m))'$, that $u^j \in C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$. We note here that the duality $(L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m))' = L^2(\Omega; \mathbb{R}^m) + W^{-1, s'}(\Omega; \mathbb{R}^m)$ is a consequence of the Duality Theorem (cf. Theorem 2.7.1 on p.32 in Bergh & Löfström [4]), because $L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m)$ is dense in both $L^2(\Omega; \mathbb{R}^m)$ and $W_0^{1, s}(\Omega; \mathbb{R}^m)$ (e.g., because $C_0^\infty(\Omega; \mathbb{R}^m)$ is dense in both $L^2(\Omega; \mathbb{R}^m)$ and $W_0^{1, s}(\Omega; \mathbb{R}^m)$). Hence, $f^j \in C([0, T])$. We further have that

$$|f^j(t)| \leq \int_{\Omega} |u^j(t, x)| |w(x)| \, dx \leq \|u^j\|_{L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))} \|w\|_{L^2(\Omega; \mathbb{R}^m)} \leq C \|w\|_{L^2(\Omega; \mathbb{R}^m)},$$

which gives the boundedness of the sequence $\{f^j\}$ in $C([0, T])$ for each w contained in $L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m)$. In order to verify equicontinuity of the sequence, note that, for all $w \in L^2(\Omega; \mathbb{R}^m) \cap W_0^{1, s}(\Omega; \mathbb{R}^m)$,

$$\begin{aligned} |f^j(t+h) - f^j(t)| &= \left| \int_{\Omega} (u^j(t+h, x) - u^j(t, x)) \cdot w(x) \, dx \right| \\ &= \left| \int_t^{t+h} \int_{\Omega} \partial_t u^j(s, x) \cdot w(x) \, dx \, ds \right| \\ &\leq \|\partial_t u^j\|_{L^{s'}(0, T; W^{-1, s'}(\Omega; \mathbb{R}^m))} |h|^{\frac{1}{s}} \|w\|_{W^{1, s}(\Omega; \mathbb{R}^m)} \\ &\leq C |h|^{\frac{1}{s}} \|w\|_{W^{1, s}(\Omega; \mathbb{R}^m)}. \end{aligned}$$

Thus by the Arzelà–Ascoli theorem there is a subsequence $\{f^{j_k}\}$, which converges uniformly as $k \rightarrow \infty$ to some function $f \in C([0, T])$. We note that as this convergence is uniform and the sequence f^{j_k} is uniformly bounded we have that

$$\lim_{k \rightarrow \infty} \int_0^T \varphi(t) \int_{\Omega} u^{j_k}(t, x) \cdot w(x) \, dx \, dt = \lim_{k \rightarrow \infty} \int_0^T \varphi(t) f^{j_k}(t) \, dt = \int_0^T \varphi(t) f(t) \, dt$$

for all $\varphi \in C^\infty([0, T])$. By the assumed weak convergence, we also have that

$$\lim_{k \rightarrow \infty} \int_0^T \varphi(t) \int_{\Omega} u^{j_k}(t, x) \cdot w(x) \, dx \, dt = \int_0^T \varphi(t) \int_{\Omega} u(t, x) \cdot w(x) \, dx \, dt$$

for all $\varphi \in C^\infty([0, T])$, allowing us to identify

$$f(t) = \int_{\Omega} u(t, x) \cdot w(x) \, dx \quad \forall t \in [0, T].$$

Thus we have shown that, for all $w \in L^2(\Omega; \mathbb{R}^m) \cap W_0^{1,s}(\Omega; \mathbb{R}^m)$,

$$\int_{\Omega} u^{j_k}(t, x) \cdot w(x) \, dx \rightarrow \int_{\Omega} u(t, x) \cdot w(x) \, dx, \quad \text{uniformly in } C([0, T]).$$

As $L^2(\Omega; \mathbb{R}^m) \cap W_0^{1,s}(\Omega; \mathbb{R}^m)$ is dense in $L^2(\Omega; \mathbb{R}^m)$, it then follows that, for all $w \in L^2(\Omega; \mathbb{R}^m)$,

$$\int_{\Omega} u^{j_k}(t, x) \cdot w(x) \, dx \rightarrow \int_{\Omega} u(t, x) \cdot w(x) \, dx, \quad \text{uniformly in } C([0, T]),$$

which in particular means that for all $t \in [0, T]$ we have $u^{j_k}(t, \cdot) \rightarrow u(t, \cdot)$ in the space $L^2(\Omega; \mathbb{R}^m)$. \square

Finally, we recall the following standard Lebesgue space interpolation result.

Lemma 4. *Let $u : [0, T] \times \Omega \rightarrow \mathbb{R}^m$, where $\Omega \subset \mathbb{R}^n$ is open and bounded, satisfy*

$$u \in L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m)) \cap L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$$

for some $p \geq 1$. Then, $u \in L^{\frac{p(n+2)}{n}}(Q_T; \mathbb{R}^m)$.

3. SPATIAL DISCRETIZATION AND CONVERGENCE OF THE SEMIDISCRETE PROBLEM

In this section we prove the existence of solutions to the system (1)–(9). In preparation for the construction of the fully discrete numerical approximation of the class of problems under consideration described in the next section, our proof of existence of large-data global-in-time Young measure solutions will be based

on performing a spatial finite element approximation. We shall therefore suppose henceforth that Ω is a Lipschitz polytope in \mathbb{R}^n . The proof presented below can be replicated on any bounded Lipschitz domain Ω by replacing the finite element basis, consisting of continuous piecewise linear basis functions satisfying a homogeneous Dirichlet boundary condition on $\partial\Omega$ by an abstract Galerkin basis. Suppose further that $0 < h_0 \ll \text{diam}(\Omega)$ and that $\{\mathcal{T}_h\}_{h \in (0, h_0]}$ is a shape-regular family of subdivisions \mathcal{T}_h of $\bar{\Omega}$ into closed simplexes Δ , where $h = \max_{\Delta \in \mathcal{T}_h} \text{diam}(\Delta)$. Consider the finite element space

$$V^h := \{v^h \in W_0^{1,\infty}(\bar{\Omega}) : v^h|_{\Delta} \text{ is affine for all } \Delta \in \mathcal{T}_h\}.$$

By V_m^h we denote the space of m -component vector-valued functions, each of whose components lies in V^h . We shall assume that $\{\mathcal{T}_h\}_{h \in (0, h_0]}$ is such that the $L^2(\Omega; \mathbb{R}^m)$ orthogonal projector P^h onto V_m^h is stable in $W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m)$; i.e., there exists a positive constant C , independent of $h \in (0, h_0]$, such that

$$(12) \quad \|D(P^h \varphi)\|_{L^{\hat{q}}(\Omega; \mathbb{R}^{m \times n})} \leq C \|D\varphi\|_{L^{\hat{q}}(\Omega; \mathbb{R}^{m \times n})} \quad \forall \varphi \in W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m),$$

where $\hat{q} := \max\{q, 2\}$. On globally quasiuniform subdivisions (12) is a consequence of a global inverse inequality and inequality (7) in [8]; we note, however, that the stability inequality (12) is in fact valid under less restrictive assumptions on the subdivision than quasiuniformity (see, for example, [6]).

Let $\{\phi_i^h\}_{i=1}^{\mathcal{N}(h)}$ be a basis for V_m^h . We shall therefore seek an approximate solution $u_h \in V_m^h$ in the form

$$u^h(t, x) = \sum_{i=1}^{\mathcal{N}(h)} \alpha_i^h(t) \phi_i^h(x), \quad t \in [0, T], \quad x \in \bar{\Omega},$$

satisfying

$$(13a) \quad (\partial_t u^h(t), v^h) + (a(Du^h(t)), Dv^h) + (Bu^h, v^h) = (F, v^h) \quad \forall v^h \in V_m^h,$$

for all $t \in [0, T]$, and

$$(13b) \quad u^h(0) = u_0^h,$$

with $u_0^h \in V_m^h$ and $u_0^h \rightarrow u_0$ (strongly) in $L^2(\Omega; \mathbb{R}^m)$ as $h \rightarrow 0_+$. The system given by (13a) is a system of ODEs for the coefficients α_i^h , $i = 1, \dots, \mathcal{N}(h)$, and a solution to (13a), (13b) exists locally by Peano's Theorem on some interval $[0, T_h) \subset [0, T]$, thanks to the assumed (local Lipschitz) continuity of a . Our existence theorem below is based on proving that u^h can be extended to the final time $T > 0$ for each $h \in (0, h_0]$, and that as $h \rightarrow 0_+$ a subsequence of the sequence of approximate solutions converges, in a sense to be made precise, to a Young measure solution of the problem.

Theorem 2. *Suppose that $\Omega \subset \mathbb{R}^n$ is a Lipschitz polytope, let $u_0 \in L^2(\Omega; \mathbb{R}^m)$, assume that*

$$p := \min_{i=1, \dots, m} p_i > \max \left\{ \frac{2n}{n+2}, 1 \right\} \quad \text{and} \quad q := \max_{i=1, \dots, m} p_i$$

satisfy $q - p < 1$, let $F \in L^{p'}(Q_T; \mathbb{R}^m) \cap L^2(Q_T; \mathbb{R}^m)$, and suppose that $a : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ is a locally Lipschitz mapping satisfying the assumptions (7)–(9). Then, there exists a Young measure solution (u, ν) of the problem (1) with data given by (5) and (6). Furthermore, there exists a subsequence (not indicated) of solutions $\{u^h\}$ to the semidiscrete problem (13a), (13b) such that

$$u^h \overset{*}{\rightharpoonup} u \text{ in } L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)), \quad Du_i^h \rightharpoonup Du_i \text{ in } L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^{m \times n})),$$

for $i = 1, \dots, m$; and, with $\hat{q} := \max\{q, 2\}$,

$$\begin{aligned} a(Du^h) &\rightharpoonup \langle \nu, a \rangle \text{ in } L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n}), & \partial_t u^h &\rightharpoonup \partial_t u \text{ in } L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)), \\ u^h &\rightarrow u \text{ in } C_w([0, T]; L^2(\Omega; \mathbb{R}^m)). \end{aligned}$$

Proof. As was noted above, (13a), (13b) is an initial-value problem for a system of ODEs for the coefficients α_i^h , and a solution exists locally by Peano's Theorem on some interval $[0, T_h) \subset [0, T]$, thanks to the assumed (local Lipschitz) continuity of a . We begin by showing that we can extend the numerical solution $u^h \in V_m^h$, defined on $[0, T_h) \times \bar{\Omega}$ up to time T for all $h \in (0, h_0]$, so that it is defined on the whole of $[0, T] \times \bar{\Omega}$. To this end, we take $v^h = u^h$ in (13a); hence, for all $\tau \in (0, T_h)$,

$$(14) \quad \int_0^\tau \int_\Omega \partial_t u^h \cdot u^h + K(Du^h)|Du^h|^2 + Bu^h \cdot u^h \, dx \, dt = \int_0^\tau \int_\Omega F \cdot u^h \, dx \, dt,$$

which, by using (4) and (8) on the left-hand side and Hölder's inequality on the right-hand side, yields, for all $\tau \in (0, T_h)$,

$$(15) \quad \begin{aligned} &\|u^h(\tau, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \int_0^\tau \int_\Omega \left[\sum_{i=1}^m (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} \right] |Du^h|^2 \, dx \, dt \\ &\leq c \left[\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \left(\int_0^T \|F\|_{L^{p'}(\Omega; \mathbb{R}^m)}^{p'} \, dt \right)^{\frac{1}{p'}} \left(\int_0^\tau \|u^h\|_{L^p(\Omega; \mathbb{R}^m)}^p \, dt \right)^{\frac{1}{p}} \right]. \end{aligned}$$

In order to bound the second term on the left-hand side of (15) from below, we note that

$$|Du_i^h|^{p_i} \leq (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_i^h|^2 \quad \text{if } p_i \geq 2,$$

and the following inequality holds for all $s \in [0, \infty)$ if $1 < p_i < 2$ and $\mu_i \neq 0$:

$$s^{p_i} \leq \begin{cases} 2^{1-\frac{p_i}{2}} (\mu_i^2 + s^2)^{\frac{p_i-2}{2}} s^2 & \text{for } s \geq |\mu_i|, \\ (\mu_i^2 + s^2)^{\frac{p_i-2}{2}} s^2 + |\mu_i|^{p_i} & \text{for } 0 \leq s \leq |\mu_i|. \end{cases}$$

Using these, we deduce that, for all $\tau \in (0, T_h)$,

$$\begin{aligned} & \|u^h(\tau, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \int_0^\tau \int_\Omega \sum_{i=1}^m |Du_i^h|^{p_i} dx dt \\ & \leq c \left[1 + \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \left(\int_0^\tau \|F\|_{L^{p'}(\Omega; \mathbb{R}^m)}^{p'} dt \right)^{\frac{1}{p'}} \left(\int_0^\tau \|u^h\|_{L^p(\Omega; \mathbb{R}^m)}^p dt \right)^{\frac{1}{p}} \right]. \end{aligned}$$

Thus, using Young's inequality and Poincaré's inequality in $W_0^{1,p}(\Omega; \mathbb{R}^m)$ in order to absorb the final factor on the right-hand side into the second term on the left-hand side, we arrive at the following energy inequality:

$$(16) \quad \sup_{t \in [0, T_h]} \|u^h(t, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \int_0^{T_h} \int_\Omega \sum_{i=1}^m |Du_i^h|^{p_i} dx dt \leq c(1 + \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}),$$

which provides uniform bounds on the sequence $\{u^h\}$, thanks to the strong convergence of $u_0^h \rightarrow u_0$ in $L^2(\Omega; \mathbb{R}^m)$ as $h \rightarrow 0_+$, which implies in particular that $\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}$ is uniformly bounded in h . Hence, the right-hand side of (16) is bounded, independent of $h \in (0, h_0]$. This means that T_h cannot be the maximal existence time, and u^h can be therefore extended from $[0, T_h)$ beyond T_h to the whole of the time interval $[0, T]$ for all $h \in (0, h_0]$. Hence,

$$(17) \quad \begin{aligned} \sup_{t \in [0, T]} \|u^h(t, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \int_0^T \int_\Omega \left[\sum_{i=1}^m (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} \right] |Du_i^h|^2 dx dt \\ \leq c(1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}) \end{aligned}$$

and also

$$(18) \quad \sup_{t \in [0, T]} \|u^h(t, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \int_0^T \int_\Omega \sum_{i=1}^m |Du_i^h|^{p_i} dx dt \leq c(1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}).$$

The uniform bound (18) implies the existence of a function

$$u \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)) \cap L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m)) \text{ with } Du_i \in L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^{m \times n}))$$

for $i = 1, \dots, m$, for which, up to a subsequence, we have

$$(19) \quad u^h \overset{*}{\rightharpoonup} u \text{ in } L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$$

and

$$(20) \quad Du_i^h \rightharpoonup Du_i \text{ in } L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^{m \times n})), \quad i = 1, \dots, m.$$

Furthermore, thanks to (8) and because the Euclidean norm of the vector $(|Du_1^h|, \dots, |Du_m^h|)^T$ is bounded by its 1-norm, we have the following bound:

(21)

$$|a(Du^h)| \leq c_1 \sum_{i=1}^m (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} \left(\sum_{k=1}^m |Du_k^h|^2 \right)^{\frac{1}{2}} \leq c_1 \sum_{i=1}^m \sum_{k=1}^m (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h|.$$

Now we seek to bound each of the terms appearing in the last sum in (21). Clearly for the case $i = k$ we have an expression of the form $(\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-1}{2}}$ appearing, which is bounded in $L^{p_i'}(Q_T)$, uniformly with respect to $h \in (0, h_0]$, by the energy estimate (18). In the case $i \neq k$ we proceed by using Hölder's inequality together with (17) and (18), our objective being to bound all of these terms in $L^{p_i'}(Q_T)$ if $p_i \geq 2$ or in $L^2(Q_T)$ if $1 < p_i < 2$. We begin with the terms for which $p_i \geq 2$:

$$\begin{aligned} & \int_0^T \int_{\Omega} \left[(\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h| \right]^{p_i'} dx dt \\ &= \int_0^T \int_{\Omega} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i(p_i-2)}{4}} |Du_k^h|^{p_i'} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i(p_i-2)}{4}} dx dt \\ &\leq \left(\int_0^T \int_{\Omega} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h|^2 dx dt \right)^{\frac{p_i'}{2}} \left(\int_0^T \int_{\Omega} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i}{2}} dx dt \right)^{1-\frac{p_i'}{2}} \\ &\leq c \left(1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p_i'}(Q_T; \mathbb{R}^m)}^{p_i'} \right)^{\frac{p_i'}{2}} \left(1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p_i'}(Q_T; \mathbb{R}^m)}^{p_i'} \right)^{1-\frac{p_i'}{2}}. \end{aligned}$$

Now for the terms in which $1 < p_i < 2$, thanks to (17), we have

$$\begin{aligned} & \int_0^T \int_{\Omega} \left[(\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h| \right]^2 dx dt \\ &= \int_0^T \int_{\Omega} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h|^2 dx dt \\ &\leq |\mu_i|^{p_i-2} \int_0^T \int_{\Omega} (\mu_i^2 + |Du_i^h|^2)^{\frac{p_i-2}{2}} |Du_k^h|^2 dx dt \\ &\leq c |\mu_i|^{p_i-2} (1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p_i'}(Q_T; \mathbb{R}^m)}^{p_i'}). \end{aligned}$$

Thus we have bounded each term in the sum appearing on the right-hand side of (21) in $L^{p_i'}(Q_T)$ (if $p_i \geq 2$) or in $L^2(Q_T)$ (if $1 < p_i < 2$), uniformly with respect to $h \in (0, h_0]$, and therefore $\{a(Du^h)\}_{0 < h \leq h_0}$ is a bounded sequence in $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$. More precisely, there exists a positive constant C , independent of h , such that

$$(22) \quad \|a(Du^h)\|_{L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})} \leq C \quad \forall h \in (0, h_0],$$

where we recall that $\hat{q} = \max\{q, 2\}$, whereby $\hat{q}' = \min\{q', 2\}$. Therefore, there is a subsequence (still indexed only by h) and a $\chi \in L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$ such that

$$(23) \quad a(Du^h) \rightharpoonup \chi \quad \text{in } L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n}).$$

Using similar estimates to those above we can also show that there exists a positive constant C , independent of h , such that

$$(24) \quad \|\partial_t u^h\|_{L^{\hat{q}}(0,T;W^{-1,\hat{q}}(\Omega;\mathbb{R}^m))} \leq C \quad \forall h \in (0, h_0].$$

Indeed, for each $\varphi \in L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$, by (13a), (12) and (22), we have that

$$\begin{aligned} \int_0^T (\partial_t u^h, \varphi) dt &= \int_0^T (\partial_t u^h, P^h \varphi) dt = - \int_0^T (a(Du^h(t)), DP^h \varphi) dt + \int_0^T (F - Bu^h, P^h \varphi) dt \\ &\leq C \|a(Du^h)\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})} \|D\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})} \\ &\quad + \left(\|F\|_{L^2(Q_T; \mathbb{R}^m)} + \|B\| \|u^h\|_{L^2(Q_T; \mathbb{R}^m)} \right) \|\varphi\|_{L^2(Q_T; \mathbb{R}^m)} \\ &\leq C (\|D\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})} + \|\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^m)}) \leq C \|\varphi\|_{L^{\hat{q}}(0,T;W^{1,\hat{q}}(\Omega;\mathbb{R}^m))}, \end{aligned}$$

and then (24) follows by noting that the dual space of $L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$ is $L^{\hat{q}'}(0, T; W^{-1,\hat{q}'}(\Omega; \mathbb{R}^m))$.

Now let $\psi \in L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$ be the strong limit, as $h \rightarrow 0_+$ and $l \rightarrow \infty$, of the sequence of functions

$$\psi^{\mathcal{N}(h),l}(t, x) = \sum_{i=1}^{\mathcal{N}(h)} \beta_i^l(t) \phi_i^h(x)$$

in the $L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$ norm, with $\beta_i^l \in C([0, T])$ and $\phi_i^h \in V_m^h$, $i = 1, \dots, \mathcal{N}(h)$, $l = 1, 2, \dots$; such a limit necessarily exists by density. We then have that

$$\begin{aligned} &\int_0^T \int_{\Omega} [\partial_t u^h \cdot \psi + a(Du^h) : D\psi + Bu^h \cdot \psi - F \cdot \psi] dx dt \\ &= \int_0^T \int_{\Omega} \partial_t u^h \cdot \psi^{\mathcal{N}(h),l} + a(Du^h) : D\psi^{\mathcal{N}(h),l} + Bu^h \cdot \psi^{\mathcal{N}(h),l} - F \cdot \psi^{\mathcal{N}(h),l} dx dt \\ &\quad + \int_0^T \int_{\Omega} \left[\partial_t u^h \cdot (\psi - \psi^{\mathcal{N}(h),l}) + a(Du^h) : (D\psi - D\psi^{\mathcal{N}(h),l}) \right. \\ &\quad \left. + Bu^h \cdot (\psi - \psi^{\mathcal{N}(h),l}) - F \cdot (\psi - \psi^{\mathcal{N}(h),l}) \right] dx dt \\ &\leq \|\partial_t u^h\|_{L^{\hat{q}}(0,T;W^{-1,\hat{q}}(\Omega;\mathbb{R}^m))} \|\psi - \psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;W_0^{1,\hat{q}}(\Omega;\mathbb{R}^m))} \\ &\quad + \|a(Du^h)\|_{L^{\hat{q}}(0,T;L^{\hat{q}}(\Omega;\mathbb{R}^{m \times n}))} \|D\psi - D\psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;L^{\hat{q}}(\Omega;\mathbb{R}^{m \times n}))} \\ &\quad + \|F\|_{L^{p'}(Q_T;\mathbb{R}^m)} \|\psi - \psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;W_0^{1,\hat{q}}(\Omega;\mathbb{R}^m))} \\ &\quad + c \|u^h\|_{L^{\hat{q}}(Q_T;\mathbb{R}^m)} \|\psi - \psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;W_0^{1,\hat{q}}(\Omega;\mathbb{R}^m))} \\ &\leq c (\|u_0\|_{L^2(\Omega;\mathbb{R}^m)}, \mu_i, T, |\Omega|, n, F) \\ &\quad \times (\|\psi - \psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;W_0^{1,\hat{q}}(\Omega;\mathbb{R}^m))} + \|D\psi - D\psi^{\mathcal{N}(h),l}\|_{L^{\hat{q}}(0,T;L^{\hat{q}}(\Omega;\mathbb{R}^{m \times n}))}). \end{aligned}$$

The right-hand side of this inequality converges to zero in the limit as $h \rightarrow 0_+$ (and therefore $\mathcal{N}(h) \rightarrow \infty$) and $l \rightarrow \infty$. Thus we have that

$$(25) \quad \lim_{h \rightarrow 0_+} \int_0^T \int_{\Omega} \partial_t u^h \cdot \psi + a(Du^h) : D\psi + Bu^h \cdot \psi \, dx \, dt = \int_0^T \int_{\Omega} F \cdot \psi \, dx \, dt,$$

for all $\psi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$.

The uniform in h bounds established in (24) and (18) now allow passage to the limit in the first and third term on the left-hand side of (25): by the uniform in h bound on $\|u^h\|_{L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))}$, there exists a subsequence (still indexed by h) which converges weakly-* to u in $L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$. This enables us to also pass to the limit in the third term of the integrand on the left-hand side of (25). In addition, thanks to (24), one can extract a further subsequence (which we continue to index by h) for which, by the uniqueness of the weak limit,

$$(26) \quad \partial_t u^h \rightharpoonup \partial_t u \quad \text{in} \quad L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)),$$

meaning that

$$\lim_{h \rightarrow 0_+} \int_0^T (\partial_t u^h, \psi) \, dt = \int_0^T \langle \partial_t u, \psi \rangle \, dt,$$

for all $\psi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$, where $\langle \cdot, \cdot \rangle$ is the duality pairing between the spaces $W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)$ and $W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m)$. Hence,

$$(27) \quad \int_0^T \langle \partial_t u, \psi \rangle \, dt + \lim_{h \rightarrow 0_+} \int_0^T \int_{\Omega} a(Du^h) : D\psi + \int_0^T \int_{\Omega} Bu \cdot \psi \, dx \, dt = \int_0^T \int_{\Omega} F \cdot \psi \, dx \, dt,$$

for all $\psi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$. The final convergence result in the statement of the theorem, that $u^h \rightarrow u$ in $C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$, is a direct consequence of (19), (26) and Lemma 3.

In order to pass to the limit in the second term in (27), we shall invoke Theorem 1. To this end, we begin by verifying that the subsequence $\{Du^{h_k}\}$ satisfies the tightness condition from Theorem 1; i.e., that

$$(28) \quad \lim_{s \rightarrow \infty} \sup_k |\{(t, x) \in ([0, T] \times \Omega) \cap Q_R(0) : |Du^{h_k}(t, x)| \geq s\}| = 0,$$

for every $R > 0$. In order to show this we set

$$A_{k, s, R} := \{(t, x) \in ([0, T] \times \Omega) \cap Q_R(0) : |Du^{h_k}(t, x)| \geq s\},$$

and apply Chebyshev's inequality:

$$\begin{aligned} |A_{k, s, R}| &\leq \frac{1}{s^p} \int_{A_{k, s, R}} |Du^{h_k}|^p \, dx \\ &\leq \frac{c(1 + \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'})}{s^p}, \end{aligned}$$

where we have used (18) along with the fact that for $r_1 \geq r_2$, we have $|s|^{r_2} \leq 1 + |s|^{r_1}$ to transition from the first line to the second. The resulting bound is independent of k (and h_k) and R , and so by passing $s \rightarrow \infty$ we deduce that (28) holds.

By taking $\mathfrak{K} = \mathbb{R}^m$ (see Remark 3 following the statement of Theorem 1), we may apply Theorem 1, to deduce the existence of a parameterized family of Young measures $(\nu_{t,x})_{(t,x) \in Q_T}$ and a further subsequence (which we still index by h) for which

$$(29) \quad a(Du^h) \rightharpoonup \langle \nu, a \rangle \quad \text{in } L^\alpha(Q_T; \mathbb{R}^{m \times n}),$$

where $1 < \alpha < \frac{p}{q-1}$ (see Remark 2, combined with the fact that $a(A) \leq c(1 + |A|^{q-1})$, which follows from (8) by recalling that $q := \max_{i=1, \dots, m} p_i$, whence $a \in E_{p/\alpha}$). By comparing (23) and (29) and noting the uniqueness of weak limits there is then a subsequence of u^h (still indexed by h) such that

$$(30) \quad a(Du^h) \rightharpoonup \langle \nu, a \rangle \quad \text{in } L^{\hat{q}'}(Q_T; \mathbb{R}^m).$$

Here we have used that $\hat{q}' = \min\{q', 2\} = \min\{\frac{q}{q-1}, 2\} \geq \min\{\frac{p}{q-1}, 2\} \geq \min\{\alpha, 2\}$, with α as above, and noted that (29) implies that $a(Du^h) \rightharpoonup \langle \nu, a \rangle$ in the space $L^{\min\{\alpha, 2\}}(Q_T; \mathbb{R}^{m \times n})$ because $\alpha \geq \min\{\alpha, 2\}$.

This allows us to complete the passage to the limit and obtain the existence of a pair (u, ν) such that

$$u \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)) \cap L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m)), \quad Du_i \in L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^{m \times n})),$$

for $i = 1, \dots, m$, with

$$\partial_t u \in L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)), \quad u \in C_w([0, T]; L^2(\Omega; \mathbb{R}^m)), \quad \langle \nu, a \rangle \in L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$$

satisfying

$$\int_0^T \langle \partial_t u, \phi \rangle + \int_\Omega \langle \nu_{t,x}, a \rangle : D\phi + Bu \cdot \phi \, dx \, dt = \int_0^T \int_\Omega F \cdot \phi \, dx \, dt$$

for all $\phi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$.

It remains to be shown that the initial condition is satisfied. By Lemma 3 it follows that we can extract a further subsequence h_l along which we have that

$$(31) \quad \int_\Omega u^{h_l}(t, x) \cdot \varphi(x) \, dx \rightarrow \int_\Omega u(t, x) \cdot \varphi(x) \, dx,$$

uniformly in $C([0, T])$, for all $\varphi \in L^2(\Omega; \mathbb{R}^m)$. Let us decompose

$$\begin{aligned} & \int_\Omega (u_0(x) - u(0, x)) \cdot \varphi(x) \, dx \\ &= \int_\Omega (u_0(x) - u^{h_l}(0, x)) \cdot \varphi(x) \, dx + \int_\Omega (u^{h_l}(0, x) - u^{h_l}(t, x)) \cdot \varphi(x) \, dx \\ & \quad + \int_\Omega (u^{h_l}(t, x) - u(t, x)) \cdot \varphi(x) \, dx + \int_\Omega (u(t, x) - u(0, x)) \cdot \varphi(x) \, dx \\ &=: I_l + II_l(t) + III_l(t) + IV(t). \end{aligned}$$

The term I_l converges to zero as $h_l \rightarrow 0_+$ by the assumed strong convergence in $L^2(\Omega; \mathbb{R}^m)$ of the discretized initial condition $u_0^{h_l}$ to u_0 , and the term $III_l(t)$ converges to zero, uniformly in $C([0, T])$ as $h_l \rightarrow 0_+$ by (31). Hence, for any $\varepsilon > 0$, there exists an $h_l^* = h_l^*(\varepsilon, \varphi) \in (0, h_0]$, such that

$$\left| \int_{\Omega} (u_0(x) - u^{h_l^*}(0, x)) \cdot \varphi(x) \, dx \right| < \frac{1}{4}\varepsilon \quad \text{and} \quad \max_{t \in [0, T]} \left| \int_{\Omega} (u^{h_l^*}(t, x) - u(t, x)) \cdot \varphi(x) \, dx \right| < \frac{1}{4}\varepsilon.$$

Consequently, for any $\varepsilon > 0$, there exists an $h_l^* \in (0, h_0]$, such that

$$\begin{aligned} & \left| \int_{\Omega} (u_0(x) - u(0, x)) \cdot \varphi(x) \, dx \right| \\ & < \frac{1}{2}\varepsilon + \left| \int_{\Omega} (u^{h_l^*}(0, x) - u^{h_l^*}(t, x)) \cdot \varphi(x) \, dx \right| + \left| \int_{\Omega} (u(t, x) - u(0, x)) \cdot \varphi(x) \, dx \right|. \end{aligned}$$

With h_l^* fixed, we now pass to the limit $t \rightarrow 0_+$ in this inequality. The term $II_l(t)$ converges to zero as $t \rightarrow 0_+$ because $u^{h_l^*}$, as a solution to the Galerkin equation (13a), is continuous and satisfies the discretized initial condition. The term $IV(t)$ converges to zero as $t \rightarrow 0_+$ since by Lemma 3 we have that $u \in C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$, the space of weakly continuous functions from the interval $[0, T]$ into $L^2(\Omega; \mathbb{R}^m)$. Thus, by passing to the limit $t \rightarrow 0_+$ in the last inequality we deduce that, for any $\varphi \in L^2(\Omega; \mathbb{R}^m)$ and any $\varepsilon > 0$,

$$\left| \int_{\Omega} (u_0(x) - u(0, x)) \cdot \varphi(x) \, dx \right| < \frac{1}{2}\varepsilon.$$

Taking $\varphi(x) = u_0(x) - u(0, x)$ and letting $\varepsilon \rightarrow 0_+$ then yields that $u_0(x) = u(0, x)$ for almost every $x \in \Omega$. \square

4. THE FULLY DISCRETE SCHEME AND ITS CONVERGENCE ANALYSIS

In the course of the proof of Theorem 2 Peano's Theorem was applied in order to show that one can solve the Galerkin equation (13a). We shall now employ an implicit Euler discretization scheme to numerically approximate the solution whose existence on $[0, T]$ is guaranteed by Peano's Theorem. Starting from our discretized initial condition $u_0^h \in V_m^h$, we inductively define u_{i+1}^h , $i = 0, \dots, N-1$, as solutions to the following fully discrete problem.

Let $\Delta t = T/N$ where $N \geq 1$, and given $u_i^h \in V_m^h$, find $u_{i+1}^h \in V_m^h$ such that

$$(32) \quad \int_{\Omega} \frac{u_{i+1}^h - u_i^h}{\Delta t} \cdot \phi^h + a(Du_{i+1}^h) : D\phi^h + Bu_{i+1}^h \cdot \phi^h \, dx = \int_{\Omega} F_{i+1} \cdot \phi^h \, dx$$

for $i = 0, \dots, N-1$ and for all $\phi^h \in V_m^h$. Here,

$$F_i(x) := \frac{1}{\Delta t} \int_{t_{i-1}}^{t_i} F(t, x) \, dt,$$

and, letting $t_i := i\Delta t$, we define on Q_T the function

$$\tilde{F}_{\Delta t}(t, x) := F_i(x) \quad \text{for } t \in (t_{i-1}, t_i], \quad i = 1, \dots, N, \quad x \in \Omega.$$

Given a function u_i^h , the fact that there exists a function u_{i+1}^h satisfying (32) is standard, and can be shown using a simple consequence of Brouwer's Fixed Point Theorem (cf., for example, Corollary 1.1 on p.279 of [13]). We begin the convergence analysis with the following estimate, which is reminiscent of the energy estimate obtained in the proof of Theorem 2, for the semidiscrete problem considered there.

Lemma 5. *There exists a positive constant c , independent of h and Δt , such that the functions u_i^h , $i = 1, \dots, N$, satisfy, for all $h \in (0, h_0]$,*

$$(33) \quad \max_{i=1, \dots, N} \|u_i^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \sum_{i=1}^N \int_{\Omega} |u_i^h - u_{i-1}^h|^2 dx + \Delta t \sum_{i=1}^N \int_{\Omega} a(Du_i^h) : Du_i^h dx \\ \leq c(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + T\|F\|_{L^2(Q_T; \mathbb{R}^m)}^2).$$

We note here that $a(\xi) : \xi \geq 0$ for all $\xi \in \mathbb{R}^{m \times n}$, and therefore each of the terms appearing on the left-hand side of the inequality (33) is nonnegative. In addition, the expression on the right-hand side of (33) is further bounded above by a constant that is independent of h and Δt .

Proof. We begin by choosing $\phi^h = u_{i+1}^h$ in (32), and noting the following identity:

$$\frac{(a-b) \cdot a}{\Delta t} = \frac{|a|^2 - |b|^2}{2\Delta t} + \frac{|a-b|^2}{2\Delta t} \quad \forall a, b \in \mathbb{R}^m.$$

Using this in (32) gives

$$\int_{\Omega} \frac{|u_{i+1}^h|^2 - |u_i^h|^2}{2\Delta t} + \frac{|u_{i+1}^h - u_i^h|^2}{2\Delta t} + a(Du_{i+1}^h) : Du_{i+1}^h + Bu_{i+1}^h \cdot u_{i+1}^h dx = \int_{\Omega} F_{i+1} \cdot u_{i+1}^h dx.$$

Now we use that $Bv \cdot v \geq 0$ for all $v \in \mathbb{R}^m$, multiply by $2\Delta t$ and sum through $i = 0, \dots, M$ for some $M < N$ to get

$$(34) \quad \int_{\Omega} |u_{M+1}^h|^2 dx + \sum_{i=0}^M \int_{\Omega} |u_{i+1}^h - u_i^h|^2 dx + 2\Delta t \sum_{i=0}^M \int_{\Omega} a(Du_{i+1}^h) : Du_{i+1}^h dx \\ \leq \int_{\Omega} |u_0^h|^2 dx + 2 \sum_{i=0}^M \int_{\Omega} \left(\int_{t_i}^{t_{i+1}} F(t, x) dt \right) \cdot u_{i+1}^h dx \\ \leq \int_{\Omega} |u_0^h|^2 dx + 2T\|F\|_{L^2(Q_T; \mathbb{R}^m)}^2 + \frac{1}{2} \max_{i=1, \dots, N} \|u_i^h\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

By omitting the second and the third term from the left-hand side of the inequality (34) noting the independence of its right-hand side of M , and then taking the

maximum over $M = 1, \dots, N - 1$, yields the desired bound on the first term on the left-hand side of the inequality (33). We then return with that bound to the inequality (34) to further estimate its right-hand side from above, whilst omitting the first term from the left-hand side of (34); thus we arrive at the desired bound on the second and third term on the left-hand side of the inequality (33). The expression on the right-hand side of (33) is clearly independent of Δt . The final assertion in the statement of the lemma is a consequence of the assumed strong convergence of u_0^h to u_0 in $L^2(\Omega; \mathbb{R}^m)$. \square

As the functions u_i^h , $i = 0, \dots, N$, are defined only on $\{t_0 = 0, t_1, \dots, t_N = T\} \times \bar{\Omega}$ rather than on the whole of \bar{Q}_T , we shall next extend them to \bar{Q}_T . Thus we define the following two functions:

$$u_{\Delta t}^h(t) := \frac{t - t_{i-1}}{\Delta t} u_i^h + \frac{t_i - t}{\Delta t} u_{i-1}^h \quad \text{for } t \in [t_{i-1}, t_i], \quad i = 1, \dots, N,$$

which is continuous and piecewise linear with respect to $t \in [0, T]$, and

$$\tilde{u}_{\Delta t}^h(t) := u_i^h \quad \text{for } t \in (t_{i-1}, t_i], \quad i = 1, \dots, N, \quad \text{and} \quad \tilde{u}_{\Delta t}^h(t) := u_0^h \quad \text{for } t \in [-\Delta t, 0],$$

which is piecewise constant with respect to $t \in [-\Delta t, T]$. With these new notations, (32) can be written as

$$(35) \quad \int_{\Omega} \partial_t u_{\Delta t}^h \cdot \phi^h + a(D\tilde{u}_{\Delta t}^h) : D\phi^h + B\tilde{u}_{\Delta t}^h \cdot \phi^h \, dx = \int_{\Omega} \tilde{F}_{\Delta t} \cdot \phi^h \, dx \quad \text{for } t \in [0, T],$$

for all $\phi^h \in V_m^h$, with $u_{\Delta t}^h(0, \cdot) = u_0^h$. We aim now to obtain uniform bounds on the sequences $\{u_{\Delta t}^h\}$ and $\{\tilde{u}_{\Delta t}^h\}$, which will allow us to complete our passage to the limit in Δt , possibly along some subsequence.

Lemma 6. *There exists a positive constant c , independent of h and Δt , such that, for all $h \in (0, h_0]$,*

$$\begin{aligned} \|\tilde{u}_{\Delta t}^h\|_{L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))}^2 &\leq c(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2), \\ \|\tilde{u}_{\Delta t}^h - \tilde{u}_{\Delta t}^h(\cdot - \Delta t)\|_{L^2(0, T; L^2(\Omega; \mathbb{R}^m))}^2 &\leq c(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2), \\ \|a(D\tilde{u}_{\Delta t}^h)\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})}^{\hat{q}'} &\leq c\left(1 + \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}\right), \\ \sum_{i=1}^m \|D(\tilde{u}_{\Delta t}^h)_i\|_{L^{p_i}(0, T; L^{p_i}(\Omega; \mathbb{R}^n))}^{p_i} &\leq c\left(1 + \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}\right), \end{aligned}$$

where $\hat{q} := \max\{q, 2\}$. Furthermore, the right-hand sides of these inequalities are bounded above by a positive constant, independent of $h \in (0, h_0]$ and Δt .

Proof. The first two bounds follow directly from the bounds on the first two terms in (33), while the fourth inequality is proved by an argument that is completely analogous to the proof of the bound on the second term appearing on the left-hand

side of (18). We shall therefore focus our attention on showing the third bound in the statement of the lemma. To this end, we begin by noting that since $a(\xi) : \xi \geq 0$ for all $\xi \in \mathbb{R}^{m \times n}$, we have that

$$|a(Du_i^h)|^{q'} = |(a(Du_i^h) : Du_i^h) K(Du_i^h)|^{\frac{q'}{2}} = (a(Du_i^h) : D\tilde{u}_{\Delta t}^h)^{\frac{q'}{2}} |K(Du_i^h)|^{\frac{q'}{2}}.$$

First suppose that $q > 2$; then, $\hat{q} = q$ and $\hat{q}' = q' \in (1, 2)$; hence, by Young's inequality,

$$|a(Du_i^h)|^{\hat{q}'} = |a(Du_i^h)|^{q'} \leq \frac{q'}{2} a(Du_i^h) : Du_i^h + \frac{2 - q'}{2} |K(Du_i^h)|^{\frac{q'}{2 - q'}}.$$

Therefore, by (33) and (8), we have that

$$\begin{aligned} \int_0^T \int_{\Omega} |a(D\tilde{u}_{\Delta t}^h)|^{\hat{q}'} dx dt &= \int_0^T \int_{\Omega} |a(D\tilde{u}_{\Delta t}^h)|^{q'} dx dt = \Delta t \sum_{i=1}^N \int_{\Omega} |a(Du_i^h)|^{q'} dx \\ &\leq \frac{q'}{2} \Delta t \sum_{i=1}^N \int_{\Omega} a(Du_i^h) : Du_i^h dx + \frac{2 - q'}{2} \Delta t \sum_{i=1}^N \int_{\Omega} |K(Du_i^h)|^{\frac{q'}{2 - q'}} dx \\ &\leq c(q) \left(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + T \|F\|_{L^2(\Omega; \mathbb{R}^m)}^2 \right) \\ &\quad + c(q) \Delta t \sum_{i=1}^N \int_{\Omega} \left(c_1 \sum_{j=1}^m (|\mu_j|^2 + |Du_{i,j}^h|^2)^{\frac{p_j - 2}{2}} \right)^{\frac{q'}{2 - q'}} dx \\ &\leq c(q) \left(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + T \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2 \right) \\ &\quad + c(c_1, q, m) \Delta t \sum_{i=1}^N \int_{\Omega} \left[\sum_{j=1}^m |\mu_j|^{\frac{(p_j - 2)q}{q - 2}} + \sum_{\substack{j=1 \\ p_j \geq 2}}^m |Du_{i,j}^h|^{\frac{(p_j - 2)q}{q - 2}} \right] dx. \end{aligned}$$

As $q := \max\{p_1, \dots, p_m\} > 2$, for any $j \in \{1, \dots, m\}$ such that $p_j \geq 2$ we have that $\frac{(p_j - 2)q}{q - 2} < p_j$; it thus follows that

$$\begin{aligned} \int_0^T \int_{\Omega} |a(D\tilde{u}_{\Delta t}^h)|^{\hat{q}'} dx dt &\leq c(q) \left(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + T \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2 \right) \\ &\quad + c(c_1, q, m) T |\Omega| \sum_{j=1}^m |\mu_j|^{\frac{(p_j - 2)q}{q - 2}} + c(q, m, c_1, p_1, \dots, p_m, |\Omega|) \Delta t \sum_{i=1}^N \int_{\Omega} \left(\sum_{j=1}^m |Du_{i,j}^h|^{p_j} \right) dx. \end{aligned}$$

The final term on the right-hand side of this inequality is then bounded analogously as the corresponding term in the proof of Theorem 2 and we therefore omit the details (cf. (18)):

$$\Delta t \sum_{i=1}^N \int_{\Omega} \left(\sum_{j=1}^m |Du_{i,j}^h|^{p_j} dx \right) \leq c(1 + \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'}).$$

Substituting this into the right-hand side of the previous inequality completes the proof of the lemma in the case when $\hat{q} = q > 2$.

If $\hat{q} = 2$ (and therefore $\hat{q}' = 2$), then $p_j \in (1, 2]$ for all $j = 1, \dots, m$. It then follows from (8) that

$$0 \leq K(A) \leq c_1 \sum_{i=1}^m |\mu_i|^{p_i-2} \quad \forall A \in \mathbb{R}^{m \times n},$$

whereby

$$|a(Du_i^h)|^{\hat{q}'} = (a(Du_i^h) : D\tilde{u}_{\Delta t}^h) K(Du_i^h) \leq c_1 \left(\sum_{i=1}^m |\mu_i|^{p_i-2} \right) (a(Du_i^h) : D\tilde{u}_{\Delta t}^h).$$

Hence, by Lemma 5,

$$\begin{aligned} \int_0^T \int_{\Omega} |a(D\tilde{u}_{\Delta t}^h)|^{\hat{q}'} dx dt &= \Delta t \sum_{i=1}^N \int_{\Omega} |a(Du_i^h)|^{\hat{q}'} dx \\ &\leq c_1 \left(\sum_{i=1}^m |\mu_i|^{p_i-2} \right) \Delta t \sum_{i=1}^N \int_{\Omega} a(Du_i^h) : Du_i^h dx \\ &\leq c(c_1, p_1, \dots, p_m, \mu_1, \dots, \mu_m, m, T) \left(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^2(\Omega; \mathbb{R}^m)}^2 \right). \end{aligned}$$

The final assertion in the statement of the lemma is a consequence of the assumed strong convergence of u_0^h to u_0 as $h \rightarrow 0_+$. That completes the proof. \square

Next, we prove the required uniform bounds on the sequence $\{u_{\Delta t}^h\}$.

Lemma 7. *There exists a positive constant c , independent of h and Δt , such that, for all $h \in (0, h_0]$,*

$$\|u_{\Delta t}^h\|_{L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))}^2 \leq c(\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^2(Q_T; \mathbb{R}^m)}^2);$$

furthermore, with $\hat{q} := \max\{q, 2\}$,

$$\|\partial_t u_{\Delta t}^h\|_{L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m))} \leq c$$

and

$$(\Delta t)^{-1} \|\tilde{u}_{\Delta t}^h - \tilde{u}_{\Delta t}^h(\cdot - \Delta t)\|_{L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m))} \leq c.$$

Proof. The first inequality is just a restatement of the bound on the first term in the inequality (33) in terms of $u_{\Delta t}^h$, and the third inequality is a restatement of the second inequality in terms of $\tilde{u}_{\Delta t}^h$. We shall therefore concentrate on the proof of the second stated inequality. The argument proceeds along the same lines as the proof of (24). For each $\varphi \in L^{\hat{q}}(0, T; W_0^{1, \hat{q}}(\Omega; \mathbb{R}^m))$ we have, using the stability

in $W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m)$ of the $L^2(\Omega; \mathbb{R}^m)$ orthogonal projector $P^h : L^2(\Omega; \mathbb{R}^m) \rightarrow V_m^h$ (cf. (12)), that

$$\begin{aligned}
\int_0^T (\partial_t u_{\Delta t}^h, \varphi) dt &= \int_0^T (\partial_t u_{\Delta t}^h, P^h \varphi) dt = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \int_{\Omega} \frac{u_{i+1}^h - u_i^h}{\Delta t} \cdot P^h \varphi dx dt \\
&= - \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \int_{\Omega} a(Du_{i+1}^h) \cdot DP^h \varphi dx dt + \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (F_{i+1} - Bu_{i+1}^h) \cdot P^h \varphi dx dt \\
&= - \int_0^T (a(D\tilde{u}_{\Delta t}^h(t)), DP^h \varphi) dt + \int_0^T (\tilde{F}_{\Delta t} - B\tilde{u}_{\Delta t}^h, P^h \varphi) dt \\
&\leq \|a(D\tilde{u}_{\Delta t}^h)\|_{L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})} \|D\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})} \\
&\quad + \left(\|\tilde{F}_{\Delta t}\|_{L^2(Q_T; \mathbb{R}^m)} + |B| \|\tilde{u}_{\Delta t}^h\|_{L^2(Q_T; \mathbb{R}^m)} \right) \|\varphi\|_{L^2(Q_T; \mathbb{R}^m)} \\
&\leq c(\|D\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})} + \|\varphi\|_{L^{\hat{q}}(Q_T; \mathbb{R}^m)}) \leq c\|\varphi\|_{L^{\hat{q}}(0, T; W^{1,\hat{q}}(\Omega; \mathbb{R}^m))},
\end{aligned}$$

where in the transition to the last line we have used the third and the first bound from Lemma 6, together with the assumed strong convergence of u_0^h to u_0 in $L^2(\Omega; \mathbb{R}^m)$ to deduce that the constant c is independent of h and Δt . Dividing by $\|\varphi\|_{L^{\hat{q}}(0, T; W^{1,\hat{q}}(\Omega; \mathbb{R}^m))}$, taking the supremum over all $\varphi \in L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$ and recalling the definition of the norm of the dual space $L^{\hat{q}'}(0, T; W^{-1,\hat{q}'}(\Omega; \mathbb{R}^m)) = (L^{\hat{q}}(0, T; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m)))'$ the stated inequality directly follows. \square

We shall now discuss the convergence of the sequence of functions $\{u_{\Delta t}^h\}$ with $h \in (0, h_0]$ fixed.

Lemma 8. *There exists a subsequence of Δt , labelled Δt_k , and a function $u^h \in C([0, T]; V_m^h)$ such that $u_{\Delta t_k}^h \rightarrow u^h$ in $C([0, T]; L^2(\Omega; \mathbb{R}^m))$, and $\partial_t u_{\Delta t_k}^h \rightarrow \partial_t u^h$ in $L^{\hat{q}'}(0, T; W^{-1,\hat{q}'}(\Omega; \mathbb{R}^m))$ for any (fixed) $h \in (0, h_0]$, as $k \rightarrow \infty$.*

Proof. We begin by noting that, for any Δt and any $h \in (0, h_0]$, the functions $u_{\Delta t}^h$ are continuous in both t and x . Thanks to the first two bounds in Lemma 7 there exists a subsequence of Δt , labelled Δt_k , and a function $u^h \in L^\infty(0, T; V_m^h)$ such that $u_{\Delta t_k}^h \xrightarrow{*} u^h$ in $L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$, and $\partial_t u_{\Delta t_k}^h \rightarrow \partial_t u^h$ in $L^{\hat{q}'}(0, T; W^{-1,\hat{q}'}(\Omega; \mathbb{R}^m))$. By Lemma 3, $u^h \in C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$ and $u_{\Delta t_k}^h \rightarrow u^h$ in $C_w([0, T]; L^2(\Omega; \mathbb{R}^m))$. As both $u_{\Delta t}^h(t, \cdot)$ and $u^h(t, \cdot)$ belong to the finite-dimensional space V_m^h for all $t \in [0, T]$, it follows that $u^h \in C([0, T]; L^2(\Omega; \mathbb{R}^m))$ and $u_{\Delta t_k}^h \rightarrow u^h$ in $u^h \in C([0, T]; L^2(\Omega; \mathbb{R}^m))$, because $u^h \in C_w([0, T]; V_m^h) = C([0, T]; V_m^h)$ and because a sequence converges in $C_w([0, T]; V_m^h)$ if, and only if, it converges in $C([0, T]; V_m^h)$ as $k \rightarrow \infty$ for any (fixed) $h \in (0, h_0]$. \square

Next, we study the convergence of the sequence of functions $\{\tilde{u}_{\Delta t}^h\}$ with $h \in (0, h_0]$ fixed.

Lemma 9. *There exists a function $\tilde{u}^h \in L^\infty(0, T; V_m^h)$ and a subsequence Δt_k such that*

$$\tilde{u}_{\Delta t_k}^h \rightharpoonup \tilde{u}^h \text{ in } \bigtimes_{j=1}^m L^{p_j}(0, T; W_0^{1,p_j}(\Omega; \mathbb{R}^m)) \text{ and } \tilde{u}_{\Delta t_k}^h \rightarrow \tilde{u}^h \text{ in } L^r(0, T; L^2(\Omega; \mathbb{R}^m)),$$

for any $r \in [1, \infty)$ and fixed $h \in (0, h_0]$, as $k \rightarrow \infty$. Furthermore, along this subsequence we have $a(D\tilde{u}_{\Delta t_k}^h) \rightarrow a(D\tilde{u}^h)$ in $L^s(Q_T; \mathbb{R}^{m \times n})$ for any $s \in [1, \hat{q}')$ and $a(D\tilde{u}_{\Delta t_k}^h) \rightarrow a(D\tilde{u}^h)$ in $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$, for any fixed $h \in (0, h_0]$, as $k \rightarrow \infty$.

Proof. Thanks to the first and fourth bound from Lemma 6, there exists a $\tilde{u}^h \in L^\infty(0, T; V_m^h)$ and a subsequence Δt_k such that $\tilde{u}_{\Delta t_k}^h \overset{*}{\rightharpoonup} \tilde{u}^h$ in $L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$ and

$$\tilde{u}_{\Delta t_k}^h \rightharpoonup \tilde{u}^h \quad \text{in } \bigtimes_{j=1}^m L^{p_j}(0, T; W_0^{1,p_j}(\Omega; \mathbb{R}^m)),$$

for any fixed $h \in (0, h_0]$, as $k \rightarrow \infty$.

By applying Theorem 1 from [9] (with $X = W_0^{1,p}(\Omega; \mathbb{R}^m)$, $B = L^2(\Omega; \mathbb{R}^m)$ and $Y = W^{-1,\hat{q}'}(\Omega; \mathbb{R}^m)$ and $r = 1$, there), we deduce that $\{\tilde{u}_{\Delta t_k}^h\}$ is relatively compact in $L^p(0, T; L^2(\Omega; \mathbb{R}^m))$, for each fixed $h \in (0, h_0]$; therefore, because it is a bounded sequence in $L^\infty(0, T; L^2(\Omega; \mathbb{R}^m))$, by interpolation $\{\tilde{u}_{\Delta t_k}^h\}$ is also relatively compact in $L^r(0, T; L^2(\Omega; \mathbb{R}^m))$, for each fixed $h \in (0, h_0]$ and all $r \in [1, \infty)$.

Hence, in particular, $\{\tilde{u}_{\Delta t_k}^h\}$ is relatively compact in $L^1(0, T; L^1(\Omega; \mathbb{R}^m))$, for each fixed $h \in (0, h_0]$. As $\tilde{u}_{\Delta t_k}^h(t, \cdot) \in V_m^h$ for all $t \in [0, T]$, and V_m^h is finite-dimensional, for each fixed $h \in (0, h_0]$, by norm equivalence in finite-dimensional spaces, it follows that $\{\tilde{u}_{\Delta t_k}^h\}$ is relatively compact in $L^1(0, T; W_0^{1,1}(\Omega; \mathbb{R}^m))$, for each fixed $h \in (0, h_0]$. Therefore $D\tilde{u}_{\Delta t_k}^h \rightarrow D\tilde{u}^h$ in $L^1(Q_T; \mathbb{R}^{m \times n})$ as $k \rightarrow \infty$, for each fixed $h \in (0, h_0]$. Thus we can extract a subsequence with respect to k (not indicated) such that $D\tilde{u}_{\Delta t_k}^h \rightarrow D\tilde{u}^h$ a.e. on Q_T as $k \rightarrow \infty$, for each fixed $h \in (0, h_0]$. Thanks to the continuity of a , we then have that $a(D\tilde{u}_{\Delta t_k}^h) \rightarrow a(D\tilde{u}^h)$ a.e. on Q_T as $k \rightarrow \infty$, for each fixed $h \in (0, h_0]$. Since by the third inequality from Lemma 6 the sequence $\{a(D\tilde{u}_{\Delta t_k}^h)\}$ is weakly compact in $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$ and hence in particular also in $L^1(Q_T; \mathbb{R}^{m \times n})$, for each fixed $h \in (0, h_0]$, it follows by Vitali's convergence theorem that $a(D\tilde{u}_{\Delta t_k}^h) \rightarrow a(D\tilde{u}^h)$ strongly in $L^1(Q_T; \mathbb{R}^{m \times n})$ as $k \rightarrow \infty$, for each fixed $h \in (0, h_0]$, and therefore also strongly in $L^s(Q_T; \mathbb{R}^{m \times n})$ as $k \rightarrow \infty$, for each fixed $h \in (0, h_0]$ and any $s \in [1, \hat{q}')$. By the uniqueness of the weak limit it then follows that $a(D\tilde{u}_{\Delta t_k}^h) \rightarrow a(D\tilde{u}^h)$ weakly in $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$, for any fixed $h \in (0, h_0]$, as $k \rightarrow \infty$. \square

We now show that the limit functions u^h and \tilde{u}^h defined above are equal when considered as elements of the space $L^2(Q_T; \mathbb{R}^m)$.

Lemma 10. *The limiting functions u^h and \tilde{u}^h are equal in the space $L^2(Q_T; \mathbb{R}^m) = L^2(0, T; L^2(\Omega; \mathbb{R}^m))$.*

Proof. We apply the triangle inequality to see that, with $\Delta t_k = T/N_k$ and $h \in (0, h_0]$,

$$\begin{aligned} \|u^h - \tilde{u}^h\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))} &\leq \|u^h - u_{\Delta t_k}^h\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))} \\ &\quad + \|u_{\Delta t_k}^h - \tilde{u}_{\Delta t_k}^h\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))} + \|\tilde{u}_{\Delta t_k}^h - \tilde{u}^h\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))}. \end{aligned}$$

The first and last terms converge to zero in the limit as $\Delta t_k \rightarrow 0_+$ thanks to the convergence results stated in Lemma 8 and Lemma 9 (with $r = 2$), respectively, for each fixed $h \in (0, h_0]$. We now show that the second term also converges to zero in the limit:

$$\begin{aligned} \|u_{\Delta t_k}^h - \tilde{u}_{\Delta t_k}^h\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))}^2 &= \int_0^T \int_{\Omega} |u_{\Delta t_k}^h - \tilde{u}_{\Delta t_k}^h|^2 dx dt \\ &= \sum_{i=1}^{N_k} \int_{\Omega} \int_{t_{i-1}}^{t_i} \left| \frac{t - t_{i-1} - \Delta t_k}{\Delta t_k} u_i^h + \frac{t_i - t}{\Delta t_k} u_{i-1}^h \right|^2 dt dx \\ &= \frac{\Delta t_k}{3} \sum_{i=1}^{N_k} \int_{\Omega} |u_i^h - u_{i-1}^h|^2 dx \\ &= \frac{\Delta t_k}{3} \|\tilde{u}_{\Delta t_k}^h - \tilde{u}_{\Delta t_k}^h(\cdot - \Delta t)\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))}^2. \end{aligned}$$

Thanks to the second inequality in Lemma 6, the right-hand side of the last inequality converges to zero in the limit of $\Delta t_k \rightarrow 0_+$ ($N_k \rightarrow \infty$ as $k \rightarrow \infty$). Hence, $u^h = \tilde{u}^h$, as has been asserted. \square

Next, we discuss the attainment of the initial condition.

Lemma 11. *The limiting function u^h satisfies the initial condition in the following sense:*

$$\lim_{t \rightarrow 0_+} \|u^h(t, \cdot) - u_0^h(\cdot)\|_{L^2(\Omega;\mathbb{R}^m)} = 0,$$

for each fixed $h \in (0, h_0]$.

Proof. Thanks to Lemma 8, for any $\epsilon > 0$ there is an $\ell \in \mathbb{N}$ such that $\|u^h(t, \cdot) - u_{\Delta t_{k_l}}^h(t, \cdot)\|_{L^2(\Omega;\mathbb{R}^m)} < \epsilon$ for all $l \geq \ell$ and all $t \in [0, T]$. Furthermore, by the triangle inequality:

$$\|u^h(t, \cdot) - u_0^h(\cdot)\|_{L^2(\Omega;\mathbb{R}^m)} \leq \|u^h(t, \cdot) - u_{\Delta t_{k_l}}^h(t, \cdot)\|_{L^2(\Omega;\mathbb{R}^m)} + \|u_{\Delta t_{k_l}}^h(t, \cdot) - u_0^h(\cdot)\|_{L^2(\Omega;\mathbb{R}^m)}$$

for any $l \in \mathbb{N}$. Hence, in particular, for $l = \ell$ (fixed),

$$\|u^h(t, \cdot) - u_0^h(\cdot)\|_{L^2(\Omega;\mathbb{R}^m)} \leq \|u^h(t, \cdot) - u_{\Delta t_{k_\ell}}^h(t, \cdot)\|_{L^2(\Omega;\mathbb{R}^m)} + \epsilon.$$

As $u_{\Delta t_{k_\ell}}^h$ is a continuous function in t , passage to the limit in this inequality, using that $\lim_{t \rightarrow 0_+} u_{\Delta t_{k_l}}^h(t, x) = u_0^h(x)$, implies that, for each $\epsilon > 0$,

$$0 \leq \limsup_{t \rightarrow 0_+} \|u^h(t, \cdot) - u_0^h(\cdot)\|_{L^2(\Omega;\mathbb{R}^m)} \leq \epsilon.$$

Thus we have the asserted result. \square

Finally, we focus on the convergence of the source term as the time step tends to 0.

Lemma 12. *We have that $\lim_{\Delta t \rightarrow 0_+} \|\tilde{F}_{\Delta t}(\cdot, x) - F(\cdot, x)\|_{L^r(0, T)}$ for all $r \in [1, \infty)$ and a.e. $x \in \Omega$.*

Proof. The proof of this result is standard and is therefore omitted. \square

We are now ready for passage to the limit in (35) as $\Delta t \rightarrow 0_+$. By Lemma 8, Lemma 9, and Lemma 12, we see that the limit function u^h satisfies

$$\int_0^T \int_{\Omega} \partial_t u^h \cdot \Phi^h + a(Du^h) : D\Phi^h + Bu^h \cdot \Phi^h \, dx \, dt = \int_0^T \int_{\Omega} F \cdot \Phi^h \, dx \, dt \quad \begin{array}{l} \forall \Phi^h \in V_m^h, \\ \forall h \in (0, h_0], \end{array}$$

with $u^h(0, x) = u_0^h(x)$ for $x \in \bar{\Omega}$. We can now pass to the limit $h \rightarrow 0_+$ by (the proof of) Theorem 2, to deduce that as the discretization parameters converge to zero the solution of the fully discrete scheme (32) converges to a Young measure solution of the problem (1)–(9) under consideration.

In preparation for the considerations in the next section, we now discuss continuous dependence of $u_{\Delta t}^h$ on the initial data. We start by considering two functions, $u_0^h \in V_m^h$ and $v_0^h \in V_m^h$, and we let u_1^h and v_1^h be respective solutions to (32) after a single time step. Subtracting the resulting equations satisfied by u_1^h from the equation satisfied by v_1^h and choosing as the test function $\phi^h = u_1^h - v_1^h$, it follows that

$$\int_{\Omega} |u_1^h - v_1^h|^2 + \Delta t (a(Du_1^h) - a(Dv_1^h)) \cdot (Du_1^h - Dv_1^h) \, dx \leq \int_{\Omega} (u_0^h - v_0^h) \cdot (u_1^h - v_1^h) \, dx,$$

which after applying Young's inequality to the right-hand side and absorbing terms, gives

$$\|u_1^h - v_1^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + 2\Delta t \int_{\Omega} (a(Du_1^h) - a(Dv_1^h)) \cdot (Du_1^h - Dv_1^h) \, dx \leq \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

As the functions u_1^h and v_1^h are expressed in terms of a finite basis, we see through the equivalence of norms in finite-dimensional spaces and the first inequality in Lemma 7 that $\|Du_1^h\|_{L^\infty(\Omega; \mathbb{R}^{m \times n})}$ and $\|Dv_1^h\|_{L^\infty(\Omega; \mathbb{R}^{m \times n})}$ are both bounded by a constant $C = C(h, \|u_0\|_{L^2(\Omega; \mathbb{R}^m)}, \|F\|_{L^2(Q_T; \mathbb{R}^m)})$. Therefore we can make use of the local Lipschitz condition satisfied by a and the inverse inequality

$$\|Du_1^h - Dv_1^h\|_{L^2(\Omega; \mathbb{R}^m)} \leq C(h) \|u_1^h - v_1^h\|_{L^2(\Omega; \mathbb{R}^m)},$$

(with $C(h) = Ch^{-1}$ when the family of triangulations \mathcal{T}^h is quasiuniform) to deduce, by suppressing in our notation for the constant $C(h, \|a\|_{\text{Lip,loc}})$ below its dependence on the (fixed) data $\|u_0\|_{L^2(\Omega; \mathbb{R}^m)}$ and $\|F\|_{L^2(Q_T; \mathbb{R}^m)}$, that

$$\|u_1^h - v_1^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 (1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t) \leq \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

Therefore by choosing Δt sufficiently small for $h \in (0, h_0]$ fixed, so that

$$0 < 1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t < 1,$$

we have thus shown the desired continuous dependence on the initial data for a single time step. By iterating this estimate, noting that the constant $C(h, \|a\|_{\text{Lip,loc}})$ remains the same regardless of the choice of the time level $i \in \{1, \dots, N\}$, we obtain that, for each $i = 1, \dots, N$, the following estimate holds:

$$(36) \quad \|u_i^h - v_i^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 (1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t)^i \leq \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

Recall that Δt and N are related via the identity $T = N\Delta t$. Hence, as $N \rightarrow \infty$,

$$(1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t)^N = \left(1 - \frac{TC(h, \|a\|_{\text{Lip,loc}})}{N}\right)^N \rightarrow e^{-TC(h, \|a\|_{\text{Lip,loc}})}.$$

To transfer this continuous dependence on the initial data to the limiting functions u^h and v^h as $\Delta t \rightarrow 0_+$, we consider $u_{\Delta t}^h$ and $v_{\Delta t}^h$, defined, for $i = 1, \dots, N$, by

$$u_{\Delta t}^h(t, x) = \frac{t - t_{i-1}}{\Delta t} u_i^h(x) + \frac{t_i - t}{\Delta t} u_{i-1}^h(x) \quad \text{for } t \in [t_{i-1}, t_i] \text{ and } x \in \bar{\Omega},$$

and, analogously,

$$v_{\Delta t}^h(t, x) = \frac{t - t_{i-1}}{\Delta t} v_i^h(x) + \frac{t_i - t}{\Delta t} v_{i-1}^h(x) \quad \text{for } t \in [t_{i-1}, t_i] \text{ and } x \in \bar{\Omega}.$$

It then follows from (36) that

$$(37) \quad \sup_{t \in [0, T]} \|u_{\Delta t}^h - v_{\Delta t}^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 (1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t)^N \leq \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

Let Δt_k be a sequence such that $u_{\Delta t_k}^h \rightarrow u^h$ and $v_{\Delta t_k}^h \rightarrow v^h$ in the space $C([0, T]; L^2(\Omega; \mathbb{R}^m))$ and $k \rightarrow \infty$ (cf. Lemma 8). Now passing $k \rightarrow \infty$ (whereby $\Delta t_k = T/N_k \rightarrow 0_+$ and $N_k \rightarrow \infty$) we see that

$$\|u^h - v^h\|_{C([0, T]; L^2(\Omega; \mathbb{R}^m))} \leq e^{\frac{1}{2}TC(h, \|a\|_{\text{Lip,loc}})} \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)}.$$

We have therefore shown the following result.

Proposition 1. *Let all of the assumptions of Theorem 2 hold, with the additional assumption that, for $h \in (0, h_0]$ fixed, $\Delta t = T/N$ is chosen so that $0 < 1 - C(h, \|a\|_{\text{Lip,loc}})\Delta t < 1$. Then, there is a unique solution to (13a), which depends continuously on the choice of discretized initial data, and which can be constructed numerically by the approximations considered in this subsection.*

Remark 4. *The above proposition in particular implies that the map S_t^h sending the initial condition u_0^h to $S_t^h u_0^h = u^h(t, \cdot)$ is (Lipschitz) continuous from V_m^h to V_m^h for all $t \in [0, T]$ and $h \in (0, h_0]$ fixed, and by the equivalence of norms in*

finite-dimensional vector spaces, so is the map sending the initial condition u_0^h to $Du^h(t, \cdot)$. Recall that we are seeking the function u^h in terms of a finite element basis, i.e., $u^h(t, x) = \sum_{i=1}^{\mathcal{N}(h)} \alpha_i^h(t) \phi_i^h(x)$. The above considerations enable us to view the map from $\mathbb{R}^{\mathcal{N}(h)}$ to $\mathbb{R}^{\mathcal{N}(h)}$, which sends the vector $\underline{\alpha}^h(0)$ of coefficients of the discretized initial condition in terms of the finite-element basis to the vector $\underline{\alpha}^h(t)$ of coefficients of u^h , in the same finite-element basis, as a continuous function. To see this, let $\underline{\alpha}^h$ and $\underline{\beta}^h$ be the coefficients of u^h and v^h respectively, and let

$$M_{i,j}^h := \frac{(\phi_i^h, \phi_j^h)}{\|\phi_i^h\|_{L^2(\Omega; \mathbb{R}^m)} \|\phi_j^h\|_{L^2(\Omega; \mathbb{R}^m)}}, \quad i, j = 1, \dots, \mathcal{N}(h).$$

The (normalized) Gram matrix $M^h = (M_{i,j}^h)_{i,j=1}^{\mathcal{N}(h)}$ is symmetric positive definite, with $C_0 I \leq M^h \leq C_1 I$, where $0 < C_0 \leq C_1$, in the sense of symmetric positive definite matrices. By expressing $u^h(t, x) - v^h(t, x) = \sum_{i=1}^{\mathcal{N}(h)} (\alpha_i^h(t) - \beta_i^h(t)) \phi_i^h(x)$, it follows that,

$$\begin{aligned} C_0 \min_{i=1, \dots, \mathcal{N}(h)} \|\phi_i^h\|_{L^2(\Omega; \mathbb{R}^m)} \max_{t \in [0, T]} \left(\sum_{i=1}^{\mathcal{N}(h)} |\alpha_i^h(t) - \beta_i^h(t)|^2 \right)^{\frac{1}{2}} \\ \leq C_0 \max_{t \in [0, T]} \left(\sum_{i=1}^{\mathcal{N}(h)} |\alpha_i^h(t) - \beta_i^h(t)|^2 \int_{\Omega} |\phi_i^h(x)|^2 dx \right)^{\frac{1}{2}} \\ \leq \max_{t \in [0, T]} \left(\sum_{i,j=1}^{\mathcal{N}(h)} (\alpha_i^h(t) - \beta_i^h(t)) (\alpha_j^h(t) - \beta_j^h(t)) \int_{\Omega} \phi_i^h(x) \cdot \phi_j^h(x) dx \right)^{\frac{1}{2}} \\ = \max_{t \in [0, T]} \left(\int_{\Omega} \sum_{i,j=1}^{\mathcal{N}(h)} (\alpha_i^h(t) - \beta_i^h(t)) \phi_i^h(x) (\alpha_j^h(t) - \beta_j^h(t)) \phi_j^h(x) dx \right)^{\frac{1}{2}} \\ = \max_{t \in [0, T]} \left(\int_{\Omega} \left| \sum_{i=1}^{\mathcal{N}(h)} (\alpha_i^h(t) - \beta_i^h(t)) \phi_i^h(x) \right|^2 dx \right)^{\frac{1}{2}} \\ = \|u^h - v^h\|_{C([0, T]; L^2(\Omega; \mathbb{R}^m))} \leq C(T, h, \|a\|_{\text{Lip, loc}}) \|u_0^h - v_0^h\|_{L^2(\Omega; \mathbb{R}^m)} \\ \leq C(T, h, \|a\|_{\text{Lip, loc}}) C_1 \max_{i=1, \dots, \mathcal{N}(h)} \|\phi_i^h\|_{L^2(\Omega; \mathbb{R}^m)} \left(\sum_{i=1}^{\mathcal{N}(h)} |\alpha_i^h(0) - \beta_i^h(0)|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Thus the map $\underline{\alpha}^h(0) \in \mathbb{R}^{\mathcal{N}(h)} \mapsto \underline{\alpha}^h(t) \in \mathbb{R}^{\mathcal{N}(h)}$ is (Lipschitz) continuous for all $t \in [0, T]$ and Δt sufficiently small so as to ensure that $0 < 1 - C(h, \|a\|_{\text{Lip, loc}}) \Delta t < 1$, with $h \in (0, h_0]$ fixed.

5. COMPUTATION OF YOUNG MEASURE SOLUTIONS

Theorem 2 and the results of Section 2.3 (in particular, Proposition 1) give criteria under which Young measure solutions exist as limits of solutions to corresponding semi- or fully-discrete problems. However, the results of Section 2.3 and Section 2.3 only give “indirect” information about the Young measure ν through its action on functions belonging to the space $E_{\frac{p}{\alpha}}$ for $\alpha \in (1, \frac{p}{q-1})$. Indeed, note that for any function b which has the form (7)–(9), we can copy the calculations used to deduce (30) to see that $b(Du^h) \rightharpoonup \langle \nu, b \rangle$ in $L^{\tilde{q}}(Q_T; \mathbb{R}^m)$. In this section we shall discuss an algorithm for the numerical approximation of Young measure solutions to systems of the form (1), which allows for a more direct approximation of the Young measure ν . The ideas presented here have been inspired by [10] and [11], where the authors develop and analyze numerical schemes for the approximation of measure-valued solutions to systems of hyperbolic conservation laws. We will show that the algorithms developed there can be adapted to systems of forward-backward parabolic equations of the form (1) exhibiting Young measure solutions. In particular, we shall demonstrate that, for the class of problems under consideration here, there are Young measure solutions for which the measure ν may be constructed as a limit of averaged sums of particular Dirac masses, each of whose support is the value taken by an appropriate approximating solution. Throughout this section, we shall be taking subsequences to approximate Young measure solutions, for which we have not shown a uniqueness result: therefore, as we have already noted in the Introduction, different subsequences could potentially converge to different Young measure solutions.

5.4 Preliminary definitions and results

We recall here a selection of definitions and results from Appendix 1 of [10] that are pertinent to the discussion herein, extending them, where necessary.

Definition 2. *We let $\mathcal{P}(\mathbb{R}^m)$ denote the set of all probability measures on \mathbb{R}^m , and define, for $r > 0$, the subset $\mathcal{P}^r(\mathbb{R}^m)$ of all probability measures $\mu \in \mathcal{P}(\mathbb{R}^m)$ for which $\langle \mu, |\xi|^r \rangle < \infty$, where ξ denotes the identity function.*

We denote by $(\mathfrak{X}, \mathcal{F}, P)$ a probability space, with \mathcal{F} being a σ -algebra of sets on the space \mathfrak{X} , and P being a probability measure, and we let $u : \mathfrak{X} \times Q_T \rightarrow \mathbb{R}^m$ be a random field (that is, a jointly measurable function). Given this, we define the law of u as follows:

$$(38) \quad \mu_{t,x}(B) := P(\{\omega \in \mathfrak{X} : u(\omega; t, x) \in B\}),$$

for Borel subsets $B \subset \mathbb{R}^m$. It is clear that $\mu_{t,x}(\mathbb{R}^m) = 1$ for all $(t, x) \in Q_T$. We have the following result.

Lemma 13. *Let $(\omega; t, x) \in \mathfrak{X} \times Q_T \mapsto u(\omega; t, x) \in \mathbb{R}^m$ be a random field; then, (38) is equivalent to*

$$\langle \mu_{t,x}, g \rangle = \int_{\mathfrak{X}} g(u(\omega; t, x)) \, dP(\omega) \quad \text{for a.e. } (t, x) \in Q_T,$$

for every continuous r -component vector function g , defined on \mathbb{R}^m , which is such that $\int_{\mathfrak{X}} |g(u(\omega; t, x))| \, dP(\omega)$ is finite for a.e. $(t, x) \in Q_T$.

Proof. As the asserted equality is understood componentwise, it suffices to prove it component-by-component. We shall therefore assume in the argument below that $r = 1$, i.e., that g is a continuous mapping from \mathbb{R}^m into \mathbb{R} such that $\int_{\mathfrak{X}} |g(u(\omega; t, x))| \, dP(\omega)$ is finite for a.e. $(t, x) \in Q_T$. Let us first consider the case when the function $g : \mathbb{R}^m \rightarrow \mathbb{R}$ is nonnegative, and take a sequence of simple functions g_n defined on \mathbb{R}^m by

$$g_n(\xi) = \sum_{k=1}^{n2^n+1} b_k^n \chi_{A_k^n}(\xi), \quad \xi \in \mathbb{R}^m, \quad n = 1, 2, \dots,$$

where A_k^n , $k = 1, \dots, n$, are disjoint measurable subsets of \mathbb{R}^m , n is a positive integer, and $b_k^n \in \mathbb{R}$, $k = 1, \dots, n$, are such that $g_n(\xi)$ increases to $g(\xi)$. This can be achieved by defining

$$A_k^n := g^{-1} \left(\left[\frac{k-1}{2^n}, \frac{k}{2^n} \right) \right), \quad 1 \leq k \leq n2^n,$$

and

$$A_{n, n2^n+1} := g^{-1}([n, \infty)).$$

Observe that $g_n(\xi) \leq g(\xi)$ for all $\xi \in \mathbb{R}^m$. Further, for a given $\xi \in \mathbb{R}^m$ we have that $g_n(\xi) = n$ for $g(\xi) \geq n$ whereas for $g(\xi) < n$ we have that

$$g_n(\xi) = \frac{1}{2^n} [2^n g(\xi)],$$

where $[y]$ denotes the greatest integer $\leq y$. As $[2^{n+1}g(\xi)] \geq 2[2^n g(\xi)]$, it follows that $g_{n+1}(\xi) \geq g_n(\xi)$ for all $\xi \in \mathbb{R}^m$. Furthermore, as $[2^n g(\xi)] \geq 2^n g(\xi) - 1$, we have that $g_n(\xi) \geq g(\xi) - 2^{-n}$ as soon as $n > g(\xi)$. Since $g(\xi) \geq g_n(\xi)$ it follows that $\lim_{n \rightarrow \infty} g_n(\xi) = g(\xi)$ for all $\xi \in \mathbb{R}^m$.

Now, let $\mathcal{F}_k^n := \{\omega \in \mathfrak{X} : u(\omega; t, x) \in A_k^n\}$. We observe that

$$\chi_{\mathcal{F}_k^n}(\omega) = \begin{cases} 1 & \text{for } \omega \in \mathcal{F}_k^n \\ 0 & \text{for } \omega \notin \mathcal{F}_k^n \end{cases} = \begin{cases} 1 & \text{for } u(\omega; t, x) \in A_k^n \\ 0 & \text{for } u(\omega; t, x) \notin A_k^n \end{cases} = \chi_{A_k^n}(u(\omega; t, x)).$$

Hence we have that

$$\begin{aligned}
\langle \mu_{t,x}, g \rangle &= \int_{\mathbb{R}^m} g(\xi) d\mu_{t,x}(\xi) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^m} g_n(\xi) d\mu_{t,x}(\xi) = \lim_{n \rightarrow \infty} \sum_{k=1}^{n2^n+1} b_k^n \mu_{t,x}(A_k^n) \\
&= \lim_{n \rightarrow \infty} \sum_{k=1}^{n2^n+1} b_k^n P(\mathcal{F}_k^n) = \lim_{n \rightarrow \infty} \int_{\mathfrak{X}} \sum_{k=1}^{n2^n+1} b_k^n \chi_{\mathcal{F}_k^n}(\omega) dP(\omega) \\
&= \lim_{n \rightarrow \infty} \int_{\mathfrak{X}} \sum_{k=1}^{n2^n+1} b_k^n \chi_{A_k^n}(u(\omega; t, x)) dP(\omega) \\
&= \lim_{n \rightarrow \infty} \int_{\mathfrak{X}} g_n(u(\omega; t, x)) dP(\omega) = \int_{\mathfrak{X}} g(u(\omega; t, x)) dP(\omega),
\end{aligned}$$

where we have applied the Monotone Convergence Theorem in the first and last lines to exchange the limit with the integral over \mathbb{R}^m . From here the extension to any, not necessarily nonnegative, function $g : \mathbb{R}^m \rightarrow \mathbb{R}$ such that $\int_{\mathfrak{X}} |g(u(\omega; t, x))| dP(\omega)$ is finite for a.e. $(t, x) \in Q_T$, is straightforward: we decompose g as $g = g^+ - g^-$, where $g^+ := \frac{1}{2}(|g| + g)$ and $g^- := \frac{1}{2}(|g| - g)$ are (respectively) the positive and negative part of g , and apply the above argument to g^+ and g^- separately, noting that both are nonnegative on \mathbb{R}^m . \square

Remark 5. *It is easy to extend Lemma 13 to functions of the form $g(t, x, u(\omega, t, x))$ where g is a Carathéodory function, i.e., it is continuous in its third variable for a.e. $(t, x) \in Q_T$ and measurable in $(t, x) \in Q_T$ for each value of the third variable, and $\int_{\mathfrak{X}} |g(t, x, u(\omega, t, x))| dP(\omega) < \infty$ for a.e. $(t, x) \in Q_T$.*

We have the following proposition, which shows that the law $\mu = \mu_{t,x}$ defined by (38) is in fact a Young measure. This is Proposition 1 in Appendix 1 of [10].

Proposition 2. *If $u : \mathfrak{X} \times Q_T \rightarrow \mathbb{R}^m$ is jointly measurable, then the law $\mu = \mu_{t,x}$ (cf. (38)) defines a Young measure.*

Proof. The proof is the same as that in [10]. \square

Finally we consider the following result adapted from [10], found therein as Proposition 2 in Appendix 1.

Proposition 3. *For every Young measure $\{\nu_{t,x}\}_{(t,x) \in Q_T} \in \mathcal{P}^r$, $r > 0$, on \mathbb{R}^m there exists a probability space (\mathfrak{X}, F, P) with $P \in \mathcal{P}^r$ and a Borel measurable function $u : \mathfrak{X} \times \Omega \rightarrow \mathbb{R}^m$ such that u has law given by ν . In particular we may choose the probability space to be the Lebesgue measure on the interval $[0, 1)$ with the Borel σ -algebra.*

Proof. The proof of this result relies on using the characterization given in Lemma 13. In [10] a proof is supplied with $g \in C_0(\mathbb{R}^m)$; we extend this result below to cover our particular class of Young measures. We give the proof in the case

of scalar-valued functions u ; the extension to vector-valued functions follows by a component-wise argument.

To extend *Proposition 2* in Appendix 1 of [10] from $g \in C_0(\mathbb{R})$ to $g \in C(\mathbb{R})$ with $\int_{\mathfrak{X}} |g(u(\omega; t, x))| dP(\omega) < \infty$ for a.e. $(t, x) \in Q_T$, let $\varphi \in C_0^\infty(\mathbb{R})$ be such that $0 \leq \varphi \leq 1$, $\varphi(\xi) \equiv 1$ for $|\xi| \leq 1$ and $\varphi(\xi) \equiv 0$ for $|\xi| \geq 2$. Let $g_k(x) := g(x)\varphi(x/k)$. Clearly, g_k is a sequence of functions in $C_0(\mathbb{R})$, $g_k(\xi) = g(\xi)$ if $|\xi| \leq k$, and g_k converges pointwise to g on \mathbb{R} as $k \rightarrow \infty$. Furthermore, $|g_k(x)| \leq |g(x)|$ for all $x \in \mathbb{R}$. Thus, thanks to the Dominated Convergence Theorem, we have that

$$\int_{\mathfrak{X}} g(u(\omega; x)) dP(\omega) = \lim_{k \rightarrow \infty} \int_{\mathfrak{X}} g_k(u(\omega; x)) dP(\omega) = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} g_k(\xi) d\nu_x(\xi) = \int_{\mathbb{R}} g(\xi) d\nu_x(\xi).$$

We remark that progressing from the left-hand side of the second equality above to the right-hand side of that equality is precisely the result of *Proposition 2* in Appendix 1 of [10]. \square

Finally, we state and prove a standard lemma regarding independent and identically distributed random variables, and their images under measurable functions.

Lemma 14. *Let (\mathfrak{X}, F, P) be a probability space and let $f : \mathfrak{X} \rightarrow \mathbb{R}^m$ be a measurable function. If two random variables, Y_1 and Y_2 , defined on \mathfrak{X} are independent and identically distributed, then the random variables $f(Y_1)$ and $f(Y_2)$ are independent and identically distributed on \mathfrak{X} .*

Proof. We first prove that the random variables $f(Y_1)$ and $f(Y_2)$ are identically distributed. We let $\mathcal{A} \subset \mathbb{R}^m$ be a measurable set. It then follows from the measurability of f that, for all $\omega \in \mathfrak{X}$,

$$P(f(Y_1(\omega)) \in \mathcal{A}) = P(Y_1(\omega) \in f^{-1}(\mathcal{A})) = P(Y_2(\omega) \in f^{-1}(\mathcal{A})) = P(f(Y_2(\omega)) \in \mathcal{A}).$$

To show independence, we let \mathcal{A}_1 and \mathcal{A}_2 be measurable sets and calculate

$$\begin{aligned} P((f(Y_1(\omega)) \in \mathcal{A}_1) \cap (f(Y_2(\omega)) \in \mathcal{A}_2)) &= P((Y_1(\omega) \in f^{-1}(\mathcal{A}_1)) \cap (Y_2(\omega) \in f^{-1}(\mathcal{A}_2))) \\ &= P(Y_1(\omega) \in f^{-1}(\mathcal{A}_1)) P(Y_2(\omega) \in f^{-1}(\mathcal{A}_2)) \\ &= P(f(Y_1(\omega)) \in \mathcal{A}_1) P(f(Y_2(\omega)) \in \mathcal{A}_2) \end{aligned}$$

for all $\omega \in \mathfrak{X}$. That completes the proof of the lemma. \square

5.5 An overview of the algorithm

The algorithms discussed below involve generating (at random) a large set of initial data, evolving each of these forward under the solution operator of a semi- or fully-discrete numerical method, and considering the arithmetic average of the resulting functions at a fixed time, before completing various limit passages. The algorithms considered here have been motivated by similar algorithms used in [10] and [11] for the approximation of measure-valued solutions to hyperbolic problems. To this end, let $Y(\Omega, \mathbb{R}^m)$ denote the set of all Young measures from Ω to \mathbb{R}^m .

Algorithm A: Let the initial data for an underlying time-dependent PDE be given as a Young measure $\sigma \in Y(\Omega; \mathbb{R}^m)$ and let Δ denote a PDE discretization parameter associated with a certain numerical scheme (in [10] and [11] a finite difference scheme is utilized, and so Δ would correspond to a vector of grid sizes in the various co-ordinate directions).

Step 1: Let $u_0 : \omega \in \mathfrak{X} \mapsto u_0(\omega; \cdot) \in L^p(\Omega; \mathbb{R}^m)$ be a random field on a probability space (\mathfrak{X}, F, P) with law σ_x ; i.e., $\sigma_x(E) = P(u_0(\omega; x) \in E)$ for all Borel sets $E \subset \mathbb{R}^m$ and $x \in \Omega$.

Step 2: Evolve the initial random field by applying a suitable numerical scheme, with solution map \mathbf{S}_t^Δ , to the initial data $u_0(\omega; \cdot)$ for every $\omega \in \mathfrak{X}$, obtaining an approximate random field $u^\Delta(\omega; \cdot, t) := \mathbf{S}_t^\Delta u_0(\omega; \cdot)$, $t \in (0, T]$.

Step 3: Define the approximate measure-valued solution μ^Δ as the law of u^Δ with respect to P , that is, for all Borel sets $E \subset \mathbb{R}^m$ and $(t, x) \in Q_T$,

$$\mu_{t,x}^\Delta(E) = P(u^\Delta(\omega; t, x) \in E).$$

Note that we should expect a slight difference compared with [10] and [11] in what we are trying to do, as the Young measures appearing in our context are generated by sequences of gradients. Next, we need a method to approximate the random field $(\omega; \cdot, \cdot) \mapsto u^\Delta(\omega; \cdot, \cdot)$, as performing these computations for every $\omega \in \mathfrak{X}$ would be infeasible. This is *Algorithm 4.3* in [11].

Algorithm B: Let, as above, Δ denote a discretization parameter, and let $M \in \mathbb{N}$. Let σ^Δ be the initial Young measure.

Step 1: From some probability space (\mathfrak{X}, F, P) draw M independent and identically distributed random fields $u_0^{\Delta,1}, \dots, u_0^{\Delta,M}$ all with the same law σ^Δ .

Step 2: For each $k \in \{1, \dots, M\}$ and for a fixed $\omega \in \mathfrak{X}$ approximate the solution to the PDE using the solution operator with initial data $u_0^{\Delta,k}(\omega)$; denote $u^{\Delta,k}(\omega; \cdot, t) := \mathbf{S}_t^\Delta u_0^{\Delta,k}(\omega; \cdot)$.

Step 3: Define the approximate measure-valued solution by

$$\mu_{t,x}^{\Delta,M} := \frac{1}{M} \sum_{k=1}^M \delta_{u^{\Delta,k}(\cdot; t, x)}.$$

Results from [11] (see *Theorem 4.5*, *Theorem 5.1* and *Corollary 5.4*) then guarantee convergence of the sequence of approximate measure-valued solutions to a measure-valued solution as one passes to the limit (diagonally). In this section we discuss in more detail these algorithms and their adaptation to the systems of parabolic PDEs which we are interested in.

5.6 Modifications for systems of forward-backward parabolic PDEs

We now describe the necessary changes to Algorithms A and B described above in order to be able to apply them to the problem (1)–(9).

Below we formulate the analogues of Algorithm A and Algorithm B, which are needed for our parabolic problem. Since in our case the solution is a Sobolev function, we make here the additional restriction that we only consider atomic initial data: that is, our initial datum is assumed to be given by a function $u_0 \in L^2(\Omega; \mathbb{R}^m)$, which we view as the atomic Young measure $\delta_{u_0} = \delta_{u_0(x)}$, $x \in \Omega$.

Algorithm C: Let the initial datum for the problem (1)–(9) be given as a function $u_0 \in L^2(\Omega; \mathbb{R}^m)$ and let $h \in (0, h_0]$ be the spatial grid size parameter.

Step 1: Let $v : \mathfrak{X} \rightarrow L^2(\Omega; \mathbb{R}^m)$ be a random field on a probability space (\mathfrak{X}, F, P) , and discretize v by a finite element approximation of random fields $v^h \in V_m^h$, so that $\|v^h(\omega; \cdot)\|_{L^2(\Omega; \mathbb{R}^m)} \leq 1$ for P -a.e. $\omega \in \mathfrak{X}$. Discretize u_0 by a finite element approximation $u_0^h \in V_m^h$, and then perturb this discretization by defining

$$u_0^{h,\epsilon}(\omega; x) := u_0^h(x) + \epsilon v^h(\omega; x), \quad \omega \in \mathfrak{X}, \quad x \in \Omega,$$

where $\epsilon \in (0, 1]$, and let $\sigma^{h,\epsilon}$ be the law of $u_0^{h,\epsilon}$; i.e., $\sigma_x^{h,\epsilon}(E) = P(u_0^{h,\epsilon}(\omega; x) \in E)$ for $\omega \in \mathfrak{X}$ and $x \in \Omega$.

Step 2: Evolve the initial random field by applying a suitable numerical scheme, with solution map \mathbf{S}_t , to the initial datum $u_0^{h,\epsilon}(\omega)$ for every $\omega \in \mathfrak{X}$, obtaining an approximate random field $u^{h,\epsilon}(\omega, t, \cdot) := \mathbf{S}_t u_0^{h,\epsilon}(\omega; \cdot)$, $t \in (0, T]$.

Step 3: Define the approximate Young measure solution $\mu^{h,\epsilon}$ as the law of $u^{h,\epsilon}$ with respect to P , that is, for all Borel sets $E \subset \mathbb{R}^m$,

$$\mu_{t,x}^{h,\epsilon}(E) = P(u^{h,\epsilon}(\omega; t, x) \in E), \quad \omega \in \mathfrak{X}, \quad (t, x) \in Q_T,$$

and define, analogously, $\nu^{h,\epsilon}$ to be the law of $Du^{h,\epsilon}$ with respect to P .

As before we need a method to approximate the random field $u_0^\epsilon(\omega; x)$, as well as the measures $\mu^{h,\epsilon}$ and $\nu^{h,\epsilon}$.

Algorithm D: Let the initial datum for the problem (1)–(9) be given as a function $u_0 \in L^2(\Omega; \mathbb{R}^m)$, let $h \in (0, h_0]$ be the spatial grid size parameter, and let $M \in \mathbb{N}$.

Step 1: From some probability space (\mathfrak{X}, F, P) draw M independent and identically distributed random fields $v^{h,1}, \dots, v^{h,M}$ such that $\|v^{h,i}(\omega; \cdot)\|_{L^2(\Omega; \mathbb{R}^m)} \leq 1$, for P -a.e. $\omega \in \mathfrak{X}$ and for all $i = 1, \dots, M$, and such that $u_0^{h,1,\epsilon}, \dots, u_0^{h,M,\epsilon}$ all have the same law $\sigma^{h,\epsilon}$, where

$$u_0^{h,k,\epsilon}(\omega; x) = u_0^h(x) + \epsilon v^{h,k}(\omega; x), \quad \omega \in \mathfrak{X}, \quad x \in \Omega,$$

with $\epsilon \in (0, 1]$. We make explicit here that $\|u_0^{h,k,\epsilon}(\omega; \cdot)\|_{L^2(\Omega; \mathbb{R}^m)} \leq \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)} + \epsilon$ for all $\omega \in \mathfrak{X}$, so that $\|u_0^{h,k,\epsilon}(\omega; \cdot)\|_{L^2(\Omega; \mathbb{R}^m)}$ can be assumed to be bounded independent of k , ϵ , and h (thanks to the assumed strong convergence of u_0^h to u_0 in $L^2(\Omega; \mathbb{R}^m)$ and because $\epsilon \in (0, 1]$).

Step 2: For each k and for a fixed $\omega \in \mathfrak{X}$ approximate the solution to the initial boundary-value problem (1)–(9) under consideration using the solution operator with initial data $u_0^{h,k,\epsilon}(\omega)$; denote $u^{h,k,\epsilon}(\omega; t, \cdot) = \mathbf{S}_t u_0^{h,k,\epsilon}(\omega; \cdot, \cdot)$.

Step 3: Define the approximate Young measure solution $(\mu_{t,x}^{h,M,\epsilon}, \nu_{t,x}^{h,M,\epsilon})$ by

$$\mu_{t,x}^{h,M,\epsilon} := \frac{1}{M} \sum_{k=1}^M \delta_{u^{h,k,\epsilon}(\cdot; t, x)} \quad \text{and} \quad \nu_{t,x}^{h,M,\epsilon} := \frac{1}{M} \sum_{k=1}^M \delta_{D u^{h,k,\epsilon}(\cdot; t, x)}.$$

5.7 Applying the algorithm to construct Young measure solutions

In what follows we will show that Algorithm D described above converges to a Young measure solution of the system of PDEs (1)–(9) under consideration as we pass $\Delta t \rightarrow 0_+$, $M \rightarrow \infty$, $h, \epsilon \rightarrow 0_+$ (in that order).

Remark 6. *The parameter ω taken from the probability space \mathfrak{X} is intended to represent the “seed” in the numerical algorithm. Different seeds $\omega \in \mathfrak{X}$ will give rise to different approximate solutions. What will be shown in our analysis of the limit process $\Delta t \rightarrow 0_+$, $M \rightarrow \infty$, $h, \epsilon \rightarrow 0_+$ is that one can pass to the limit in M (along a subsequence) to obtain convergence to a quantity which is independent of $\omega \in \mathfrak{X}$, for P -a.e. $\omega \in \mathfrak{X}$.*

Step 1: Passage to the limits $\Delta t \rightarrow 0_+$, $M \rightarrow \infty$. We shall use the semidiscrete numerical scheme discussed in Section 2.3, augmented with the time-stepping procedure discussed in Section 2.3. By Lemma 1 we have that, for any $\omega \in \mathfrak{X}$, any $\epsilon > 0$, any positive integer k and any $h \in (0, h_0]$ and $\Delta t = T/N$, $N \in \mathbb{N}$, we can construct, via the time-stepping procedure from Section 2.3, a function $u_{\Delta t}^{h,k,\epsilon}(\omega; \cdot, \cdot)$, which for any $h \in (0, h_0]$ and any $\Delta t > 0$ (sufficiently small, for a given fixed $h \in (0, h_0]$) uniquely solves the fully discrete scheme and depends continuously on the discretized initial condition $u_0^{h,k,\epsilon}$. For the rest of the analysis we shall assume for the sake of brevity that the passage to the limit $\Delta t \rightarrow 0_+$ has already been made for $\omega \in \mathfrak{X}$ and $h \in (0, h_0]$ fixed, by repeating the analysis performed in Section 2.3, and we shall therefore consider the limiting function $u^{h,k,\epsilon}(\omega; \cdot, \cdot)$ resulting from the passage to the limit $\Delta t \rightarrow 0_+$ instead of $u_{\Delta t}^{h,k,\epsilon}(\omega; \cdot, \cdot)$. Passage to the limit $\Delta t \rightarrow 0_+$, for reasons that will be explained in due course, must happen before the limit passage $M \rightarrow \infty$. In particular we remark that as we are at this stage considering finitely many perturbations of the discretized initial condition, the existence of a suitable subsequence $\Delta t_j \rightarrow 0_+$, which works for each $k \in \{1, \dots, M\}$, is straightforward.

For each initial datum $u_0^{h,k,\epsilon}(\omega; \cdot)$ (for fixed $\omega \in \mathfrak{X}$ and $k \in \{1, \dots, M\}$), we have from (18) the following energy estimate:

$$(39) \quad \begin{aligned} \|u^{h,k,\epsilon}(\omega)\|_{L^\infty(0,T;L^2(\Omega;\mathbb{R}^m))}^2 + \sum_{i=1}^m \|Du_i^{h,k,\epsilon}(\omega)\|_{L^{p_i}(0,T;W_0^{1,p_i}(\Omega;\mathbb{R}^{m \times n}))}^{p_i} \\ \leq c(1 + \|u_0^{h,k,\epsilon}(\omega)\|_{L^2(\Omega;\mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T;\mathbb{R}^m)}^{p'}). \end{aligned}$$

Furthermore, (22) implies, with $\hat{q} := \max\{q, 2\}$ and $\hat{q}' = \frac{\hat{q}}{\hat{q}-1}$, that

$$(40) \quad \|a(Du^{h,k,\epsilon})(\omega)\|_{L^{\hat{q}'}(Q_T;\mathbb{R}^{m \times n})} \leq C$$

and (24) yields

$$(41) \quad \|\partial_t u^{h,k,\epsilon}(\omega)\|_{L^{\hat{q}'}(0,T;W^{-1,\hat{q}'}(\Omega;\mathbb{R}^m))} \leq C,$$

where C is a positive constant that is independent of h, k and ϵ . This follows from the form of the perturbed initial condition in Algorithm D, which guarantees that the perturbed initial data are bounded uniformly in h, k and ϵ . Summing the equations satisfied by $u^{h,k,\epsilon}(\omega; \cdot, \cdot)$ over $k = 1, \dots, M$ and dividing by M we see that the functions $u^{h,k,\epsilon}$ satisfy, for all $\psi^h \in L^q(0, T; V_m^h)$, the equality:

$$(42) \quad \begin{aligned} \int_0^T \int_\Omega \frac{1}{M} \sum_{k=1}^M \partial_t u^{h,k,\epsilon}(\omega; t, x) \cdot \psi^h(t, x) + \frac{1}{M} \sum_{k=1}^M a(Du^{h,k,\epsilon}(\omega; t, x)) : D\psi^h(t, x) \, dx \, dt \\ + \int_0^T \int_\Omega \frac{1}{M} \sum_{k=1}^M Bu^{h,k,\epsilon}(\omega; t, x) \cdot \psi^h(t, x) \, dx \, dt = \int_0^T \int_\Omega F \cdot \psi^h \, dx \, dt. \end{aligned}$$

Integration by parts in time then yields, for all $\psi^h \in W_0^{1,1}(0, T; V_m^h)$,

$$(43) \quad \begin{aligned} \int_0^T \int_\Omega \frac{1}{M} \sum_{k=1}^M u^{h,k,\epsilon}(\omega; t, x) \cdot (\partial_t \psi^h(t, x) - B^T \psi^h(t, x)) \, dx \, dt \\ - \int_0^T \int_\Omega \frac{1}{M} \sum_{k=1}^M a(Du^{h,k,\epsilon}(\omega; t, x)) : D\psi^h(t, x) - F(t, x) \cdot \psi^h(t, x) \, dx \, dt = 0. \end{aligned}$$

We now wish to let $M \rightarrow \infty$ in (43). To this end, we need to prove certain convergence results that will enable passage to this limit. We begin by considering the sequence of functions $\{\frac{1}{M} \sum_{k=1}^M a(Du^{h,k,\epsilon}(\omega; \cdot, \cdot))\}_{M \geq 1}$, for a fixed $\omega \in \mathfrak{X}$, which is, thanks to (22), bounded in the $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$ norm, uniformly in $h \in (0, h_0]$ and $\epsilon \in (0, 1]$, and as such there is a subsequence $\{M_n\}_{n \geq 1}$ and a function $\chi^{h,\epsilon}(\omega; \cdot, \cdot) \in L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$ such that $\frac{1}{M_n} \sum_{k=1}^{M_n} a(Du^{h,k,\epsilon}(\omega; \cdot, \cdot)) \rightarrow \chi^{h,\epsilon}(\omega; \cdot, \cdot)$ in $L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$, as $n \rightarrow \infty$. We aim to identify this function $\chi^{h,\epsilon}(\omega; \cdot, \cdot)$ for $\omega \in \mathfrak{X}$, by identifying the limit in some weaker space and appealing to the uniqueness of weak limits.

Let us consider the initial datum $u_0^{h,\epsilon}(\omega; x) = u_0^h(x) + \epsilon v^h(\omega; x)$ that is independent of the initial data $u_0^{h,1,\epsilon}, \dots, u_0^{h,M,\epsilon}$, and has the same distribution as

these initial data (meaning in particular that the law of $u_0^{h,\epsilon}$ is given by $\sigma^{h,\epsilon}$), and satisfies $\|v^h(\omega; \cdot)\|_{L^2(\Omega; \mathbb{R}^m)} \leq 1$ for P -almost every $\omega \in \mathfrak{X}$. We evolve this according to the algorithm for each $\omega \in \mathfrak{X}$ to obtain a function $u^{h,\epsilon}(\omega; t, x)$. We denote the law of $u^{h,\epsilon}$ by $\mu^{h,\epsilon}$, and similarly denote the law of $Du^{h,\epsilon}$ by $\nu^{h,\epsilon}$. From Theorem 2 we see that the following regularity results hold:

$$(44) \quad u^{h,\epsilon} \in L^\infty \left(\mathfrak{X}; \bigotimes_{i=1}^m L^{p_i}(0, T; W_0^{1,p_i}(\Omega; \mathbb{R}^m)) \right) \cap L^\infty(\mathfrak{X} \times (0, T); L^2(\Omega; \mathbb{R}^m)),$$

$$(45) \quad \partial_t u^{h,\epsilon} \in L^\infty(\mathfrak{X}; L^{\hat{q}}(0, T; W^{-1,\hat{q}}(\Omega; \mathbb{R}^m))),$$

$$(46) \quad a(Du^{h,\epsilon}) \in L^\infty(\mathfrak{X}; L^{\hat{q}}(Q_T; \mathbb{R}^{m \times n})).$$

Analogously to (39)–(41), the norms of these functions in the respective spaces can be bounded by a positive constant C , independent of h and ϵ . Next, we formulate an analogue of *Theorem 11* from [10].

Lemma 15. *Let $\alpha \in (1, \frac{p}{q-1})$ and suppose that $\gamma \in L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})$ (understood as an $m \times n$ matrix-valued function). Along the subsequence $\{M_n\}_{n \geq 1}$ we have*

$$\int_0^T \int_\Omega \langle \nu_{\omega;t,x}^{h,M_n,\epsilon}, \gamma(t, x, \cdot) \rangle dx dt \rightarrow \int_0^T \int_\Omega \langle \nu_{t,x}^{h,\epsilon}, \gamma(t, x, \cdot) \rangle dx dt,$$

in $L^2(\mathfrak{X})$ as $n \rightarrow \infty$. In particular, for each $\gamma \in L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})$ there is a subsequence of M_n such that for P -a.e. $\omega \in \mathfrak{X}$ the convergence is pointwise.

Proof. Given a random field $\eta : \mathfrak{X} \rightarrow L^1(Q_T)$ we define its expectation with respect to the probability measure P as

$$\mathbb{E}(\eta) := \int_{\mathfrak{X}} \eta(\omega) dP(\omega).$$

Similarly to [10] we define, for $k = 1, \dots, M$, $i = 1, \dots, m$ and $j = 1, \dots, n$, the following quantities:

$$G_{i,j}(\omega) = \int_0^T \int_\Omega \gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x)) dx dt,$$

and

$$G_{i,j}^k(\omega) = \int_0^T \int_\Omega \gamma_{i,j}(t, x, Du^{h,k,\epsilon}(\omega; t, x)) dx dt.$$

We need to show that $G_{i,j}^1, \dots, G_{i,j}^{M_n}$ are independent and identically distributed.

We begin by remarking that as the initial data $u_0^{h,1,\epsilon}, \dots, u_0^{h,M_n,\epsilon}$ are independent and identically distributed (in the variable $\omega \in \mathfrak{X}$), and as the solution operator \mathbf{S}_t mapping an initial condition \tilde{u}_0 to $\tilde{u}(t, \cdot)$ is a continuous map (see Lemma 1), the functions $u^{h,k,\epsilon}$ are all independent and identically distributed as

well (see Lemma 14). Furthermore, by the remarks following the statement of Theorem 1, the functions $Du^{h,k,\epsilon}$ are also independent and identically distributed.

Recall that the functions $\gamma_{i,j}$ are measurable, they are continuous in their third argument, and satisfy the growth rate coming from the space $E_{\frac{p}{\alpha}}$. This implies that the functions $\gamma_{i,j}(t, x, Du^{h,k,\epsilon}(\omega; t, x))$ are independent and identically distributed with respect to $\omega \in \mathfrak{X}$. From Tonelli's Theorem it then follows that the functions $G_{i,j}^k$ are measurable in ω , and thus by Lemma 14 are also independent and identically distributed in $\omega \in \mathfrak{X}$.

We now compute the L^2 error with respect to the probability measure P of the following quantity:

$$\left(\mathbb{E}(G_{i,j}(\omega)) - \frac{1}{M_n} \sum_{k=1}^{M_n} G_{i,j}^k(\omega) \right)^2,$$

and as in [10], using the fact that the random variables $G_{i,j}^k$ are independent and identically distributed, we can reduce this to

$$(47) \quad \mathbb{E} \left(\left(\mathbb{E}(G_{i,j}(\omega)) - \frac{1}{M_n} \sum_{k=1}^{M_n} G_{i,j}^k(\omega) \right)^2 \right) = \frac{1}{M_n} (\mathbb{E}((G_{i,j})^2) - (\mathbb{E}(G_{i,j}))^2) \leq \frac{1}{M_n} \mathbb{E}((G_{i,j})^2).$$

Given $\alpha \in (1, \frac{p}{q-1})$ we have that

$$\begin{aligned} \mathbb{E}((G_{i,j})^2) &= \int_{\mathfrak{X}} \left(\int_0^T \int_{\Omega} \gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x)) \, dx \, dt \right)^2 \, dP(\omega) \\ &\leq \int_{\mathfrak{X}} \|\gamma_{i,j}\|_{L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})}^2 \|1 + |Du^{h,\epsilon}(\omega)|_{\frac{p}{\alpha}}\|_{L^{\alpha}(Q_T)}^2 \, dP(\omega). \end{aligned}$$

Next, we can bound the norm $\|1 + |Du^{h,\epsilon}(\omega)|_{\frac{p}{\alpha}}\|_{L^{\alpha}(Q_T)}$ independently of h , ϵ and ω by using the energy estimate (18) (analogously to (40)). Thus, since P is a probability measure, we have that $\mathbb{E}((G_{i,j})^2)$ is bounded, and so the right-hand side of (47) converges to zero as $M_n \rightarrow \infty$.

We wish to apply Lemma 13 (note also Remark 5). In order to do this we must show that

$$\int_{\mathfrak{X}} |\gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x))| \, dP(\omega) < \infty \quad \text{for a.e. } (t, x) \in Q_T.$$

Note that if we show that

$$\int_{\mathfrak{X}} \int_0^T \int_{\Omega} |\gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x))| \, dx \, dt \, dP(\omega) < \infty,$$

then the desired bound is true for almost every $(t, x) \in Q_T$, which is sufficient for our purposes. We see, using that $\gamma \in L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})$ with $\alpha \in (1, \frac{p}{q-1})$, that

$$\begin{aligned}
& \int_{\mathfrak{X}} \int_0^T \int_{\Omega} |\gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x))| dx dt dP(\omega) \\
&= \int_{\mathfrak{X}} \int_0^T \int_{\Omega} |\gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x))| \frac{1 + |Du^{h,\epsilon}(\omega; t, x)|^{\frac{p}{\alpha}}}{1 + |Du^{h,\epsilon}(\omega; t, x)|^{\frac{p}{\alpha}}} dx dt dP(\omega) \\
&\leq \int_{\mathfrak{X}} \int_0^T \int_{\Omega} \|\gamma_{i,j}(t, x, \cdot)\|_{E_{\frac{p}{\alpha}}} (1 + |Du^{h,\epsilon}(\omega; t, x)|^{\frac{p}{\alpha}}) dx dt dP(\omega) \\
&\leq \int_{\mathfrak{X}} \|\gamma_{i,j}\|_{L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})} \|1 + |Du^{h,\epsilon}(\omega; \cdot, \cdot)|^{\frac{p}{\alpha}}\|_{L^{\alpha}(Q_T)} dP(\omega) \\
&\leq \|\gamma_{i,j}\|_{L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})} \int_{\mathfrak{X}} |Q_T|^{\frac{1}{\alpha}} + \|Du^{h,\epsilon}(\omega; \cdot, \cdot)\|_{L^p(Q_T; \mathbb{R}^{m \times n})}^{\frac{p}{\alpha}} dP(\omega) \\
&\leq \|\gamma_{i,j}\|_{L^{\alpha'}(Q_T; E_{\frac{p}{\alpha}})} \left(|Q_T|^{\frac{1}{\alpha}} + \|Du^{h,\epsilon}\|_{L^{\infty}(\mathfrak{X}; L^p(Q_T; \mathbb{R}^{m \times n}))}^{\frac{p}{\alpha}} \right),
\end{aligned}$$

which is finite by the assumptions on γ and by (44). Then, by applying Lemma 13, we see that:

$$\begin{aligned}
\mathbb{E}(G_{i,j}(\omega)) &= \int_{\mathfrak{X}} \int_0^T \int_{\Omega} \gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x)) dx dt dP(\omega) \\
&= \int_0^T \int_{\Omega} \left(\int_{\mathfrak{X}} \gamma_{i,j}(t, x, Du^{h,\epsilon}(\omega; t, x)) dP(\omega) \right) dx dt \\
&= \int_0^T \int_{\Omega} \langle \nu_{i,j;t,x}^{h,\epsilon}, \gamma_{i,j}(t, x, \cdot) \rangle dx dt.
\end{aligned}$$

Repeating these calculations for all components $\gamma_{i,j}$ gives us the convergence of

$$\int_0^T \int_{\Omega} \langle \nu_{\omega;t,x}^{h,M_n,\epsilon}, \gamma(t, x, \cdot) \rangle dx dt \rightarrow \int_0^T \int_{\Omega} \langle \nu_{t,x}^{h,\epsilon}, \gamma(t, x, \cdot) \rangle dx dt,$$

in $L^2(\mathfrak{X})$ as $M_n \rightarrow \infty$, from which we deduce the existence of a subsequence (not relabelled) converging for P -almost every $\omega \in \mathfrak{X}$ using standard results in measure theory. \square

What we would like to conclude is that by the above theorem we have, for P -almost every $\omega \in \mathfrak{X}$, the weak-star convergence of $\nu_{\omega}^{h,M_n,\epsilon} = \nu_{\omega,\cdot,\cdot}^{h,M_n,\epsilon}$ to $\nu^{h,\epsilon} = \nu_{\cdot,\cdot}^{h,\epsilon}$ as $n \rightarrow \infty$. This does not follow immediately, as a-priori there is *no reason* why the subsequence of M_n along which we have pointwise convergence should be *independent* of the function γ (a point not clarified in [10] or [11]). To get around this we note that as $E_{\frac{p}{\alpha}}$ is separable (see [17]), the space $L^{\alpha}(Q_T; E_{\frac{p}{\alpha}})$ is also separable for our choice of α . Therefore there is a countable dense subset $\{\gamma_i\}_{i=1}^{\infty}$. The desired weak-star convergence then follows from a diagonal argument.

Recall from the calculations prior to Lemma 15 that, as $M_n \rightarrow \infty$ ($n \rightarrow \infty$),

(48)

$$\int_0^T \int_{\Omega} \frac{1}{M_n} \sum_{k=1}^{M_n} a(Du^{h,k,\epsilon}(\omega; t, x)) : D\psi(t, x) \, dx \, dt \rightarrow \int_0^T \int_{\Omega} \chi^{h,\epsilon}(\omega; t, x) : D\psi(t, x) \, dx \, dt$$

for every $\psi \in L^{\hat{q}}([0, T]; W_0^{1,\hat{q}}(\Omega; \mathbb{R}^m))$, and for all $\omega \in \mathfrak{X}$. Furthermore, by Lemma 15, there is a subsequence of M_n such that, as $M_n \rightarrow \infty$ ($n \rightarrow \infty$),

$$(49) \quad \int_0^T \int_{\Omega} \frac{1}{M_n} \sum_{k=1}^{M_n} a(Du^{h,k,\epsilon}(\omega; t, x)) : D\psi(t, x) \, dx \, dt \rightarrow \int_0^T \int_{\Omega} \langle \nu_{t,x}^{h,\epsilon}, a \rangle : D\psi(t, x) \, dx \, dt$$

for every $\psi \in L^{\alpha}(0, T; W_0^{1,\alpha'}(\Omega; \mathbb{R}^m))$, and for P -a.e. $\omega \in \mathfrak{X}$.

As both (48) and (49) hold for $\psi \in L^{\max\{\hat{q}, \alpha'\}}(0, T; W_0^{1,\max\{\hat{q}, \alpha'\}}(\Omega; \mathbb{R}^m))$, by the uniqueness of the weak limit we can identify $\chi^{h,\epsilon}(\omega; \cdot, \cdot) = \langle \nu_{t,x}^{h,\epsilon}, a \rangle$ for P -a.e. $\omega \in \mathfrak{X}$. As the right-hand side of this equality is independent of $\omega \in \mathfrak{X}$, it follows that $\chi^{h,\epsilon}$ is also independent of ω , and we shall therefore suppress the dependence on ω in our notation and write $\chi^{h,\epsilon}(t, x)$ instead of $\chi^{h,\epsilon}(\omega; t, x)$, noting that $\chi^{h,\epsilon} = \langle \nu_{t,x}^{h,\epsilon}, a \rangle \in L^{\hat{q}'}(Q_T; \mathbb{R}^{m \times n})$. Thus, by combining (48) and (49), which guarantee weak convergence in $L^{\hat{q}'}(0, T; W_0^{1,\hat{q}'}(\Omega; \mathbb{R}^m))$ and $L^{\alpha}(0, T; W_0^{1,\alpha'}(\Omega; \mathbb{R}^m))$, respectively, for P -a.e. $\omega \in \mathfrak{X}$, and therefore, by uniqueness of the weak limit, also in the function space $L^{\max\{\hat{q}', \alpha\}}(0, T; W_0^{1,\max\{\hat{q}', \alpha\}}(\Omega; \mathbb{R}^m))$, for P -a.e. $\omega \in \mathfrak{X}$, we have that

$$(50) \quad \int_0^T \int_{\Omega} \frac{1}{M_n} \sum_{k=1}^{M_n} a(Du^{h,k,\epsilon}(\omega; t, x)) : D\psi(t, x) \, dx \, dt \rightarrow \int_0^T \int_{\Omega} \langle \nu_{t,x}^{h,\epsilon}, a \rangle : D\psi(t, x) \, dx \, dt$$

for all $\psi \in L^{\min\{\hat{q}, \alpha'\}}(0, T; W_0^{1,\min\{\hat{q}, \alpha'\}}(\Omega; \mathbb{R}^m))$, for P -a.e. $\omega \in \mathfrak{X}$.

The terms involving the measures $\mu^{h,\epsilon}$ are treated completely analogously. Thus, by passing to the limit $M_n \rightarrow \infty$ ($n \rightarrow \infty$) in (43), we have that for all $\psi^h \in W_0^{1,1}(0, T; V_m^h)$ the following identity is satisfied:

$$\int_0^T \int_{\Omega} \langle \mu_{t,x}^{h,\epsilon}, \xi \rangle \cdot (\partial_t \psi^h(t, x) - B^T \psi^h(t, x)) - \langle \nu_{t,x}^{h,\epsilon}, a \rangle : D\psi^h(t, x) + F(t, x) \cdot \psi^h(t, x) \, dx \, dt = 0.$$

Step 2: Passing $h, \epsilon \rightarrow 0_+$. By Lemma 15 and noting that $a \in E_{\frac{p}{\alpha}}$, it follows that $\langle \nu_{\omega, \cdot, \cdot}^{h, M_n, \epsilon}, a \rangle \rightharpoonup \langle \nu_{\cdot, \cdot, \cdot}^{h, \epsilon}, a \rangle$ weakly in $L^{\alpha}(Q_T; \mathbb{R}^{m \times n})$, for $\alpha \in (1, \frac{p}{q-1})$, and for P -a.e. $\omega \in \mathfrak{X}$. Thus, by weak lower-semicontinuity of the norm function, we have that $\|\langle \nu_{\cdot, \cdot, \cdot}^{h, \epsilon}, a \rangle\|_{L^{\alpha}(Q_T; \mathbb{R}^{m \times n})} \leq \liminf_{n \rightarrow \infty} \|\langle \nu_{\omega, \cdot, \cdot}^{h, M_n, \epsilon}, a \rangle\|_{L^{\alpha}(Q_T; \mathbb{R}^{m \times n})}$ for P -a.e. $\omega \in \mathfrak{X}$.

Let $1 < \alpha < \frac{p}{q-1}$ and consider the function $\tilde{a}(\xi) = 1 + |\xi|^{\frac{p}{\alpha}}$; then, $\tilde{a} \in E_{p/\alpha}$. It then follows from Lemma 15 that $\langle \nu_{\omega, \cdot, \cdot}^{h, M_l, \epsilon}, \tilde{a} \rangle \rightharpoonup \langle \nu_{\cdot, \cdot, \cdot}^{h, \epsilon}, \tilde{a} \rangle$ in $L^{\alpha}(Q_T; \mathbb{R}^{m \times n})$ as $l \rightarrow \infty$ for P -a.e. $\omega \in \mathfrak{X}$. Then, by weak lower-semicontinuity of the norm function $\|\cdot\|_{L^{\alpha}(Q_T; \mathbb{R}^{m \times n})}$, we have that

$$\begin{aligned} \|\langle \nu_{\cdot, \cdot, \cdot}^{h, \epsilon}, \tilde{a} \rangle\|_{L^{\alpha}(Q_T; \mathbb{R}^{m \times n})} &\leq \liminf_{l \rightarrow \infty} \|\langle \nu_{\omega, \cdot, \cdot}^{h, M_l, \epsilon}, \tilde{a} \rangle\|_{L^{\alpha}(Q_T; \mathbb{R}^{m \times n})} \\ &= \liminf_{l \rightarrow \infty} \left\| \frac{1}{M_l} \sum_{k=1}^{M_l} (1 + |Du^{h,k,\epsilon}(\omega; \cdot, \cdot)|^{\frac{p}{\alpha}}) \right\|_{L^{\alpha}(Q_T)} \\ &\leq \liminf_{l \rightarrow \infty} \frac{1}{M_l} \sum_{k=1}^{M_l} \left\| 1 + |Du^{h,k,\epsilon}(\omega; \cdot, \cdot)|^{\frac{p}{\alpha}} \right\|_{L^{\alpha}(Q_T)} \end{aligned}$$

$$\begin{aligned}
&\leq \liminf_{l \rightarrow \infty} |Q_T|^{\frac{1}{\alpha}} + \frac{1}{M_l} \sum_{k=1}^{M_l} \left\| Du^{h,k,\epsilon}(\omega; \cdot, \cdot) \right\|_{L^p(Q_T; \mathbb{R}^{m \times n})}^{\frac{p}{\alpha}} \\
(51) \quad &\leq \liminf_{l \rightarrow \infty} |Q_T|^{\frac{1}{\alpha}} + \frac{1}{M_l} \sum_{k=1}^{M_l} \left[c \left(1 + \|u_0^{h,k,\epsilon}(\omega)\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'} \right) \right]^{\frac{1}{\alpha}} \\
&\leq \liminf_{l \rightarrow \infty} |Q_T|^{\frac{1}{\alpha}} + \frac{1}{M_l} \sum_{k=1}^{M_l} \left[c \left(1 + 2\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + 2\epsilon^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'} \right) \right]^{\frac{1}{\alpha}} \\
&= |Q_T|^{\frac{1}{\alpha}} + c \left(3 + 2\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \|F\|_{L^{p'}(Q_T; \mathbb{R}^m)}^{p'} \right)^{\frac{1}{\alpha}} \leq C, \quad \text{for } P\text{-a.e. } \omega \in \mathfrak{X},
\end{aligned}$$

where C is a positive constant, independent of $h \in (0, h_0]$ and $\epsilon \in (0, 1]$. Here, in the transition from the left-hand side of the second inequality to its right-hand side we have used the triangle inequality; in the transition from the left-hand side of the third inequality to its right-hand side we have again used the triangle inequality; in the transition from the left-hand side of the fourth inequality to its right-hand side we have used the energy estimate (39), followed by recalling that $\|u_0^{h,k,\epsilon}(\omega)\|_{L^2(\Omega; \mathbb{R}^m)} \leq \|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)} + \epsilon$, with $\epsilon \in (0, 1]$, and noting that by the assumed strong convergence of u_0^h to u_0 in $L^2(\Omega; \mathbb{R}^m)$ one can bound $\|u_0^h\|_{L^2(\Omega; \mathbb{R}^m)}$ by a constant, independent of $h \in (0, h_0]$.

Having shown that $\|\langle \nu^{h,\epsilon}, \tilde{a} \rangle\|_{L^\alpha(Q_T)}$ is bounded by a constant, independent of h and ϵ , we now compute

$$\begin{aligned}
\|\nu^{h,\epsilon}\|_{L^\alpha(Q_T; E'_{p/\alpha})} &= \sup_{\|\gamma\|_{L^{\alpha'}(Q_T; E_{p/\alpha})} = 1} \int_{Q_T} \langle \nu_{t,x}^{h,\epsilon}, \gamma(t, x, \cdot) \rangle dx dt \\
&= \sup_{\|\gamma\|_{L^{\alpha'}(Q_T; E_{p/\alpha})} = 1} \int_{Q_T} \left\langle \nu_{t,x}^{h,\epsilon}, \frac{\gamma(t, x, \cdot)(1 + |\cdot|^{\frac{p}{\alpha}})}{1 + |\cdot|^{\frac{p}{\alpha}}} \right\rangle dx dt \\
&= \sup_{\|\gamma\|_{L^{\alpha'}(Q_T; E_{p/\alpha})} = 1} \int_{Q_T} \int_{A \in \mathbb{R}^{m \times n}} \frac{\gamma(t, x, A)(1 + |A|^{\frac{p}{\alpha}})}{1 + |A|^{\frac{p}{\alpha}}} d\nu_{t,x}^{h,\epsilon}(A) dx dt \\
&\leq \sup_{\|\gamma\|_{L^{\alpha'}(Q_T; E_{p/\alpha})} = 1} \int_{Q_T} \|\gamma(t, x, \cdot)\|_{E_{p/\alpha}} |\langle \nu_{t,x}^{h,\epsilon}, \tilde{a} \rangle| dx dt \\
&\leq \|\langle \nu^{h,\epsilon}, \tilde{a} \rangle\|_{L^\alpha(Q_T)},
\end{aligned}$$

which we can further bound from above by a positive constant C , independent of $h \in (0, h_0]$ and $\epsilon \in (0, 1]$, using (51). Hence, for each $\epsilon \in (0, 1]$ there is a subsequence of the sequence of measures $\{\nu^{h_j, \epsilon}\}$ indexed by h_j and a measure $\nu^\epsilon \in L^\alpha(Q_T; E'_{p/\alpha})$ such that for every $\gamma \in E_{p/\alpha}$ we have that

$$\langle \nu^{h_j, \epsilon}, \gamma \rangle \rightarrow \langle \nu^\epsilon, \gamma \rangle$$

in $L^\alpha(Q_T; \mathbb{R}^{m \times n})$ as $j \rightarrow \infty$. In particular this holds for $\gamma = a \in E_{p/\alpha}$, $1 < \alpha < \frac{p}{q-1}$, that is:

$$\langle \nu^{h_j, \epsilon}, a \rangle \rightarrow \langle \nu^\epsilon, a \rangle$$

in $L^\alpha(Q_T; \mathbb{R}^{m \times n})$ for any such $\alpha > 1$. Similar calculations hold for the family of measures $\mu^{h,\epsilon}$.

Taking the final limit in ϵ is now straightforward, and is done similarly to the above. In particular we note that we can in fact take a diagonal subsequence of (h, ϵ) and pass to

the limit along this subsequence, obtaining the existence of a two measures μ and ν satisfying, for all $\psi \in W_0^{1,1}(0, T; W_0^{1,\alpha'}(\Omega; \mathbb{R}^m))$ (contained in $W_0^{1,1}(0, T; W_0^{1,\min\{\hat{q}, \alpha'\}}(\Omega; \mathbb{R}^m)) \subset L^{\min\{\hat{q}, \alpha'\}}(0, T; W_0^{1,\min\{\hat{q}, \alpha'\}}(\Omega; \mathbb{R}^m))$):

$$(52) \quad \int_0^T \int_{\Omega} \langle \mu_{t,x}, \xi \rangle \cdot (\partial_t \psi(t, x) - B^T \psi(t, x)) - \langle \nu_{t,x}, a \rangle : D\psi(t, x) + F(t, x) \cdot \psi(t, x) \, dx \, dt = 0.$$

Step 3: Interpretation as a Young measure solution. Next, we discuss how the pair of measures (μ, ν) satisfying (52) can be interpreted as a Young measure solution. Recall that $\mu^{h,\epsilon}$ is the law of the random variable $u^h = u^{h,\epsilon}(\omega; t, x)$, and that, for each $\omega \in \mathfrak{X}$ and each $t \in [0, T]$, the function $u^{h,\epsilon}(\omega; t, \cdot) \in V_m^h \subset W_0^{1,\infty}(\Omega; \mathbb{R}^m)$ is a linear combination of finite element basis functions over the domain Ω . Therefore,

$$x \in \Omega \mapsto \langle \mu_{t,x}^{h,\epsilon}, \xi \rangle = \int_{\mathfrak{X}} u^{h,\epsilon}(\omega; t, x) \, dP(\omega), \quad t \in [0, T],$$

is weakly differentiable with respect to $x \in \Omega$, and thanks to the definition of weak derivative and Fubini's theorem,

$$D\langle \mu_{t,x}^{h,\epsilon}, \xi \rangle = \int_{\mathfrak{X}} Du^{h,\epsilon}(\omega; t, x) \, dP(\omega) = \langle \nu_{t,x}^{h,\epsilon}, \xi \rangle, \quad (t, x) \in (0, T] \times \Omega.$$

Using the weak-star convergence of $\mu^{h,\epsilon}$ and $\nu^{h,\epsilon}$ to, respectively, μ and ν established in Step 2, in conjunction with the regularity of $u^{h,\epsilon}$ coming from (44)–(46), and the fact that \mathfrak{X} is a space of finite P -measure, we see that by passing $h, \epsilon \rightarrow 0_+$ along subsequences, we get the existence of a function

$$u \in L^\infty \left(\mathfrak{X}; \bigotimes_{i=1}^m L^{p_i}(0, T; W_0^{1,p_i}(\Omega; \mathbb{R}^m)) \right)$$

such that, for $(t, x) \in (0, T] \times \Omega$,

$$\langle \mu_{t,x}, \xi \rangle = \int_{\mathfrak{X}} u(\omega; t, x) \, dP(\omega) \quad \text{and} \quad \langle \nu_{t,x}, \xi \rangle = \int_{\mathfrak{X}} Du(\omega; t, x) \, dP(\omega).$$

This comes from the fact that, for example,

$$\langle \mu^{h,\epsilon}, \xi \rangle \rightharpoonup \langle \mu, \xi \rangle,$$

and we have that

$$\int_0^T \int_{\Omega} \int_{\mathfrak{X}} u^{h,\epsilon}(\omega; t, x) \cdot \phi(\omega; t, x) \, dP(\omega) \, dx \, dt \rightarrow \int_0^T \int_{\Omega} \int_{\mathfrak{X}} u(\omega; t, x) \cdot \phi(\omega; t, x) \, dP(\omega) \, dx \, dt$$

for all $\phi \in L^1(\mathfrak{X}; L^\alpha(0, T; W_0^{1,\alpha}(\Omega; \mathbb{R}^m)))$. However, as \mathfrak{X} has finite P -measure, we can choose test functions which are constant in the parameter ω to see that

$$\int_0^T \int_{\Omega} \left(\int_{\mathfrak{X}} u^{h,\epsilon}(\omega; t, x) \, dP(\omega) \right) \cdot \phi(t, x) \, dx \, dt \rightarrow \int_0^T \int_{\Omega} \left(\int_{\mathfrak{X}} u(\omega; t, x) \, dP(\omega) \right) \cdot \phi(t, x) \, dx \, dt$$

for all $\phi \in L^\alpha(0, T; W_0^{1,\alpha}(\Omega; \mathbb{R}^m))$, which gives the desired representation for $\langle \mu_{t,x}, \xi \rangle$ (the corresponding term for $\langle \nu_{t,x}, \xi \rangle$ follows similarly).

Now we define the function

$$U(t, x) := \int_{\mathfrak{X}} u(\omega; t, x) \, dP(\omega) = \langle \mu_{t,x}, \xi \rangle.$$

The function U is locally integrable over Q_T and it therefore has a well-defined distributional derivative DU . If we can demonstrate that $DU(t, x) = \langle \nu_{t,x}, \xi \rangle$ in a suitable sense, then we will have shown that the pairing (U, ν) is a Young measure solution of the system (1)–(9) under consideration. What must first be shown is that it makes sense to speak of DU as a function, rather than as a distribution. To that end, note that the function $u \in L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m))$ satisfies $u|_{\partial\Omega} = 0$, and so we can extend the function u (and hence the function U) by zero from $[0, T] \times \bar{\Omega}$ to the whole of $[0, T] \times \mathbb{R}^n$ so that the extended function (still denoted by u) belongs to $L^p(0, T; W^{1,p}(\mathbb{R}^n; \mathbb{R}^m))$. We consider the difference quotient of U , defined by

$$D_i^\delta U = \frac{U(t, x + \delta e_i) - U(t, x)}{\delta},$$

where e_i is the unit vector in the i -th co-ordinate direction. We then have that

$$\begin{aligned} \|D_i^\delta U(t, \cdot)\|_{L^{p_i}(\Omega; \mathbb{R}^m)} &= \left\| \frac{U(t, \cdot + \delta e_i) - U(t, \cdot)}{\delta} \right\|_{L^{p_i}(\Omega; \mathbb{R}^m)} \\ &= \left\| \int_{\mathfrak{X}} \frac{u(\omega; t, \cdot + \delta e_i) - u(\omega; t, \cdot)}{\delta} \, dP(\omega) \right\|_{L^{p_i}(\Omega; \mathbb{R}^m)} \\ &\leq \int_{\mathfrak{X}} \left\| \frac{u(\omega; t, \cdot + \delta e_i) - u(\omega; t, \cdot)}{\delta} \right\|_{L^{p_i}(\Omega; \mathbb{R}^m)} \, dP(\omega) \\ &\leq \int_{\mathfrak{X}} \|D_i u(\omega; t, \cdot)\|_{L^{p_i}(\Omega; \mathbb{R}^m)} \, dP(\omega), \quad \forall \delta \in (0, 1), \quad i = 1, \dots, n. \end{aligned}$$

where in the transition to the last line we have used that, by Jensen's inequality,

$$\begin{aligned} \left\| \frac{u(\omega; t, \cdot + \delta e_i) - u(\omega; t, \cdot)}{\delta} \right\|_{L^{p_i}(\Omega; \mathbb{R}^m)}^{p_i} &= \left\| \int_0^1 D_i u(\omega; t, \cdot + s\delta) \, ds \right\|_{L^{p_i}(\Omega; \mathbb{R}^m)}^{p_i} \\ &\leq \int_0^1 \|D_i u(\omega; t, \cdot + s\delta)\|_{L^{p_i}(\Omega; \mathbb{R}^m)}^{p_i} \, ds \\ &= \int_0^1 \int_{\Omega} |D_i u(\omega; t, x + s\delta)|^{p_i} \, dx \, ds \leq \int_0^1 \int_{\Omega} |D_i u(\omega; t, x)|^{p_i} \, dx \, ds = \|D_i u(\omega; t, \cdot)\|_{L^{p_i}(\Omega; \mathbb{R}^m)}^{p_i}. \end{aligned}$$

Therefore, by a standard characterization of Sobolev functions in terms of difference quotients, $D_i U(t, \cdot) \in L^{p_i}(\Omega; \mathbb{R}^{m \times n})$, $i = 1, \dots, n$, for almost every $t \in [0, T]$. By integrating over $t \in [0, T]$ in the above inequality it follows that $D_i U \in L^{p_i}(Q_T; \mathbb{R}^{m \times n})$, $i = 1, \dots, n$, and in particular $U \in L^p(0, T; W_0^{1,p}(\Omega; \mathbb{R}^m))$, as required. By an analogous argument,

$$U \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)), \quad \partial_t U \in L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)),$$

thanks to the regularity results

$$u(\omega; \cdot, \cdot) \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)), \quad \partial_t u(\omega; \cdot, \cdot) \in L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)), \quad \text{for a.e. } \omega \in \mathfrak{X},$$

which follow from the regularity properties of $u^{h,\epsilon}$ stated in (44)–(46) using the weak lower-semi-continuity of the norm function.

Next, we write

$$V(t, x) := \langle \nu_{t,x}, \xi \rangle = \int_{\mathfrak{X}} Du(\omega; t, x) dP(\omega),$$

and let $\varphi \in C_0^\infty(Q_T; \mathbb{R}^{m \times n})$ be a test function. We then compute, using Fubini's Theorem, with $D = D_x$ and $\operatorname{div} = \operatorname{div}_x$,

$$\begin{aligned} \int_0^T \int_{\Omega} DU(t, x) : \varphi(t, x) dx dt &= - \int_0^T \int_{\Omega} U(t, x) \cdot \operatorname{div} \varphi(t, x) dx dt \\ &= - \int_0^T \int_{\Omega} \int_{\mathfrak{X}} u(\omega; t, x) \cdot \operatorname{div} \varphi(t, x) dP(\omega) dx dt \\ &= - \int_{\mathfrak{X}} \int_0^T \int_{\Omega} u(\omega; t, x) \cdot \operatorname{div} \varphi(t, x) dx dt dP(\omega) \\ &= \int_{\mathfrak{X}} \int_0^T \int_{\Omega} Du(\omega; t, x) : \varphi(t, x) dx dt dP(\omega) \\ &= \int_0^T \int_{\Omega} \int_{\mathfrak{X}} Du(\omega; t, x) : \varphi(t, x) dP(\omega) dx dt \\ &= \int_0^T \int_{\Omega} V(t, x) : \varphi(t, x) dx dt \quad \forall \varphi \in C_0^\infty(Q_T; \mathbb{R}^{m \times n}). \end{aligned}$$

Hence, $DU(t, x) = V$, and because $V := \langle \nu, \xi \rangle$, it follows that $DU = \langle \nu, \xi \rangle$, as desired. Thus we have shown that the pair (U, ν) is a Young measure solution of the system (1)–(9) under consideration.

Step 4: Attainment of the initial condition. As

$$\partial_t U^{h,\epsilon} \in L^{\hat{q}'}(0, T; W^{-1, \hat{q}'}(\Omega; \mathbb{R}^m)),$$

and

$$U^{h,\epsilon} \in \bigtimes_{i=1}^m L^{p_i}(0, T; W_0^{1, p_i}(\Omega; \mathbb{R}^m)) \cap L^\infty(0, T; L^2(\Omega; \mathbb{R}^m)),$$

it follows from Lemma 3 that there is a subsequence (h_j, ϵ_j) of (h, ϵ) along which

$$(53) \quad \int_{\Omega} U^{h_j, \epsilon_j}(t, x) \cdot \varphi(x) dx \rightarrow \int_{\Omega} U(t, x) \cdot \varphi(x) dx$$

uniformly in $C([0, T])$ for all $\varphi \in L^2(\Omega; \mathbb{R}^m)$. Therefore to show that $U(0, \cdot) = u_0$ we let $\varphi \in L^2(\Omega; \mathbb{R}^m)$ and compute:

$$\begin{aligned} (54) \quad \int_{\Omega} (U(0, x) - u_0(x)) \cdot \varphi(x) dx &= \int_{\Omega} (U(0, x) - U^{h_j, \epsilon_j}(0, x)) \cdot \varphi(x) dx \\ &\quad + \int_{\Omega} (U^{h_j, \epsilon_j}(0, x) - U^{h_j, \epsilon_j}(t, x)) \cdot \varphi(x) dx \\ &\quad + \int_{\Omega} \left(U^{h_j, \epsilon_j}(t, x) - \int_{\mathfrak{X}} u_0^{h_j, \epsilon_j}(\omega, x) dP(\omega) \right) \cdot \varphi(x) dx \\ &\quad + \int_{\Omega} \left(\int_{\mathfrak{X}} u_0^{h_j, \epsilon_j}(\omega, x) dP(\omega) - u_0(x) \right) \cdot \varphi(x) dx \\ &=: I_j + II_j(t) + III_j(t) + IV_j. \end{aligned}$$

The term I_j converges to zero as $j \rightarrow \infty$ by the uniform convergence in $C([0, T])$ stated in (53). The term $II_j(t)$ converges to zero as $t \rightarrow 0_+$ as it holds that $u^{h_j, \epsilon_j}(\omega; t, x) \rightarrow u^{h_j, \epsilon_j}(\omega, 0, x)$ by continuity of $t \in [0, T] \mapsto u^{h_j, \epsilon_j}(\omega; t, x)$, $\omega \in \mathfrak{X}$, $x \in \Omega$, and we can apply the Dominated Convergence Theorem to interchange this limit with the integral over \mathfrak{X} that appears in the definition of U^{h_j, ϵ_j} . The term $III_j(t)$ also converges to zero as $t \rightarrow 0_+$ as we have that the numerical functions u^{h_j, ϵ_j} satisfy the initial condition, and we may once again apply the Dominated Convergence Theorem to interchange the limit with the integral over \mathfrak{X} . Finally for the term IV_j we compute:

$$\begin{aligned} \left| \int_{\mathfrak{X}} u_0^{h_j, \epsilon_j}(\omega, x) dP(\omega) - u_0(x) \right| &= \left| \int_{\mathfrak{X}} u_0^{h_j} + \epsilon_j v^{h_j}(\omega; x) - u_0(x) dP(\omega) \right| \\ &\leq |u_0^h(x) - u_0(x)| + \epsilon_j \int_{\mathfrak{X}} |v^{h_j}(\omega, x)| dP(\omega). \end{aligned}$$

Integrating this over Ω we get

$$\begin{aligned} &\int_{\Omega} \left| \int_{\mathfrak{X}} u_0^{h_j, \epsilon_j}(\omega, x) dP(\omega) - u_0(x) \right| dx \\ &\leq \|u_0^{h_j} - u_0\|_{L^1(\Omega; \mathbb{R}^m)} + \epsilon_j \int_{\mathfrak{X}} \|v^{h_j}(\omega, \cdot)\|_{L^1(\Omega; \mathbb{R}^m)} dP(\omega) \\ &\leq \|u_0^{h_j} - u_0\|_{L^2(\Omega; \mathbb{R}^m)} + \epsilon_j |\Omega|^{\frac{1}{2}} \int_{\mathfrak{X}} \|v^{h_j}(\omega, \cdot)\|_{L^2(\Omega; \mathbb{R}^m)} dP(\omega) \\ &\leq \|u_0^{h_j} - u_0\|_{L^2(\Omega; \mathbb{R}^m)} + \epsilon_j |\Omega|^{\frac{1}{2}}, \end{aligned}$$

which converges to zero as $j \rightarrow \infty$ by the assumed strong convergence of our discretized initial condition u_0^h to u_0 in $L^2(\Omega; \mathbb{R}^m)$. In summary then, $\lim_{j \rightarrow \infty} I_j = 0$, $\lim_{t \rightarrow 0_+} II_j(t) = 0$ for all $j \geq 1$, $\lim_{t \rightarrow 0_+} III_j(t) = 0$ for all $j \geq 1$, and $\lim_{j \rightarrow \infty} IV_j = 0$. Passing to the limit $t \rightarrow 0_+$ in (54) with $j \geq 1$ kept fixed and then with $j \rightarrow \infty$ we deduce that

$$\int_{\Omega} (U(0, x) - u_0(x)) \cdot \varphi(x) dx = 0 \quad \forall \varphi \in L^2(\Omega; \mathbb{R}^m).$$

Thus we have that $U(0, x) = u_0(x)$ for almost every $x \in \Omega$, showing that the initial condition is satisfied.

In summary, we have that that the pair (U, ν) is a Young measure solution to the model problem satisfying the initial condition $U(0, \cdot) = u_0(\cdot)$. For clarity we state the following theorem which collects the most relevant properties proved in Steps 1–4 above.

Theorem 3. *Algorithms C and D converge to a Young measure solution of the problem (1)–(9) in the event that the initial Young measure is a Dirac mass, $\delta_{u_0(x)}$, concentrated on an initial datum $u_0 \in L^2(\Omega; \mathbb{R}^m)$.*

Remark 7. *If we have a (nonatomic) measure-valued initial datum σ , Steps 1–4 above provide an outline of an algorithm for proving the existence of a Young measure solution, assuming that all of the M “perturbations” to σ can be chosen to be bounded, independent of M . This requires being able to find a random variable u_0 whose law is given by the initial Young measure σ . The existence of such a random variable is guaranteed by Proposition 3.*

Remark 8. *In [10] and [11], in their analysis of semidiscrete schemes the authors are able to interchange the order in which the limits of the spatial discretization parameter and*

the parameter ϵ tending to zero and $M \rightarrow \infty$ are taken. For us here the passages to the limits in $h \rightarrow 0_+$, $\epsilon \rightarrow 0_+$ and $\Delta t \rightarrow 0_+$ are all interchangeable, but the passage in M is problematic from this point of view. The reason for this is that, while the limits obtained in the proof of Theorem 15 are independent of h , computing these limits in the first place required us to use the independence and the “identicalness” of the distributions of the functions $u^{h,k,\epsilon}$. This required that the solution operator \mathbf{S}_t was measurable. However, for the model problem under consideration here, we were only able to show continuity of this operator on the functions $u^{h,k,\epsilon}$ for each fixed h , and were unable to transfer this property to the limiting function as we pass $h \rightarrow 0_+$, meaning that within the context of the algorithms, we must take a limit passage in M before the limit passage in h . If instead one could show that the operator \mathbf{S}_t was measurable when considered as a mapping from the nondiscretized initial condition $u_0 \in L^2(\Omega; \mathbb{R}^m)$, we would then be able to interchange the limits in any way we desired to, as this property would be sufficient for the independence and “identicalness” of the distributions to carry over to the limiting functions as we take $h \rightarrow 0_+$.

Acknowledgement. We are grateful to Professor Mike Cullen (UK Met Office, Exeter) for stimulating discussions and for suggesting to us the problem studied in this paper. This work was supported by the UK Engineering and Physical Sciences Research Council [EP/L015811/1].

REFERENCES

1. J. M. Ball. A version of the fundamental theorem for Young measures. In *PDEs and Continuum Models of Phase Transitions*, volume 344 of *Lecture Notes of Physics*, pages 207–215. Springer-Verlag, 1989.
2. R. J. Beare and M. J. P. Cullen. Validating weather and climate models at small Rossby numbers: including a boundary layer. *Quarterly Journal of the Royal Meteorological Society*, 142(700):2636–2645, 2016.
3. G. Belletini, G. Fusco, and N. Guglielmi. A concept of solution and numerical experiments for forward-backward diffusion equations. *Discrete and Continuous Dynamical Systems*, 16(4):783 – 842, 2006.
4. J. Bergh and J. Löfström. *Interpolation Spaces: An Introduction*, volume 223 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1976.
5. C. Carstensen and T. Roubíček. Numerical approximation of Young measures in non-convex variational problems. *Numer. Math.*, 84:395–415, 2000.
6. M. Crouzeix and V. Thomée. The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces. *Math. Comp.*, 48(178):521–532, 1987.
7. S. Demoulini. Young measure solutions for a nonlinear parabolic equation of forward-backward type. *SIAM J. Math. Anal.*, 27:376–403, 1996.
8. J. Douglas, Jr., T. Dupont, and L. Wahlbin. The stability in L^q of the L^2 -projection into finite element function spaces. *Numer. Math.*, 23:193–197, 1974/75.
9. M. Dreher and A. Jüngel. Compact families of piecewise constant functions in $L^p(0, T; B)$. *Nonlinear Anal.*, 75(6):3072–3077, 2012.
10. U. S. Fjordholm, R. Käppeli, S. Mishra, and E. Tadmor. Construction of approximate entropy measure valued solutions for hyperbolic systems of conservation laws. *Foundations of Computational Mathematics*, 17(3):763827, 2017.

11. U. S. Fjordholm, S. Mishra, and E. Tadmor. On the computation of measure-valued solutions. *Acta Numerica*, 25:567–679, 2016.
12. J. Frehse and M. Specovius-Neugebauer. Hölder continuous Young measure solutions to coercive non-monotone parabolic systems in two space dimensions. *Applicable Analysis*, 90(1):67–84, 2011.
13. V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations: Theory and algorithms*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986.
14. J. Haškovec and Ch. Schmeiser. Stochastic particle approximation for measure valued solutions of the 2D Keller-Segel system. *J. Stat. Phys.*, 135(1):133–151, 2009.
15. S. Kim and B. Yan. Convex Integration and Infinitely Many Weak Solutions to the Perona–Malik Equation in all dimensions. *SIAM Journal on Mathematical Analysis*, 47(4):2770–2794, 2015.
16. S. Kim and B. Yan. On Lipschitz solutions for some forwardbackward parabolic equations. *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, 35(1):65 – 100, 2018.
17. D. Kinderlehrer and P. Pedregal. Gradient Young measures generated by sequences in Sobolev spaces. *The Journal of Geometric Analysis*, 4(1), 1994.
18. J.-L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications. Vol. I*. Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
19. W. A. Strauss. On continuity of functions with values in various Banach spaces. *Pacific J. Math.*, 19:543–551, 1966.
20. B. L. T. Thanh. *Forward-Backward Parabolic Equations*. PhD thesis, Sapienza University of Rome, 2013.
21. B. L. T. Thanh, F. Smarrazzo, and A. Tesei. Sobolev regularization of a class of forwardbackward parabolic equations. *Journal of Differential Equations*, 257(5):1403 – 1456, 2014.