QUARTERLY OF APPLIED MATHEMATICS
VOLUME , NUMBER 0
XXXX XXXX, PAGES 000{000

DISCRETIZING ADVECTION EQUATIONS WITH ROUGH VELOCITY FIELDS ON NON-CARTESIAN GRIDS

By

PIERRE-EMMANUEL JABIN (P.{E. Jabin. Department of Mathematics and Huck Institutes, Pennsylvania State University, State College, PA 16801, USA)

and

DATONG ZHOU (D. Zhou. Department of Mathematics, Pennsylvania State University, State College, PA 16801, USA)

Abstract. We investigate the properties of discretizations of advection equations on non-cartesian grids and graphs in general. Advection equations discretized on non-cartesian grids have remained a long-standing challenge as the structure of the grid can lead to strong oscillations in the solution, even for otherwise constant velocity elds. We introduce a new method to track oscillations of the solution for rough velocity elds on any graph. The method in particular highlights some inherent structural conditions on the mesh for propagating regularity on solutions.

1. Introduction.

1.1. Discretized advection equations. We introduce a new framework to study the regularity of discretized advection equations. Our method is able to provide quantitative regularity estimates by extending the kernel based approach initially introduced at the continuum level in [5, 4] and further studied in [35, 31]. This is particularly helpful when investigating the convergence of numerical schemes for coupled non-linear systems.

To be more specic, we study discretizations of the classical linear continuity equation,

$$\partial_t u(t,x) + \operatorname{div}_x\, b(t,x)u(t,x) = 0, \quad t \in \mathbb{R}_+,\ x \in \mathbb{R}^d, \tag{1.1}$$

Those discretized equations usually involve calculating the dynamics of a discrete density u_i that is dened on each cell of a grid or mesh. We specically focus on upwind schemes that read

$$\frac{du_i}{dt} = \frac{1}{{}_i}\sum_{i'} \quad a_{i,i'}\,u_{i'} \quad a_{i',i}\,u_i \ . \tag{1.2}$$

~~2020 Mathematic~~s Subject Classication. Primary .
P. E. Jabin and D. Zhou were partially supported by NSF DMS Grant 2049020, 2219397, and 2205694.
E-mail address: pejabin@psu.edu
E-mail address: dbz5086@psu.edu

The factor ${}_i$ represents some notion of volume of the cell i on the mesh and the coefficients $a_{i,i'}, a_{i',i}$ are related to the flux between two cells i and i'. We refer to Section 1.4 for the precise formulas for the schemes that we consider.

The continuity equation constitutes a key relation in a large variety of models, in which the velocity field $b(t,x)$ is typically related to the density $u(t,x)$ in different ways. As mentioned above, it is this coupling between b and u that makes strong compactness of the density, instead of weak convergence, an essential ingredient and one of the common difficulties when trying to prove convergence for the whole system, whether from numerical approximations or some other approximate system. We present a few typical examples below that motivate our investigations and have natural applications in biology and fluid mechanics.

The velocity field b in (1.1) can first be related to the density u by some convolution $b = K \star g(u)$ for some non-linear function g or by the Poisson equation

$$b(t,x) = -\nabla_x \phi(t,x), \qquad -\Delta_x \phi(t,x) = g(u(t,x)), \tag{1.3}$$

which corresponds to choosing the fundamental solution of the Laplacian as the kernel K. There exist already many examples of such systems in applications: We briefly mention [43] for swarming or [16, 38] for models of chemotaxis. The function g represents a non-linear dependence on the density u in the equation for b, which can capture more complex phenomena in the model such as logistic effects.

In a somewhat similar spirit, non-linear continuity equations may be considered such as

$$\partial_t u(t,x) + \mathrm{div}_x\, \big(b(t,x) f(u(t,x))\big) = 0, \qquad t \in \mathbb{R}_+,\ x \in \mathbb{R}^d. \tag{1.4}$$

This type of non-linear flux combines non-linear scalar conservation laws with linear advection. Models such as (1.4) are found for example in some biological settings, where the speed of micro-organisms is impacted by their local density; see, for instance, the discussion of Keller-Segel model in [38]. We expect that the results in this paper can be extended to such non-linear models, but for the sake of simplicity in this article, we consider only the linear continuity equation and exclude such nonlinearity in the flux from the rest of our discussion.

The continuity equation (1.1) is also naturally a critical component of compressible fluid dynamics such as the barotropic compressible Navier-Stokes system

$$\begin{cases} \dfrac{\partial u}{\partial t} + \mathrm{div}(bu) = 0 \\[2mm] \dfrac{\partial}{\partial t}(bu) + \mathrm{div}\,(b \otimes b)u + \nabla p - \mathrm{div}(\sigma(b)) = f \quad p = P(u), \end{cases}$$

with appropriate boundary conditions if considered in a bounded domain. In this system, the velocity field b is coupled with the density u by another evolution PDE, leading to an even more complex non-linearity than in the previous examples.

A related model is the Stokes system,

$$\begin{cases} \dfrac{\partial u}{\partial t} + \operatorname{div}(b u) = 0 \\ -\Delta b + \nabla p = f \\ p = P(u); \end{cases}$$

which considerably simplifies the momentum equation and the relation between b and u.

There exists a large literature on the numerical analysis of compressible Navier-Stokes system to which we cannot do justice in a few sentences. We only briefly mention [28, 27] for the compressible Navier-Stokes and [24, 25, 29] for the Stokes system. To the best of our knowledge however, the numerical analysis of these systems is only well-understood on Cartesian meshes or staggered grids, for example in [30], that still rely on Cartesian mesh for the density. Generally speaking, the regularity of discretized continuity equations such as (1.2) remains poorly understood on non-cartesian meshes, which leads to the main motivation and focus in the present work.

1.2. Renormalized solutions. Even without discretization, the well-posedness for advection equations such as (1.1) is in itself a delicate question when the velocity field b is not Lipschitz. By introducing the concept of renormalized solution, the uniqueness and compactness of (1.1) was first obtained in [23] for velocity fields $b \in W^{1,p}$. This was later improved in [2, 7] to only $b \in BV$ with $\operatorname{div} b \in L^1$.

Renormalized solutions are based on a simple but essential observation: Assume that b and u are smooth and satisfy the continuity equation (1.1). Then for all $\beta \in C^1(\mathbb{R})$, $\beta(u)$ is a solution of

$$\partial_t \beta(u) + \operatorname{div}(b\,\beta(u)) + \operatorname{div} b\,(\beta'(u)u - \beta(u)) = 0. \tag{1.5}$$

A weak solution of (1.1) with a non-smooth field b is said to be renormalized iff (1.5) holds in distributional sense for all $\beta \in C^1(\mathbb{R})$ with $|\beta(\xi)| \le C|\xi|$. Moreover, equation (1.1) with a fixed field b is said to be renormalized iff all its weak solutions are renormalized. Basically, the renormalization property consists in stating that if u is a weak solution then non-linear functions of u are also solutions, with appropriate corrections if $\operatorname{div} b \neq 0$. This directly implies the uniqueness of a weak solution u: Consider two weak solutions u, v with $u(0,x) = v(0,x)$, if $u - v$ is a renormalized solution, then $|u - v|$ is also a weak solution. Hence $\|u(t,.) - v(t,.)\|_{L^1(\mathbb{R}^d)} \le \|u(0,.) - v(0,.)\|_{L^1(\mathbb{R}^d)} = 0$ and $u = v$. Combining the uniqueness and the renormalization property directly provides compactness in the appropriate L^p_{loc} sense as one can prove that

$$\text{weak-* } \lim \beta(u_n) = \beta(\text{weak-* } \lim u_n).$$

A common critical part in the proof of the renormalization property is a so-called commutator estimate. Consider a classical convolution kernel K_ε together with $K_\varepsilon \star_x u$ where u solves (1.1). Commutator estimates arise when trying to write a similar equation on $K_\varepsilon \star_x u$. One then has that

$$\partial_t (K_\varepsilon \star_x u) + \operatorname{div}\left(b\, K_\varepsilon \star_x u\right) = R_\varepsilon;$$

where the remainder term R_ε can be written as a commutator

$$R_\varepsilon(x) = \int \big(b(x) - b(y)\big)\cdot\nabla K_\varepsilon(x-y)\,u(y)\,dy + \int \mathrm{div}_x\, b(x)\,K_\varepsilon(x-y)\,u(y)\,dy.$$

Commutator estimates then consists in proving that $R_\varepsilon \to 0$. If it is possible to prove this, then it is straightforward to deduce the renormalization property (1.5) by writing an equation on $\chi(K_\varepsilon \star_x u)$ and passing to the limit $\varepsilon \to 0$.

Renormalized solutions are also connected to some form of propagation of quantitative regularity. It had already been noticed in [3] that renormalized solutions lead to some approximate differentiability on the solution. But the first explicit propagation of regularity was obtained in [20] at the level of the characteristics flow. The characteristics method used in [20] proved very fruitful with many later extensions. One can mention the study of SDEs in [18, 26, 44], the question of mixing under incompressible flows in [11, 32, 40], well-posedness for velocity fields with less than 1 derivative in [17, 33], and velocity fields obtained through a singular integral in [6, 8, 19].

The corresponding regularity at the PDE level can be derived by directly quantifying oscillations on the solution. A first method to do so was introduced in [5, 4] for non-linear continuity equations of the form (1.4). In the linear case, sharper estimates were obtained in [35] and [15] through a somewhat similar approach. We also mention [13, 14] which combines those methods with a new notion of weights; this was applied to the com-pressible Navier-Stokes equation with a large variety of laws of state and stress tensors. A very different quantitative approach at the PDE level was studied in [41, 42], using certain optimal transport distances. All those results only propagate a weaker notion of regularity, weaker than full differentiability, usually some sort of log of derivative. It is indeed not possible in general to bound any kind of Sobolev regularity on density when the velocity field is merely Sobolev; see some counterexamples for example in [1, 34].

The approach that we follow in the present paper is inspired by the quantitative semi-norms introduced in [5, 4], which we briefly describe for this reason. The local compactness of a sequence of bounded functions $u_k \in L^p(\mathbb{R}^d)$ with $1 \le p < \infty$ follows from the following property:

$$\limsup_n \int_{\mathbb{R}^{2d}} \bar K_h(x-y)\,|u_n(x) - u_n(y)|^p\,dx\,dy \to 0 \quad \text{as } h \to 0; \tag{1.6}$$

where $\{\bar K_h\}_{h>0}$ is any family of classical convolution kernels. Scaling (1.6) with a given rate of convergence in h leads to various notions of semi-norms that measure intermediate regularity between L^p and $W^{s,p}$ for any $s > 0$, and all of such regularities are strong enough to imply local compactness in L^p.

The particular family of kernels $\{\bar K_h\}$ proposed in [5, 4] results in semi-norms corresponding to a sort of log-scale derivatives that we denote here by $\dot W^p_{\log,\theta}$. The $\dot W^p_{\log,\theta}$-regularity defined by kernels $\{K_h\}$ was then proved to be propagated by (1.1) when the velocity field $b \in W^{1,p}$, $\mathrm{div}\, b$ is bounded and $\mathrm{div}\, b$ is compact or enjoys some similar $W^p_{\log,\theta}$-regularity.

Hence with such assumptions on b, the solutions of (1.1) are compact if the initial data are $W^p_{\log,\theta}$-regular. The bounds in [35] and [15] yield some more precise log-scale derivatives based on somewhat similar semi-norms. The corresponding spaces have also received increasing attention in other settings, see for instance [9].

When trying to extend the idea of quantifying oscillations in [5, 4] to our discrete setting, it appears natural to introduce an approximation of the continuous kernel K_h on the mesh. In other words, we would like to estimate the regularity of the discrete density by something like

$$\limsup_n \sum_{i,j} \tilde{K}^h_{i,j,(n)} |u_{i,(n)} - u_{j,(n)}|^p{}_{i,(n)}{}_{j,(n)} \to 0, \qquad \text{as } h \to 0, \tag{1.7}$$

where $(\tilde{K}^h_{i,j})_{i,j}$ is an approximation of kernel K_h and the double integral is replaced by a double summation over the mesh.

The main issue however is to identify the right family of kernels $(\tilde{K}^h_{i,j})_{i,j}$ so that the corresponding semi-norms are propagated by the discrete advection equation. This turns out to be extremely challenging on non-cartesian grids as a straightforward discretization of the kernels $\{K_h\}$ used for the continuous equation does not appear to work. The main technical contribution of this paper is a general method to nd admissible approximation $(\tilde{K}^h_{i,j})_{i,j}$, extending the results in [5, 4] to upwind schemes. This leads to the study of a non-symmetric diusion equation on the mesh which we can solve and bound when the mesh show periodic patterns (the exact denition is given in Section 1.4).

1.3. Some of the issues with non-cartesian grids. At rst glance, it may not be appar-ent why non-cartesian meshes lead to such additional diculty. Eq. (1.2) may in fact be seen as an advection equation on a graph where the actual velocity eld is correspond to some projection of the original velocity eld that incorporates the structure of the graph. This means that the graph's topology can lead to additional oscillations in the solution in itself. This is made apparent in the following straightforward example, that we are grateful to T. Gallouët and R. Herbin for pointing out. This shows that even for very smooth or actually constant velocity elds at the continuous level, one may have strong oscillations in the solution at the discrete level.

Example 1.1. Consider the constant velocity eld $b(x) \equiv (1,0)$ in dimension 2 and the following non-cartesian discretization: Let h_0 be the discretization parameter, and use $\mathbb{Z}^2$ to index the cells. The cell indexed by $(i,j) \in \mathbb{Z}^2$ is given by

$$C_{(j,k)} = \begin{cases} [jh, (j+1)h) \times [kh, (k+1)h) & \text{if } k \text{ is even;} \\ [jh/2, (j+1)h/2) \times [kh, (k+1)h) & \text{if } k \text{ is odd:} \end{cases}$$

That simply means that we keep the vertical discretization h, but alternate a row with horizontal discretization h, with another row with discretization $h/2$.

Consider a discrete density $(u_{j,k})_{j,k}$ solving the upwind scheme (1.2) over such a mesh for a discretization of the constant velocity eld $b = (1,0)$. Assume that the initial data $(u_{j,k}(0))_{j,k}$ is bounded in discrete $W^{1,1}$-norm, uniformly in h_0. Then for any $t > 0$, $(u_{j,k}(t))_{j,k}$ is bounded in the discrete $W^{s,1}$-norm, uniformly in h_0, if and only if $s < 1/2$.

This type of spurious oscillations created by the mesh itself are one of the reasons why the aforementioned quantitative methods (either in ODE or PDE level) have not been extended to non-Cartesian meshes. In fact, there exist only very few qualitative results of strong convergence for non Lipschitz velocity elds and non-Cartesian meshes.

One can nevertheless mention [10] which relies on the renormalization property at the limit. However because this kind of approach is not quantitative, it requires some a priori knowledge of the compactness of the divergence of the velocity field. This appears to make handling coupled non-linear models such as (1.1)-(1.3) out of reach.

When one is not trying to handle at the same time non-Cartesian meshes and non-Lipschitz coefficients, stronger results can be obtained. On non-Cartesian meshes, we refer for instance to [36, 37] for divergence-free velocity field that are Lipschitz in both space and time, and to [21] for autonomous (time-independent) Lipschitz velocity fields with non-zero divergence. For non-Lipschitz velocity fields on Cartesian meshes, one can obtain quantitative convergence results in some suitable weak distances. When the velocity field is in the appropriate Sobolev space with one-sided bounded divergence, the upwind scheme was proved to converge at rate of $1/2$ in [39] in some weak topology. When the velocity field is one-sided Lipschitz continuous, the convergence with rate $1/2$ of the upwind scheme in Wasserstein distance was proved in [22].

To the best of our knowledge however, this article is the first to provide a general approach to the compactness of solutions to discrete advection equations with non-Lipschitz coefficients and non-Cartesian meshes, even if we still require some restrictions on the mesh such as periodic patterns.

Furthermore, the compactness result in this paper is directly applicable to some of the coupled systems discussed at the beginning of this introduction. We are in particular able to derive the compactness of discretizations of the non-linear coupled system (1.1)-(1.3). The exact result is stated later in this first section. We remark here that the velocity field b obtained from u through (1.3) is naturally bounded in $W^{1,p}$ for all $1 < p \leq \infty$, if we assume $g(u) \in L^1 \cap L^\infty$. Of course we cannot know a priori the compactness of $\operatorname{div} b$ but we have the simple relation $\operatorname{div} b = g(u)$. This is where quantitative, explicit estimates prove critical as we are able to conclude through some sort of Gronwall argument.

However more complex coupled systems would present unique challenges for our approach: This is notably the case of compressible fluid dynamics. Energy estimates would provide Sobolev, H^1_x bounds on the velocity. However the divergence of velocity is generically unbounded, which would prevent us from applying our method in any straightforward manner. Instead this would likely require the introduction of weights such as was done in [13, 14] at the continuous level.

1.4. A basic example of setting for the linear continuity equation. Considering the linear continuity equation (1.1), we introduce here its basic discretization on a polygon mesh $(\mathcal{C}, \mathcal{F}) = \{V_i\}_{i\in\mathcal{V}}, \{S_{i,j}\}_{(i,j)\in\mathcal{E}}$ over a bounded domain $\mathbb{R}^d$, which we defined as the following:

- The pair of indices $(\mathcal{V}, \mathcal{E})$ form a finite graph.
- Each cell V_i is a d-dim polygon in $\mathbb{R}^d$. The intersection of two cells V_i and V_j is nonempty if and only if $(i, j) \in \mathcal{E}$. In that case $S_{i,j} = V_i \cap V_j$ is a $(d-1)$-dim polygon in $\mathbb{R}^d$.
- The domain is covered by the mesh: $\bigcup_{i\in\mathcal{V}} V_i$.

We define the discretization size of the mesh as $\delta x = \sup_{i\in\mathcal{V}} \operatorname{diam}(V_i)$, where $\operatorname{diam}(\cdot)$ represents the diameter.

As a rst example, we consider the following semi-discrete upwind scheme

$$
\begin{cases}
\dfrac{d}{dt}u_i(t) = \dfrac{1}{|V_i|}\displaystyle\sum_{i':(i;i')\in E} \left(a_{i;i'}(t)\,u_{i'}(t) - a_{i';i}(t)\,u_i(t)\right); & \text{if } i \in V \text{ and } V_i + B_x \\
u_i(t) \equiv 0; & \text{if } i \in V \text{ but } V_i + B_x \not\subset \\
a_{i;j}(t) = \dfrac{1}{|B_x|}\displaystyle\int_{B_x}\int_{S_{i;j}} \left(b(x+y;t)\cdot N_{i;j}\,dydx\right)^+; & \text{if } (i;j) \in E \text{ and } S_{i;j} + B_x \\
u_i(0) = \dfrac{1}{|V_i|}\displaystyle\int_{V_i} u_0(x)\,dx; & \text{if } i \in V \text{ and } V_i + B_x
\end{cases}
\tag{1.8}
$$

where $(u_i(t))_{i\in V}$ are the discrete density on the mesh, $N_{i;j}$ is the unit normal vector on $S_{i;j}$, satisfying $N_{i;j} = -N_{j;i}$. The functions $b(t;x)$ and $u_0(x)$ are respectively the velocity eld and initial condition in the linear continuity equation (1.1). Also, throughout the paper, for $s \in \mathbb{R}$ we use the notation $s^+ = s \vee 0 = \max\{s;0\}$ and $s^- = -(s \wedge 0) = -\min\{s;0\}$.

The total mass on the mesh is given by $\sum_{i\in V} u_i|V_i|$. It is easy to verify that the scheme conserves mass except near the boundary of , where some leaking may occur. Such leaking eect can be controlled by no ux (no outward ux) condition of the velocity eld or by a priori estimating the distribution of density.

Before we can rigorously state any compactness result, we still need to clarify our assumptions on the mesh. Throughout the paper, for $A;B \subset \mathbb{R}^d$, we use the notation

$$A + B = \{x \in \mathbb{R}^d; x = a + b; a \in A; b \in B\}:$$

Also, for $x \in \mathbb{R}^d$ we denote $A + x = A + \{x\}$, which is a translation of set A on $\mathbb{R}^d$.

We say that a mesh has a periodic pattern if the following holds:

Definition 1.2. Let $(C;F) = \left(\{V_i\}_{i\in V};\{S_{i;j}\}_{(i;j)\in E}\right)$ be a polygon mesh over $\mathbb{R}^d$ and let $V_0 \subset V$. The mesh is periodic with pattern V_0 if it satises the following properties: The set $\bigcup_{i\in V_0} V_i$ is connected and one has $\bigcup_{i\in V_0} V_i + B_x$. There exists a translation group action

$$[m](V_i) = V_i + \sum_{k=1}^{d} m_k L_k; \quad \forall m \in \mathbb{Z}^d; i \in V;$$

where $L_1;:::;L_d \in \mathbb{R}^d$ are linearly independent vectors, such that

$$\sum_{m\in\mathbb{Z}^d}\sum_{i\in V_0} 1_{[m]V_i}(x) = 1 \quad \text{for a.e. } x \in \mathbb{R}^d:$$

Moreover, there exists an injective map

$$\begin{aligned} : V &\to \mathbb{Z}^d \times V_0; \\ i &\mapsto ([n];i_0): \end{aligned}$$

If $(i) = ([n];i_0)$ and $([m]+[n];i_0) \in (V)$, dene $[m](i) = {}^{-1}([m]+[n];i_0) \in V$. Then one has

$$[m](V_i) = V_{[m](i)}; \quad \text{if } i \in V \text{ and } [m](i) \in V:$$

If the mesh is periodic with pattern V_0 we call $|V_0|$ the pattern size; of course for a given mesh, the choice of V_0 and $|V_0|$ may not be unique. If one can choose $V^0 \subset V, E^0 \subset E$ such that $(C^0, F^0) = \left(\{V_i\}_{i\in V^0}, \{S_{i,j}\}_{(i,j)\in E^0}\right)$ forms a mesh over Ω^0, and is a periodic mesh by the definition above, then we say that (C, F) is periodic over Ω^0.

We also require some additional assumptions on the meshes, though those are rather standard. Throughout the discussion, any mesh (C, F) of our interest should satisfy that for all $i \in V$ and $x \in \mathbb{R}^d$:

$$C^{-1}\delta x \le \operatorname{diam}(V_i) \le C\delta x, \qquad C^{-1}(\delta x)^d \le |V_i| \le C(\delta x)^d, \quad \#\{k \in V : B(x,\delta x) \cap V_k \neq \emptyset\} \le C, \tag{1.9}$$

for some uniform constant C. These conditions exclude some pathological situations where some parts of the mesh would be too singular in some regard.

Finally, we observe that since we are considering the limit to the continuous equation, then we naturally expect the discretization size to vanish. Namely, let

$$\left(C^{(n)}, F^{(n)}\right) = \left(\{V_{i,(n)}\}_{i\in V^{(n)}}, \{S_{i,j,(n)}\}_{(i,j)\in E^{(n)}}\right), \qquad n \in \mathbb{N}_+$$

be a family of meshes and let $\delta x_{(n)}$, $n \in \mathbb{N}_+$ denote the discretization sizes, then we ask that

$$\delta x_{(n)} \to 0 \text{ as } n \to \infty. \tag{1.10}$$

We are now ready to state a first example of our compactness result:

Theorem 1.3. *Consider $T > 0$ and a bounded domain $\Omega \subset \mathbb{R}^d$ with piecewise smooth boundary. Let $b(t,x)$ be a velocity field with $b \in L^\infty_t L^\infty_x \cap L^q_t(W^{1,q}_x) \cap L^1_t(W^{s,1}_x)([0,T]\times\Omega)$ and divergence $\operatorname{div}_x b \in L^1_t(L^\infty_x) \cap L^1_t(W^{s,1}_x)([0,T]\times\Omega)$, for some $1 < q \le \infty$, $0 < s \le 1$. Let $u_0 \in L^1 \cap W^{s,1}(\Omega)$ be the initial data.*

Consider a sequence of polygonal meshes $\{(C^{(n)}, F^{(n)})\}_{n=1}^{\infty}$ over Ω, having discretiza-tion size $\delta x_{(n)} \to 0$, satisfying the structural assumptions (1.9) with some uniform con-stant, and being periodic on Ω with their pattern size also uniformly bounded. Let $(u_{i,(n)}(t))_{i\in V^{(n)}}$ be solutions to the semi-discrete scheme (1.8) and denote by $u_{(n)}(t,x)$ the piecewise constant function extending $(u_{i,(n)}(t))_{i\in V^{(n)}}$. Assume finally that the total mass $\sum_{i\in V^{(n)}} u_{i,(n)}(T)|V_{i,(n)}| \to \|u_0\|_{L^1}$ as $n \to \infty$. Then

$$u_{(n)}(t,x) \text{ is compact in } L^1([0,T]\times\Omega).$$

The proof of the theorem is postponed to Section 2.3, where it follows from the prop-agation of some discrete regularity of the form (1.7).

1.5. **The more complete setting.** We demonstrate the potential of our method by also deriving the compactness for a simple non-linear coupled system, namely

$$\begin{cases} \partial_t u(t,x) + \operatorname{div}_x\left(b(t,x)u(t,x)\right) = 0 \\ b(t,x) = \nabla_x \phi(t,x), \qquad -\Delta_x \phi(t,x) = g(u(t,x)). \end{cases} \tag{1.11}$$

However, the setting of polygon meshes described above may no longer be the most appropriate. The difficulty comes from the coupling of numerical schemes between the elliptic Poisson equation, for which one may want to use finite elements for example, and the hyperbolic advection equation for which we use upwind schemes. This is one of the motivations for our more general formulation.

We dene cell functions, face functions and meshes that replace the polygonal cells. We discuss later in subsection 1.6 how the previous polygon meshes can be related to this formulation.

Definition 1.4. Consider a piecewise dierentiable function with value 0 1 and vector values functions $f n_j g_{j=1}^m$ on R^d. Then is said to be a cell function with $f n_j g_{j=1}^m$ as its face functions if

$$P_{j=1}^m n_j(x) = \quad r(x) \quad \text{for a.e. } x\ 2\ R^d \qquad \text{and} \qquad \text{supp}\, n_j \ \ \text{supp}\,; \qquad 8j\ 2\ V:$$

With the cell functions and face functions dened, we give the following denition of meshes.

Definition 1.5. We dene as a generalized mesh over R^d a pair $(C; F) = f_i g_{i2V}; f n_{i;j} g_{(i;j)2E}$ satisfying the following conditions:

The pair of indices $(V;E)$ forms a nite graph.

If $i\ 2\ V$ and supp $_i$, then $_i$ must be a cell function with $f n_{j;i} g_{(j;i)2E}$ as its face functions.

Finally,

$$P_{i2V}\ {}_i(x) = 1; \qquad 8x\ 2 \qquad \text{and} \qquad n_{i;j} = \quad n_{j;i}; \qquad 8i; j\ 2\ V:$$

We also extend $f n_{i;j} g_{(i;j)2E}$ to $f n_{i;j} g_{i;j2V}$ by dening $n_{i;j} = n_{j;i} = 0$ for $(i;j)\ 2\ E$.

The discretization size of a mesh $(C;F)$ is dened as $x = \max_{i2V} \text{diam}(\text{supp}\,_i)$. The volume of cell $i\ 2\ V$ is dened as $_i = k_i k_{L^1}$.

The semi-discrete scheme we consider in this paper is of form

$$\begin{cases} \dfrac{d}{dt} u_i(t) = \dfrac{1}{;^i} \displaystyle\sum_{j:(i;j)2E} a_{i;j}(t)\, u_j(t) \quad a_{j;i}(t)\, u_i(t) \ ; & \text{if } i\ 2\ V \text{ and supp}\,_i \\ u_i(t)\ \ 0; & \text{if } i\ 2\ V \text{ but supp}\,_i * \end{cases} \tag{1.12}$$

Given $b(t;x)$ and $u_0(x)$ as the eld and initial condition in linear continuity equation (1.1) respectively, we choose the coecients and initial data in the scheme as

$$\begin{cases} u_i(0) = \dfrac{1}{_i} \displaystyle\int_{R^d} {}_i(x)\, u_0(x)\, dx; & \text{if } i\ 2\ V \text{ and supp}\,_i \\ a_{i;j}(t) = \displaystyle\int_{R^d} b(t;x)\ n_{i;j}(x) \ ^{+} dx; & \text{if } (i;j)\ 2\ E \text{ and supp}\, n_{i;j} \end{cases} \tag{1.13}$$

Notice that if supp $_i$, then for all $j\ 2\ V$, either one has $(i;j)\ 2\ E$, supp $n_{i;j}$ $_i$ or one has $(i;j)\ 2\ E$, $n_{i;j} = 0$ and supp $n_{i;j}$ trivially holds. Hence $a_{i;j}$ and $a_{i;j}$ in (1.12) are always well-dened. In addition, we let

$$a_{i;j}(t)\ \ 0 \quad \text{if } (i;j)\ 2\ E \text{ or supp}\, n_{i;j} * :$$

Then the summation in (1.12) can be taken over all $j\ 2\ V$, instead of only j such that $(i;j)\ 2\ E$, and the scheme is essentially unchanged. In some of the later calculations, this adaption can be convenient.

The structural assumptions to meshes we have made should also be adapted. In particular, the new denition of being periodic is the following:

Definition 1.6. Let $(C; F) = f_i g_{i2V}; fn_{i;j} g_{(i;j)2E}$ be a mesh over R^d and let V_0 V. We say that $(C; F)$ is a periodic mesh with pattern V_0 if it satises the following properties:

The set S_{i2V_0} supp $_i$ is connected and one has S_{i2V_0} supp $_i$

.

There exists a translation group action

$$8 \quad ([m]_i)\ x = {}_i\ x \quad P^{d}_{k=1} m_k L_k \ ; \quad < \quad 8[m]\ 2\ Z^d; i; j\ 2\ V;$$
$$> \ ([m] n_{i;j})\ x = n_{i;j}\ x \quad P^{d}_{k=1} m_k L_k \ ;$$

where $L_1; :::; L_n$ 2 R^d are linearly independent vectors, such that

$$X_{[m]2Z^d} X_{i2V_0} [m]_i(x) = 1 \quad 8x\ 2\ R^d:$$

Moreover, there exists an injective map

$$: V\ !\ \ Z^d\ V_0;$$
$$i\ !\ \ ([n]; i_0):$$

If $(i) = ([n]; i_0)$ and $([m] + [n]; i_0)$ 2 (V), dene $[m](i) = {}^1 ([m] + [n]; i_0)$ 2 V. Then one has

$$[m]_i = {}_{[m](i)}; \qquad [m] n_{i;j} = n_{[m](i);[m](j)}; \qquad \text{if } i; [m](i)\ 2\ V:$$

If the mesh is periodic with pattern V_0 we call jV_0j the pattern size. As before, for a given mesh, the choice of V_0 and jV_0j may not be unique. If one can choose V^0 $V; E^0$ E such that $(C^0; F^0) = f_i g_{i2V^0}; fn_{i;j} g_{(i;j)2E^0}$ forms a mesh over ^{0}, and is a periodic mesh by the denition above, then we say that $(C; F)$ is periodic over ^{0}.

The other structural assumptions on the mesh can be adapted in a straightforward manner. We limit our discussion to meshes $(C; F)$ that satisfy that for all i 2 V, $(j; j^0)$ 2 E and x 2 R^d:

$$C^{\ 1} x \ \text{diam}(\text{supp}_i)\ \ C x; \qquad C^{\ 1}(x)^d\ \ k_i k_{L^1}\ \ C(x)^d; \ x\, k n_{j;j^0} k_{L^1}\ \ C; \qquad fk\ 2\ V : (\text{supp}_k) \setminus B(x; x) = ?g\ \ C; \tag{1.14}$$

for some uniform constant C. Also, we assume that the discretization size vanishes when considering a family of meshes,

$$C^{(n)}; F^{(n)} = f_{i;(n)} g_{i2V^{(n)}}; fn_{i;j;(n)} g_{(i;j)2E^{(n)}}; \qquad n\ 2\ N_+$$

that (1.10) holds where $x_{(n)}$, n 2 N_+ denote the discretization sizes as given in Denition 1.5.

With this more general formulation, one can couple the upwind scheme for advection and the nite elements for Poisson equation in the following way: Consider convex bounded domains with piecewise smooth boundary $_v$ $_e$ R^d. Let $(P; N)$ be a nite element discretization of $_e$, where P is the set of shape functions and N is the set of nodal variables. Choose the mesh $(C; F) = f_i g_{i2V}; fn_{i;j} g_{(i;j)2E}$ over $_v$ as in Denition 1.5 such that C P.

The coupled system (1.11) is numerically discretized through by (1.12). The coe-cients $(a_{i;j})_{i;j \in V}$ derive (1.13) where the eld $b(t;x)$ is now a solution of the variational problem

$$
\begin{cases}
(t;\cdot) \in P; & \int_{\mathbb{R}^d} \nabla v(x) \cdot \nabla (t;x)\,dx = \int_{\mathbb{R}^d} v(x) g(t;x)\,dx; \quad \forall v \in P; \\
g(t;x) = \sum_{i \in V} g\,\big(u_i(t)\big)\, {}_i(x); \\
b(t;x) = \nabla (t;x):
\end{cases}
\tag{1.15}
$$

We only consider here Dirichlet boundary conditions for (1.15), as Neumann boundary conditions would require an extra condition $\int_{\mathbb{R}^d} g(t;x)\,dx = 0$ for all t, which does not naturally hold when g contains some nonlinear function of the density u.

When investigating this more complex coupling, we require further structural assumptions on the pair of nite element $(P;N)$ and the mesh $(C;F)$ of our interest. Namely, the exact solution of $= g$ and its approximated solution of the nite element variational method

$$
\in P; \quad \int \nabla v(x) \cdot \nabla (x)\,dx = \int_{\mathbb{R}^d} v(x) u(x)\,dx; \quad \forall v \in P;
$$

are assumed to satisfy the priori estimates

$$
\begin{aligned}
&\| \|_{H^1(\Omega)} \le C x \| \|_{H^2(\Omega)}; \\
&\| \|_{W^{1;1}(\Omega)} \le C \| \|_{W^{1;1}(\Omega)}:
\end{aligned}
\tag{1.16}
$$

Such a priori estimates can be proved under rather mild conditions on the nite element discretization; we refer to Section 5.4 and 8.1 of [12].

We are now ready to state our main theorem on this coupled system, whose proof is again postponed to Section 2.3.

Theorem 1.7. Consider bounded domains $\Omega_v \subset \Omega \subset \mathbb{R}^d$ with piecewise smooth bound-ary, a sequence of nite element discretizations $\{(P^{(n)}; N^{(n)})\}_{n=1}^{1}$ on Ω and a sequence of meshes

$$
\{(C^{(n)}; F^{(n)})\}_{n=1}^{1} = \big\{ \{ {}_{i;(n)} \}_{i \in V^{(n)}}; \{ n_{i;j;(n)} \}_{i;j \in V^{(n)}} \big\}_{n=1}^{1}
$$

over Ω_v as in Denition 1.5, satisfying $C^{(n)} \subset P^{(n)}$. Assume that the discretization size $x_{(n)} \to 0$, the meshes $\{(C^{(n)}; F^{(n)})\}_{n=1}^{1}$ satisfy the structural assumptions (1.14) by some uniform constant, and each mesh $(C^{(n)}; F^{(n)})$ is periodic on Ω_v with pattern size uniformly bounded. Moreover, assume that the nite element discretization $(P^{(n)}; N^{(n)})$ satisfy the a priori estimates (1.16) with some uniform constants.

Consider bounded, Lipschitz and concave nonlinearity $g : [0;+1) \to \mathbb{R}$ with $g(0) = 0$. Assume that the initial data u_0 satises $u_0 \in L^1 \cap W^{s;1}(\Omega_v)$ for some $s > 1$, and $\mathrm{dist}(\mathrm{supp}\, u_0; \partial\Omega_v) > 0$. For all $n \in \mathbb{N}_+$, let $(u_{i;(n)}(t))_{i \in V^{(n)}}$ and $(a_{i;j;(n)})_{i;j \in V^{(n)}}$ be the solution of the coupled scheme (1.12), (1.13) and (1.15) solved on $(C^{(n)}; F^{(n)})$ and $(P^{(n)}; N^{(n)})$. Dene

$$
u_{(n)}(t;x) = \sum_{i \in V^{(n)}} {}_{i;(n)}(x)\, u_{i;(n)}(t):
$$

Then there exists $T > 0$ such that

$$u_{(n)} \text{ is compact in } L^1([0;T]\ R^d):$$

Moreover, T could be arbitrarily large by choosing large $_v$ such that dist(supp u_0; @ $_v$) ! 1 .

1.6. Connection between the two settings. We now discuss why the polygon meshes in Section 1.4 can be understood as a special case of the more general setting in Section 1.5. Starting with any polygon mesh $(C;F) = fV_i g_{i2V}; fS_{i;j} g_{(i;j)2E}$ over R^d with discretization size x, one can construct a mesh as in Denition 1.5 through the following process. First, add more cells to $(C;F)$ if necessary, to ensure + B_x S $_{i2V}$ V_i. Second, construct the extended mesh $f_i g_{i2V}; fn_{i;j} g_{(i;j)2E}$ with the cell and face functions

$$_i(x) = \frac{1}{jB_r(0)j} \int_{B_r(0)} 1_{V_i}(x \quad y)\,dy; \quad 8i\ 2\ V;$$

$$n_{i;j}(x) = \int_{S_{i;j}} \frac{1}{jB_r(0)j} 1_{B_r(0)}(x \quad y) N_{i;j}\,dy; \quad 8(i;j)\ 2\ E: \tag{1.17}$$

where $N_{i;j}$ is the unit normal vector of $S_{i;j}$. It is then straightforward to check that if i 2 V and supp $_i$, then $_i$ is indeed a cell function with $fn_{j;i} g_{(j;i)2E}$ as its face functions. Also, one has P $_{i2V}$ $_i(x)$ = 1, 8x 2 and $n_{i;j} = \ n_{j;i}$, 8i; j 2 V. Therefore, this construction does yield a mesh over as in Denition 1.5.

With this construction, the upwind scheme (1.12) for $(f_i g; fn_{i;j} g)$ with coecients (1.13) and the upwind scheme (1.8) for $(fV_i g; fS_{i;j} g)$ are very similar. The conditions supp $_i$ and supp $n_{i;j}$ are now nothing but V_i + B_x and $S_{i;j}$ + B_x . It is also immediate to see that

$$_i = \int {}_i(x)\,dx = jV_ij:$$

Finally, the coecients $a_{i;j}$ in (1.13) (when supp $n_{i;j}$) now read

$$a_{i;j} = \frac{1}{jB_x(0)j} \int_{B_x} \int_{S_{i;j}} b(x+y)\ N_{i;j}{}^{+}\,dydx;$$

which is only slightly dierent from the coecients $a_{i;j}$ in (1.8), though we do emphasize the order of $()^+$ and integration in this formula. Notice that if $b(x)$ is constant, then the integrand $b\ N_{i;j}$ is also constant, hence $a_{i;j}$ given by (1.13) and (1.8) coincide. So, when $b(x)$ has $W^{1;p}$ regularity, we can naturally expect the two ways of determining $a_{i;j}$ to dier only by a term that is vanishing in L^p as discretization size goes to zero.

As mentioned earlier, both compactness results in Theorem 1.3 and Theorem 1.7 are derived by propagating some discrete regularity like (1.7), where the discrete density $(u_{i;(n)}(t))_{i2V^{(n)}}$ are both governed by the upwind scheme, with the coecients originally dened in dierent ways but now formulated all in the setting of Section 1.5. In Section 2, we give the precise denition of such regularity as Denition 2.3 and state the propagation of such regularity by the upwind scheme as Theorem 2.9. Theorem 2.9 can then be applied to prove both Theorem 1.3 and Theorem 1.7.

While the main elements of the proofs rely on the same result, namely Theorem 2.9, we do need to mention that some settings in Theorem 1.3 and Theorem 1.7 are not identical. Apart from the aforementioned choice of $a_{i;j}$, the way we extend discrete density to continuous functions are also slightly dierent: In Theorem 1.3, u_n is dened as piecewise constant on each cell, while in Theorem 1.7, u_n is reconstructed from cell functions and is thus not piecewise constant. Nevertheless, these dierences are only minor issues once all necessary denitions and notations are properly introduced, which we do in Section 2.

Let us also remark that u_n and u_n could be made identical by formally choosing $_i$ $= I_{V_i}$. However such choice would come with some additional issues. The indicator functions are not cell functions according to Denition 1.4 because they are not even continuous. One may still try to understand the gradients ∇_i and $n_{i;j}$ in distributional sense to have that $\nabla_i = \sum_j n_{i;j}$ and

$$\int_{\mathbb{R}^d} f(x)\, n_{i;j}(x)\, dx = \int_{S_{i;j}} f(x)\, N_{i;j}\, dx \quad \forall (i;j) \in E; f \in C_c^1(\mathbb{R}^d;\mathbb{R}^d):$$

In such cases we formally have

$$a_{i;j} = \int_{S_{i;j}} b(x)\, N_{i;j}{}^{+}\, dx:$$

where the extra mollication in the current choice is removed. But this extra mollication appears to be necessary for our formulation. For example integrating on $S_{i;j}$ without any mollication would require trace embedding and in turn more stringent conditions on the mesh, which we try to avoid.

2. Main technical results of the paper. The goal of this section is to introduce the technical setting that we need for our approach and to state the main precise, quantitative results that underlies our compactness results. First, we introduce some necessary notations. Then in subsection 2.2, we introduce the discrete kernel and semi-norm we use to prove compactness, which is modied from the continuous kernel and semi-norm introduced in [5, 4]. We next state Theorem 2.9 about the the propagation of regularity on periodic meshes. This is the main quantitative result in the paper and, in particular, Theorem 1.3 and Theorem 1.7 are deduced from it. The proof of Theorem 2.9 depends on multiple lemmas and theorems, which we state in subsections 3.1, 3.2 and 3.3. But the actual proofs of these lemmas and theorems are postponed to later sections.

2.1. Denitions and notations. Consider a mesh $(C;F)$ over $\mathbb{R}^d$ as in Deni-tion 1.5. We introduce the following notations:

$\overline{V} = \{i \in V : \operatorname{supp}_i \cap = \emptyset\};$ $V = \{i \in V : \operatorname{supp}_i\}$

$g; V = \{i \in V : \exists j;\ j = i \text{ or } (i;j) \in E;\ \text{s.t. } \operatorname{supp}_j \cap = \emptyset\};$

$E = \{(i;j) \in E : \operatorname{supp}_k \cap = \emptyset \text{ for } k = i;j\};$ $E = \{(i;j) \in E : \operatorname{supp} n_{i;j}\}$

$g:$

For a function $f \in L^1_{loc}(\mathbb{R}^d)$ (or $L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d)$), dene the "projection-to-cell operator" $P_C f$ as

$$(P_C f)_i = \begin{cases} \dfrac{1}{_i}\displaystyle\int_{\mathbb{R}^d} f(x)_i(x)\, dx; & \forall i \in V; \end{cases} \tag{2.1}$$

0;   8i 2 (V n V):

Moreover, for $f \in L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d)$, define the "projection-to-face operator" $P_F f$ as

$$(P_F f)_{i,j} = \begin{cases} \displaystyle\int_{\mathbb{R}^d} \big(f(x)\cdot n_{i,j}(x)\big)_+ \,dx, & \forall (i,j)\in E,\\ 0, & \forall (i,j)\in (V^2\setminus E). \end{cases} \tag{2.2}$$

With these notations, the coefficients and initial data in (1.13) can be rewritten as $(a_{i,j})_{i,j\in V} = P_F b$ and $(u_i(0))_{i\in V} = P_C u_0$. Next, we define the discrete divergence of $(a_{i,j})_{i,j\in V}$ as

$$D_k = D\big((a_{i,j})_{i,j\in V}\big)_k = \begin{cases} \displaystyle\frac{1}{|\Omega_k|}\sum_{i\in V} \big(a_{i,k} - a_{k,i}\big), & \forall k\in V,\\ 0, & \forall k\in (\bar V\setminus V). \end{cases} \tag{2.3}$$

The definition of discrete divergence is justified by the following observation: When choosing $(a_{i,j})_{i,j\in V} = P_F b$, one has

$$\begin{aligned} \forall k\in V,\quad D_k &= \frac{1}{|\Omega_k|}\sum_{i\in V}\left(\int_{\mathbb{R}^d} \big(b(x)\cdot n_{i,k}(x)\big)_+ dx - \int_{\mathbb{R}^d} \big(b(x)\cdot n_{k,i}(x)\big)_+ dx\right)\\ &= \frac{1}{|\Omega_k|}\sum_{i\in V}\left(\int_{\mathbb{R}^d} \big(b(x)\cdot n_{i,k}(x)\big)_+ dx - \int_{\mathbb{R}^d} \big(b(x)\cdot n_{i,k}(x)\big)_- dx\right)\\ &= \frac{1}{|\Omega_k|}\sum_{i\in V}\int_{\mathbb{R}^d} b(x)\cdot n_{i,k}(x)\,dx\\ &= -\frac{1}{|\Omega_k|}\int_{\mathbb{R}^d} b(x)\cdot\nabla\chi_k(x)\,dx = \frac{1}{|\Omega_k|}\int_{\mathbb{R}^d} \operatorname{div}_x b(x)\,\chi_k(x)\,dx. \end{aligned}$$

Hence, at least on V,

$$D(P_F b) = P_C\, \operatorname{div}_x b. \tag{2.4}$$

For any $(v_i)_{i\in V}$, we define its discrete L^p norm by

$$\|(v_i)_{i\in V}\|_{L^p(C)} = \left(\sum_{i\in V} |v_i|^p\,|\Omega_i|\right)^{1/p}.$$

In addition, we define the L^p norm for the discretized velocity field $(a_{i,j})_{i,j\in V}$ by

$$\|(a_{i,j})_{i,j\in V}\|_{L^p(F)} = \left(\sum_{i,j\in V} |a_{i,j}|^p\right)^{1/p} (\delta x)^{d/p-(d-1)},$$

where the factor $(\delta x)^{d/p-(d-1)}$ attempts to account for the expected size of the faces.

One motivation to define the discrete norms as above, and especially the scaling factor in $L^p(F)$, is that we can easily bound them by their continuous counterparts. It is easy to verify the following proposition when $p = 1$ and $p = \infty$, and the general case follows by an interpolation.

Proposition 2.1. Let $(C;F)$ be a mesh satisfying (1.14), then one has the following inequalities:

$$\|P_F b\|_{L^p(F)} \le C\|b\|_{L^p(\Omega)},\qquad \|P_C u\|_{L^p(C)} \le C\|u\|_{L^p(\Omega)},$$

where the constant C in the above inequalities only depends on the constants in the structural assumptions (1.14).

In this paper, we consider a sequence of discrete densities $u_{(n)} = (u_{i;(n)}(t))_{i\in V^{(n)}}$ defined on a sequence of meshes $(C^{(n)}; F^{(n)})$. Since the n-th density is always defined on the n-th mesh, as an abuse of notation, we write $\|u_{(n)}\|_{L^p(C^{(n)})}$ as $\|u_{(n)}\|_{L^p(C)}$ for simplicity. Similarly, for the discrete coefficients $a_{(n)} = (a_{i;j;(n)})_{i;j\in V^{(n)}}$ on the meshes we write $\|a_{(n)}\|_{L^p(F^{(n)})}$ as $\|a_{(n)}\|_{L^p(F)}$.

The following notations are also useful in later discussions: For each $i \in V$, define the "barycenter" of cell function χ_i by

$$x_i = \frac{1}{|\chi_i|}\int x\,\chi_i(x)\,dx.$$

For any $(v_i)_{i\in V}$, define its extension to $\mathbb{R}^d$ by

$$v = \sum_{i\in V} v_i\chi_i.$$

2.2. Compactness via quantitative regularity estimates. In this subsection we introduce the explicit semi-norms that we are going to use in the paper, together with lemmas and propositions about some basic properties of those objects. The proof of all lemmas and propositions are postponed to Section 7.

The following continuous kernels and semi-norms are introduced in [5, 4] to prove the compactness of density:

Definition 2.2. Define the kernel K^h for all $h > 0$ by

$$K^h(x) = \frac{\Phi(x)}{(|x| + h)^d}; \quad \forall x \in \mathbb{R}^d;$$

where Φ is some smooth function with compact support in $B(0;2)$ and s.t. $\Phi = 1$ inside $B(0;1)$. Then for $1 \le p < \infty$, $0 < \theta < 1$, the semi-norm $\|\cdot\|_{p;\theta}$ for density $u \in L^p(\mathbb{R})$ is defined as

$$\|u\|^p_{p;\theta} = \sup_{h\le 1/2} |\log h|^{-\theta} \int_{\mathbb{R}^{2d}} K^h(x-y)|u(x)-u(y)|^p\,dxdy. \tag{2.5}$$

We define the corresponding discretization of such kernels and semi-norms.

Definition 2.3. Consider a mesh $(C;F)$ over $\mathbb{R}^d$, on which there exists a discrete density $(u_i)_{i\in V}$ such that $\operatorname{supp} u_i \subset V$. Assume that

$$\sum_{i\in V}\chi_i(x) = 1; \quad \forall x \in \Omega + B(0;4). \tag{2.6}$$

Given any so-called virtual coordinates $\tilde{x} = (\tilde{x}_i)_{i\in V} \in (\mathbb{R}^d)^V$, we define an approximate kernel $\tilde{K}^h_{i;j}$ on the mesh by

$$\tilde{K}^h_{i;j} = K^h(\tilde{x}_i - \tilde{x}_j); \quad \forall i;j \in V:$$

Then for $1 \le p < \infty$, $0 < \theta < 1$ and $0 < h_0 < 1/2$, the discrete semi-norm $\|\cdot\|_{h_0;p;\theta;\tilde{x}}$ on the mesh is defined as

$$\|u\|^p_{h_0;p;\theta;\tilde{x}} = \sup_{h_0\le h\le 1/2} |\log h|^{-\theta} \sum_{i;j\in V} \tilde{K}^h_{i;j}|u_i - u_j|^p|\chi_i||\chi_j|: \tag{2.7}$$

The following lemma is the cornerstone of deriving compactness from the discrete regularity in Denition 2.3, in the particular case where $x 2 (R^d)^V$ in Denition 2.3 is simply chosen by the barycenters, i.e. $x_i = x_i$. In that case, we use the specic notation

$$K^h_{i;j} = K^h(x_i \quad x_j) \quad \text{and} \quad k \; k_{h_0;p;} = k \; k_{h_0;p;;(x_i)_{i2V}} \text{ to}$$

specify the discrete semi-norm derived from the barycenters $(x_i)_{i2V}$.

Lemma 2.4. Let $(C; F)$ be a mesh as in Denition 1.5 over R^d such that (1.14) and (2.6) hold. Consider a discrete function $(u_i)_{i2V}$ on the mesh, satisfying the bound

$$k(u_i)_{i2V} k_{h_0;p;} \quad L$$

for some $0 < h_0 < 1=2$. Dene the renormalized kernel $K^h(x) = K^h(x)=kK^h k_{L^1}$ and let u be the extension of $(u_i(t))_{i2V}$ to R^d. Then

$$8h > h_0; \quad ku \quad K^h \; ? \; u k_{L^p} \; {}^p C j \log h j \quad {}^1$$

where the constant C only depends on L and the constant in structural assumptions (1.14). If the mesh is given by a polygon mesh $(fV_i g; fS_{i;j} g)$ via (1.17), then the above inequality also holds when u is replaced by the piecewise constant extension.

Consider a sequence of meshes $(C^{(n)}; F^{(n)})$ and a sequence of discrete density $u_{(n)} = (u_{i;(n)})_{i2V^{(n)}}$ dened on them. It is just natural to study the compactness of such discrete densities on dierent meshes by some sort of extension on R^d.

To see how Lemma 2.4 helps to derive compactness of such sequence, assume that one has uniform boundedness

$$\sup_{0<h<1=2} \; \limsup_{n \; ! \; 1} ku_{(n)} k_{h_0;1;} < 1 \, ;$$

and uniform boundedness of discrete $L^p(C)$ norm. Then for any xed $h > 0$, the dier-ence between extended functions $u_{(n)}$ and their mollications are uniformly bounded by $C j \log h j$ ^{1} up to discarding nitely many terms of the sequence. On the other hand for any xed h greater than 0, the sequence of mollied functions is locally compact. Therefore, the sequence of extended functions is also locally compact.

However there are several big issues that one should be aware of when moving from the continuous to the discrete setting:

The kernel parameter h has to be bounded from below in Denition 2.3, because a kernel too sharp is not suitable for a coarser grid. Generically h_0 should be chosen much greater than the discretization size x. But for a sequence of meshes with x converging to zero, h_0 could be chosen converging to zero as well (with a possibly much slower speed). As we just discussed below Lemma 2.4, this sort of regularity in asymptotic sense would be sucient to continue our discussion of compactness.

Moreover, in Denition 2.2 the integral is taking on R^{2d}, while in Denition 2.3 we are restricted to a nite double summation. Nevertheless, any kernel K^h has bounded support in ball $B(0; 2)$, hence for any density u with bounded support, the double integral in (2.5) can be taken on $\operatorname{supp} u + B(0; 2)$ ^{2} instead of R^{2d}. Therefore, to reasonably approximate the integral in (2.5), it is natural to made the additional assumption (2.6) for the summation in (2.7). The larger ball

$B(0;4)$ is used for the convenience of later analysis. Starting from a mesh $(C;F)$ over $\mathbb{R}^d$ on which the upwind scheme (1.12) is dened, one can always put additional cell functions to make (2.6) hold. The scheme is not really aected as the density is set as zero at any $i \notin V$. For this reason, when discussing quantitative regularity, we always add (2.6) as part of our assumption to meshes.

The more delicate issue and the one that leads to most technical diculties in this paper is how to choose the virtual coordinates $(\tilde{x}_i)_{i \in V}$. While it would seem natural to take $(\tilde{x}_i)_{i\in V} = (x_i)_{i\in V}$, the corresponding semi-norm does not seem to be propagated well on the scheme (1.12). This will force the use of $(\tilde{x}_i)_{i\in V} = (x_i)_{i\in V}$ to obtain semi-norms that we can propagate well. On the other hand, by Lemma 2.4 we can clearly see compactness from the semi-norms induced by $(x_i)_{i\in V}$, but not from semi-norms induced by arbitrary $(\tilde{x}_i)_{i\in V}$. Therefore, we will also have to show that the approximate kernels $K^h(\tilde{x}_i \quad \tilde{x}_j)$ are equivalent to $K^h(x_i \quad x_j)$, for a choice of virtual coordinates $(\tilde{x}_i)_{i\in V}$ that are only slightly dierent from the barycenters $(x_i)_{i\in V}$.

We can make the last issue somewhat more precise by a more general estimate that consider the discrete kernels as some sort of perturbation of the continuous kernels.

Lemma 2.5. Consider measurable functions $f_i; g_i : \mathbb{R}^d \to \mathbb{R}^d$, $i = 1;2$ and $0 < h_1 < 1{=}4$, such that $|x \quad f_i(x)| \quad h_1$, $|x \quad g_i(x)| \quad h_1$, $\forall x \in \mathbb{R}^d; i = 1;2$. Consider the kernels

$$K_f^h(x;y) = K^h(f_1(x);f_2(y)) = \frac{(|f_1(x) \quad f_2(y)|)}{(|f_1(x) \quad f_2(y)| + h)^d};$$

$$K_g^h(x;y) = K^h(g_1(x);g_2(y)) = \frac{(|g_1(x) \quad g_2(y)|)}{(|g_1(x) \quad g_2(y)| + h)^d}:$$

Then for $1 \quad p < 1$ and $0 < h < 1{=}2$,

$$\int_{\mathbb{R}^{2d}} K_g^h(x;y)|u(x) \quad v(y)|^p\,dxdy \quad (1 + Ch_1{=}h) \int_{\mathbb{R}^{2d}} K_f^h(x;y)|u(x) \quad v(y)|^p\,dxdy; \tag{2.8}$$

where the constant C only depends on the xed choice of in the denition of kernels K^h.

Notice that the double summation in (2.7) can be rewritten as a double integral form by carefully choosing some function $f = f_1 = f_2$ and a piecewise constant, which we state as the next lemma:

Lemma 2.6. Consider a mesh $(C;F)$ as in Denition 1.5 over $\mathbb{R}^d$ with discretization size $x < 1{=}16$, such that (2.6) hold. Introduce some measurable sets $(V_i)_{i\in V} \quad \mathbb{R}^d$ such that

$$|V_i| = {}_i = \int_{\mathbb{R}^d} {}_i; \qquad \sup_{x \in V_i} |x \quad x_i| < 2x; \qquad \forall i \in V; \qquad \text{and} \qquad V_i \cap V_j = \emptyset; \qquad \forall i;j \in V:$$

Dene the piecewise constant extension $u^V = \sum_{i\in V} u_i 1_{V_i}$ for the discrete density function. Then

$$\operatorname{supp} u^V \quad [\quad + B(0;1) \quad + B(0;3) \quad \bigcup_{i\in V} V_i \; :$$

Assume that the virtual coordinates $(\tilde x_i)_{i\in V}$ are such that $|\tilde x_i - x_i| < h_2$, $\forall i\in V$ and for some $0<h_2<1/16$. Define $f:\mathbb{R}^d\to\mathbb{R}^d$ as

$$f(x)=\begin{cases}\tilde x_i, & \text{for } x\in V_i,\ i\in V;\\ x, & \text{for } x\notin \bigcup_{i\in V} V_i.\end{cases}$$

Then $|x-f(x)|\le 2\delta x+h_2<1/4$, $\forall x\in\mathbb{R}^d$. Moreover the double summation in (2.7) can be rewritten as a double integral:

$$\sum_{i,j\in V}\tilde K^h_{i,j}|u_i-u_j|^p\omega_{ij}=\int_{\mathbb{R}^{2d}}K^h\big(f(x)-f(y)\big)|u^V(x)-u^V(y)|^p\,dx\,dy;$$

for all $1\le p<\infty$ and $0<h<1/2$.

From these two lemmas, one may deduce the following proposition supporting our use of $(\tilde x_i)_{i\in V}=(x_i)_{i\in V}$.

Proposition 2.7. Consider a mesh $(\mathcal{C},\mathcal{F})$ with discretization size $\delta x<h_2<1/16$, such that (1.14) and (2.6) hold. Let $(\tilde x^{(1)}_i)_{i\in V}$, $(\tilde x^{(2)}_i)_{i\in V}\in(\mathbb{R}^d)^V$ be two sets of virtual coordinates on the mesh such that

$$\forall i\in V;\ k=1,2;\qquad |\tilde x^{(k)}_i-x_i|<h_2.$$

Then for $1\le p<\infty$, $0<\theta<1$ and $0<h_0<1/2$, the two resulting semi-norms and equivalent and satisfy

$$\|u\|_{h_0,p,\theta,\tilde x^{(2)}}\le(1+Ch_2/h_0)\,\|u\|_{h_0,p,\theta,\tilde x^{(1)}};$$

where the constant C is fixed.

The proposition implies that the equivalence of semi-norms can be derived from the closeness of virtual coordinates. Therefore, a large part of our technical analysis is actually devoted to finding appropriate $(\tilde x_i)_{i\in V}$ ensuring the propagation of regularity while remaining reasonably close to barycenters $(x_i)_{i\in V}$.

The next proposition is also a consequence of Lemma 2.5:

Proposition 2.8. Consider a mesh $(\mathcal{C},\mathcal{F})$ such that (1.14) and (2.6) hold. Then for $1\le p<\infty$,

$$\int_{\mathbb{R}^{2d}}K^h(x,y)|u(x)-u(y)|^p\,dx\,dy\le(1+C\delta x/h)\sum_{i,j\in V}K^h(x_i-x_j)|u_i-u_j|^p\omega_{ij};$$

and

$$\begin{aligned}\sum_{i,j\in V}K^h_{i,j}|(P_{\mathcal{C}}u)_i-(P_{\mathcal{C}}u)_j|^p\omega_{ij}&=\sum_{i,j\in V}K^h_{i,j}\left|\frac{1}{\omega_i}\int_{\mathbb{R}^d}u(x)\phi_i(x)\,dx-\frac{1}{\omega_j}\int_{\mathbb{R}^d}u(y)\phi_j(y)\,dy\right|^p\omega_{ij}\\ &\le(1+C\delta x/h)\int_{\mathbb{R}^{2d}}K^h(x,y)|u(x)-u(y)|^p\,dx\,dy.\end{aligned}$$

for some constant C depending only on p and the constants in the structural assumptions in (1.14).

This proposition ensures that the regularity of the extended function is comparable to the regularity of the discrete density and vice versa, which is needed in our proof of Lemma 2.4 and Theorem 1.7.

2.3. Our main quantitative regularity result. We are now ready to state our main quantitative theorem about the propagation of regularity on periodic mesh.

Theorem 2.9. Consider $T > 0$, and a bounded domain $\Omega \subset \mathbb{R}^d$ with piecewise smooth boundary. Let $\{(C^{(n)}, F^{(n)})\}_{n=1}^{\infty}$ be a sequence of meshes over $\Omega \subset \mathbb{R}^d$ as in Definition 1.5, having discretization size $\delta x_{(n)} \to 0$, satisfying the structural assumptions (1.14) and (2.6) by some uniform constant, and being periodic on Ω with pattern size uniformly bounded.

For all $n \in \mathbb{N}_+$, $t \in [0,T]$, let $(a_{i,j,(n)}(t))_{i,j\in V^{(n)}}$ be the coefficients of the upwind scheme (1.12) on $(C^{(n)}, F^{(n)})$ and let $D_{(n)}(t) = (D_{i,(n)}(t))_{i\in V^{(n)}}$ be the discrete divergence defined as in (2.3). Let $u_{(n)} = (u_{i,(n)}(t))_{i\in V^{(n)}}$ be a sequence of discrete density solved by the upwind scheme. With some $1 \le p < q \le \infty$ and $0 < s \le 1$, assume that there exists a sequence of velocity field $\tilde b_{(n)}(t,x)$, bounded uniformly in $L^q_t(W^{1,q}_x) \cap L^p_t(W^{s,p})([0,T] \times \Omega)$, and approximating the coefficients $(a_{i,j,(n)}(t))_{i,j\in V^{(n)}}$ with vanishing error

$$\left\| (a_{i,j,(n)})_{i,j\in V^{(n)}} - \tilde P_{F^{(n)}} \tilde b_{(n)} \right\|_{L^p([0,T]\times F^{(n)})} \to 0 \quad \text{as } n \to \infty.$$

Assume moreover that the solutions have uniformly bounded norms $\sup_n \|u_{(n)}\|_{L^\infty_t L^p_x([0,T]\times C^{(n)})} < \infty$, and that mass leaking vanishes

$$\|u_{(n)}(0)\|_{L^1(C)} - \|u_{(n)}(T)\|_{L^1(C)} \to 0 \quad \text{as } n \to \infty.$$

Then for all $\theta \ge \max\{1 - 1/q, 1/2\}$, $0 < h_0 < 1/2$, there exists sufficiently large $N \in \mathbb{N}^+$ such that for all $n \ge N$,

$$
\begin{aligned}
\|u_{(n)}(t)\|_{h_0,1,\theta} \le{}& \|u_{(n)}(0)\|_{h_0,1,\theta} \\
&+ C \int_0^t \Big(\|\operatorname{div} \tilde b_{(n)}(s)\|_{L^1(C)} \|u_{(n)}(s)\|_{h_0,1,\theta} + \|\tilde b_{(n)}(s)\|_{W^{1,q}} \|u_{(n)}(s)\|_{L^p(C)} \\
&+ \|D'_{(n)}(s)\|_{L^1(C)} \|u_{(n)}(s)\|_{h_0,1,\theta} + \|u_{(n)}(s)\|_{L^p(C)} \|D_{(n)}(s)\|_{h_0,p,p(\theta-1/p)} \Big) \, ds \\
&+ L_2 + L_3 \ ;
\end{aligned}
\tag{2.9}
$$

with additional terms due to discretization

$$
\begin{aligned}
L_2 &= C\,|\log h_0|/h^2_0\,\varepsilon(x_{(n)})\int_0^t \left(\|a_{(n)}(s)\|_{L^q(F)} + \|\tilde b_{(n)}(s)\|_{W^{1,q}}\right)\|u_{(n)}(s)\|_{L^{q'}(C)}\,ds \\
&\quad + C(|\log h_0|^{-1})\left(\|u_{(n)}(0)\|_{L^1(C)} - \|u_{(n)}(t)\|_{L^1(C)}\right); \\
L_3 &= C(|\log h_0|/h_0)\Big(\|a_{(n)} - P_{F^{(n)}} b_{(n)}\|_{L^p([0,T]\times F)} + (x_{(n)})^{s/(1+s)}\|b_{(n)}\|_{L^p_t W^{s,p}_x} \\
&\quad + (x_{(n)})^{\frac{1+(1/p-1/q)}{1/p-1/q}}\left(\|a_{(n)}\|_{L^q([0,T]\times F)} + \|\tilde b_{(n)}\|_{L^q(W^{1,q})}\right)\Big)\|u_{(n)}\|_{L^\infty_t L^p_x([0,t]\times C)}; \\
\varepsilon &= \exp\left(C(1/h_0)(x_{(n)})^{s/(1+s)}\right).
\end{aligned}
\tag{2.10}
$$

The constant C in (2.9) only depends on Ω, the exponents p,q and the constant in the structural assumption (1.14), while the constant C in (2.10) also depends on T, the exponent s and the constant bounding pattern size. Nevertheless, none of the constants depends on h_0 or $x_{(n)}$. The index $N \in \mathbb{N}^+$ is chosen to make $x_{(n)}$ sufficiently small, which only depends on h_0 and the constant bounding pattern size.

In particular, for any fixed $h_0 > 0$, the additional terms L_2, L_3 converge to zero and ε converges to one as $n \to \infty$.

Proving Theorem 2.9 is the main technical challenge of the paper. We split our proof into three theorems, namely Theorem 3.1, 3.3 and 3.4. These three theorems are stated in subsection 3.1, 3.2 and 3.3 and we conclude Section 3 by how they are used to prove Theorem 2.9. Each of the three theorems requires its own proof on which we spend an entire section after Section 3.

Before we move to the proof, we conclude this section by showing how to deduce Theorem 1.3 and Theorem 1.7 from Theorem 2.9.

Proof of Theorem 1.3. Let us begin with the discussion of mesh properties as Theorem 1.3 is stated in the setting of polygon meshes. We first recall the construction (1.17) in Section 1.6, restated here

$$
\begin{aligned}
\psi_i(x) &= \frac{1}{|B_r(0)|}\int_{B_r(0)} 1_{V_i}(x-y)\,dy; \qquad \forall i \in V; \\
n_{i,j}(x) &= \int_{S_{i,j}} \frac{1}{|B_r(0)|} 1_{B_r(0)}(x-y) N_{i,j}\,dy; \qquad \forall (i,j) \in E.
\end{aligned}
$$

for the entire sequence $\{(C^{(n)}, F^{(n)})\}_{n=1}^\infty$. As an abuse of notation, we still use $\{(C^{(n)}, F^{(n)})\}_{n=1}^\infty$ to denote the generated sequence $\{(\{\psi_{i,(n)}\}_{i\in V^{(n)}}, \{n_{i,j,(n)}\}_{(i,j)\in E^{(n)}})\}_{n=1}^\infty$. It is easy to verify that if the polygon meshes satisfy the structural assumptions (1.9), then the constructed new meshes as in Definition 1.5 satisfy the structural assumptions (1.14), with a possibly larger constant. Moreover, as explained in Section 2.2, one can always put additional cell functions to make (2.6) hold. This yields a sequence of meshes that fulfills the requirements of Theorem 2.9.

Now we define a linear operator P'_F as an alternate of the "projection-to-face operator" P_F, such that in Theorem 1.3, the coefficients of upwind scheme on each $(C^{(n)}, F^{(n)})$ is chosen exactly by $(a_{i;j;(n)})_{i;j\in V^{(n)}} = P^0_{F^{(n)}} b$. Such operator P^0_F is given by

$$(P^0_F f)_{i;j} = \begin{cases} \left(\int_{\mathbb{R}^d} f(x)\cdot n_{i;j}(x)\,dx\right)^+; & \forall (i;j)\in E \\ 0; & \forall (i;j)\in (V^2\setminus E): \end{cases}$$

It is easy to verify the divergence identity $D(P^0_F b) = P_C \operatorname{div}_x b$ by the same approach with which we obtain (2.4). Also, it is straightforward that $\|P^0_F b\|_{L^p(F)} \le \|P_F b\|_{L^p(F)} \le C\|b\|_{L^p(\Omega)}$.

We can then do some a priori estimates for the norms required by Theorem 2.9. To avoid writing too many index (n) in the calculation, let $(C;F) = (\{\Omega_i\}_{i\in V}; \{n_{i;j}\}_{(i;j)\in E})$ be any mesh in the sequence of meshes we consider. Firstly, notice that for $i \in V^\circ$, one has

$$\begin{aligned} \frac{du_i}{dt} &= \frac{1}{|\Omega_i|}\sum_{j\in V}\left(a_{i;j}u_j - a_{j;i}u_i\right) \le \frac{1}{|\Omega_i|}\left(\sup_{k\in V} u_k \sum_{j\in V} a_{i;j} - u_i \sum_{j\in V} a_{j;i}\right) \\ &= -D_i \sup_{k\in V} u_k + \frac{1}{|\Omega_i|}\sum_{j\in V} a_{j;i}\left(\sup_{k\in V} u_k - u_i\right); \end{aligned} \tag{2.11}$$

and for all $i \in V \setminus V^\circ$, one has $u_i \equiv 0$. By the assumption $\operatorname{div}_x b \in L^1_t L^\infty_x$, one can conclude $\|D(t)\|_{L^\infty(C)} \le C\|\operatorname{div}_x b\|_{L^\infty}$ and

$$\frac{d}{dt}\sup_{k\in V} u_k \le \sup_{i\in V}(-D_i)\sup_{k\in V} u_k \le C\|\operatorname{div}_x b\|_{L^\infty}\sup_{k\in V} u_k: \tag{2.12}$$

The constants C just above do not depend on (n), so that $(u_{i;(n)})_{i\in V^{(n)}}$ and $(D_{i;(n)})_{i\in V^{(n)}}$ have uniform a priori bound in $L^\infty_t L^\infty_x([0;T]\times C^{(n)})$. By Hölder estimate one can obtain uniform bound in any $L^\infty_t L^p_x([0;T]\times C^{(n)})$ where $1 \le p \le \infty$.

Secondly, for any $s > 0$ and $1 \quad 1/p$, the semi-norm of the divergence is bounded by

$$\|(D_i(t))_{i\in V}\|_{h_0;p;p(s-1/p)} \le C\|\operatorname{div}_x b(t)\|_{p;p(s-1/p)} \le C\|\operatorname{div}_x b(t)\|_{W^{s;p}}:$$

The first inequality is an application of the divergence identity $D(P^0_F b) = P_C \operatorname{div}_x b$ and Proposition 2.7, while the second inequality is due to Sobolev estimates. Similarly,

$$\|(u_i(0))_{i\in V}\|_{h_0;1;s} \le C\|u_0\|_{W^{s;1}}:$$

The constants C again do not depend on (n). This gives uniform bounds to the semi-norm of $(D_{i;(n)}(t))_{i\in V^{(n)}}$ and $(u_{i;(n)}(0))_{i\in V^{(n)}}$.

We want to apply Theorem 2.9 with $\tilde{b}_{(n)}(t;x) = b(t;x)$, $p = 1$ and $q = \bar{q}$ (recall that $b \in L^{\bar{q}}_t(W^{1;\bar{q}}_x) \cap L^1_t(W^{s;1}_x)$). But the issue remains is that our newly defined P^0_F is not identical to P_F. Hence, the L^1 difference

$$\left\|(a_{i;j;(n)})_{i;j\in V^{(n)}} - P_{F^{(n)}}\tilde{b}_{(n)}\right\|_{L^1([0;T]\times F^{(n)})} = \left\|P^0_{F^{(n)}} b - P_{F^{(n)}} b\right\|_{L^1([0;T]\times F^{(n)})}$$

is not zero, and one has to argue it converges to zero as $n \to \infty$ and as $x_{(n)} \to 0$. This is guaranteed by the following proposition whose proof is postponed to Section 7.

Proposition 2.10. *Let (C, F) be a mesh as in Definition 1.5 over $\mathbb{R}^d$ such that (1.14) hold. Assume that each face function $n_{i,j} \in F$ is of form $n_{i,j}(x) = N_{i,j} w_{i,j}(x)$, $\forall x \in \mathbb{R}^d$, where $N_{i,j} \in S^{d-1}$ is a unit vector and $w_{i,j}$ is a scalar function. Then for $1 \le p \le \infty$,*

$$\|P_F \tilde{b} - P_F b\|_{L^p([0,T]F)} \le C\,\delta x\,\|b\|_{L^p_t(W^{1,p}_x)},$$

where the constant C only depends on p and the constant in the structural assumption (1.14).

We only have to observe that the construction (1.17) indeed ensures that any $n_{i,j} \in F$ is of form $n_{i,j}(x) = N_{i,j} w_{i,j}(x)$. Hence this proposition applies to the setting of Theorem 1.3. And we can finally apply Theorem 2.9 to obtain (2.9) with $\tilde{b}_{(n)}(t,x) = b(t,x)$, $p = \infty$ and $q = \bar{q}$.

By Gronwall estimate one can conclude

$$C^1_{\log,\theta} = \sup_{0<h_0<1/2} \limsup_{n \to \infty} \|u_{(n)}(t)\|_{h_0,1,\theta} < \infty$$

for some $0 < \theta < 1$. This directly implies compactness in space of the density $u_{(n)}$. Compactness in time now follows by reproducing the Aubin-Lions argument in the semi-discrete setting.

For any $0 < h < 1/2$, let $h_0 = h$. By the previous estimates, one can choose $N(h) \in \mathbb{N}_+$ such that

$$\sup_{n \ge N(h)} \|u_{(n)}(t)\|_{h,1,\theta} < 2C^1_{\log,\theta}.$$

By Lemma 2.4, one has

$$\|K_h \star_x u_{(n)}(t) - u_{(n)}(t)\|_{L^1} \le C|\log h|^{-\theta} \quad \text{for } n \ge N(h),\ t \in [0,T]; \tag{2.13}$$

where C depends on $C^1_{\log,\theta}$ and the total mass of $u_{(n)}$.

On the other hand for any fixed $h > 0$, $U_{h,C} = \{K_h \star u : \|u\|_{L^1} < C,\ \mathrm{supp}\, u\}$ is a compact set by Arzelà-Ascoli theorem, on which we consider $t \to K_h \star_x u_{(n)}(t)$, the trajectory of mollified density. Notice that

$$V_h(t) = \int_{\mathbb{R}^d} \left|\frac{d}{dt} K_h \star_x u_{(n)}(t,x)\right| dx = \int_{\mathbb{R}^d} \left|\sum_{i \in V(n)} \left(\int_{C_{i,(n)}} K_h(y-x)\,dy\right) \frac{d}{dt} u_{i,(n)}(t)\right| dx.$$

By (1.2) we can bound $V_h(t)$ by

$$
\begin{aligned}
V_h(t) &= \int_{\mathbb{R}^d} \Big| \sum_{i\in V^{(n)}} \int_{C_{i;(n)}} K^h(y-x)\,dy \; \frac{1}{|C_i|} \sum_{i':(i;i')\in E^{(n)}} \left(a_{i;i';(n)}\, u_{i';(n)} - a_{i';i;(n)}\, u_{i;(n)} \right) \Big|\, dx \\
&= \int_{\mathbb{R}^d} \Big| \sum_{(i;i')\in E^{(n)}} \left(\frac{1}{|C_i|}\int_{C_{i;(n)}} K^h(y-x)\,dy - \frac{1}{|C_{i'}|}\int_{C_{i';(n)}} K^h(z-x)\,dz \right) a_{i;i';(n)}\, u_{i';(n)} \Big|\, dx \\
&\le \sum_{(i;i')\in E^{(n)}} \int_{\mathbb{R}^d} \left| \frac{1}{|C_i|}\int_{C_{i;(n)}} K^h(y-x)\,dy - \frac{1}{|C_{i'}|}\int_{C_{i';(n)}} K^h(z-x)\,dz \right| dx \; a_{i;i';(n)}\, u_{i';(n)} \\
&\le \sum_{(i;i')\in E^{(n)}} L_h\, \delta x_{(n)}\; a_{i;i';(n)}\, u_{i';(n)} \\
&\le C L_h \|b(t)\|_{L^1} \|u_{(n)}(t)\|_{L^1(C^{(n)})};
\end{aligned}
$$

where L_h denotes the Lipschitz constant of K^h and C depends on the constant in Proposition 2.1. Since we assume $b \in L^1_t L^1_x$ and have obtained uniform a priori bound of $(u_{i;(n)})_{i\in V^{(n)}}$ in $L^1_t L^1_x([0;T] \times C^{(n)})$, we have $V_h(t)$ uniformly bounded in t.

Therefore, for any fixed $h > 0$, $K^h \ast u_{(n)}$ is equicontinuous as a trajectory in $U_{h;C} \subset L^1(\mathbb{R}^d)$. By Arzela–Ascoli theorem, $\{K^h \ast u_{(n)}\}$ is compact for all $h > 0$, which implies $\{u_{(n)}\}$ is also compact thanks to (2.13).

We now turn to the proof of Theorem 1.7 which requires more work but follows somewhat similar steps.

Proof of Theorem 1.7. As in the proof of Theorem 1.3, our goal is to apply Theorem 2.9. As a comparison, this time we begin with a sequence of meshes as in Definition 1.5 so the mesh properties are obvious, but since we are discussing a coupled system, we also need to actually derive regularity estimates on the velocity field.

As before, when there is no ambiguity, we omit the index (n) by letting $(C;F);(P;N)$ be any pair of mesh and finite element in the sequence, and let $(u_i)_{i\in V}$, $(a_{i;j})_{i;j\in V}$ be the discrete solution.

Step 1: Discrete a priori bounds. The discrete divergence at $i \in V$ is given by

$$
\begin{aligned}
D_i &= \frac{1}{|C_i|} \sum_{i'\in V} \left(a_{i';i} - a_{i;i'} \right) = \frac{1}{|C_i|} \sum_{i'\in V} \int_{\mathbb{R}^d} \left(\left(b(y)\cdot n_{i';i}(y)\right)^{+} - \left(b(y)\cdot n_{i';i}(y)\right)^{-} \right) dy \\
&= \frac{1}{|C_i|} \sum_{i'\in V} \int_{\mathbb{R}^d} b(y)\cdot n_{i';i}(y)\,dy = \frac{1}{|C_i|} \int_{\mathbb{R}^d} b(y)\cdot\left(-\nabla\varphi_i(y)\right) dy \\
&= \frac{1}{|C_i|} \int_{\mathbb{R}^d} g(y)\,\varphi_i(y)\,dy.
\end{aligned}
$$

The last identity is due to the assumption that $\varphi_i \in C \cap P$ and (1.15).

By Proposition 2.8, for all $h \leq x$, one has

$$\sum_{i,j\in V} K^h_{i,j}\,|D_i - D_j|\,\omega_i\omega_j = \sum_{i,j\in V} K^h_{i,j}\left|\frac{1}{\omega_i}\int_{\mathbb{R}^d} g(x)\chi_i(x)\,dx - \frac{1}{\omega_j}\int_{\mathbb{R}^d} g(y)\chi_j(y)\,dy\right|\omega_i\omega_j$$
$$\leq C\int_{\mathbb{R}^{2d}} K^h(x,y)\,|g(x) - g(y)|\,dx\,dy$$
$$\leq C\sum_{i,j\in V} K^h_{i,j}\,|g(u_i) - g(u_j)|\,\omega_i\omega_j$$
$$\leq C L_g \sum_{i,j\in V} K^h_{i,j}\,|u_i - u_j|\,\omega_i\omega_j.$$

where L_g be the Lipschitz constant of g. Hence,

$$\|(D_i(t))_{i\in V}\|_{h_0,1,\omega} \leq C L_g\,\|(u_i(t))_{i\in V}\|_{h_0,1,\omega} \tag{2.14}$$

Also, from the above discussion it is straightforward to see that

$$D_i(t) = \frac{1}{\omega_i}\int_{\mathbb{R}^d} g(y)\,\chi_i(y)\,dy = \frac{1}{\omega_i}\int_{\mathbb{R}^d}\sum_{j\in V} g\big(u_j(t)\big)\chi_j(y)\,\chi_i(y)\,dy$$
$$= \sum_{j\in V} A_{i,j}\,g\big(u_j(t)\big),$$

where the coefficients satisfies

$$A_{i,j} \geq 0, \qquad \sum_{j\in V} A_{i,j} = 1.$$

Thus

$$\sup_{i\in V} D_i(t)_{i\in V} \leq \sup\Big\{ g(a) \,\big|\, a \in [\inf_{i\in V} u_i(t)_{i\in V},\ \sup_{i\in V} u_i(t)_{i\in V}]\Big\},$$
$$\inf_{i\in V} D_i(t)_{i\in V} \geq \inf\Big\{ g(a) \,\big|\, a \in [\inf_{i\in V} u_i(t)_{i\in V},\ \sup_{i\in V} u_i(t)_{i\in V}]\Big\}.$$

Since we assumed that $g \in L^1(\mathbb{R})$, this means that the divergence is bounded uniformly in n. By (2.11) and (2.12), we also obtain a uniform in n bound of $(u_i)_{i\in V}$ in $L^1([0,T]\,C)$ for any $T > 0$.

Recalling moreover that ϕ is the solution of $-\Delta\phi = \varrho$ (with Dirichlet BC on Ω_e), one also has a uniform bound on $\nabla\phi$ in $L^\infty_t L^2_x \cap L^2_t H^1_x$.

Step 2: Control of mass leaking by Markovian interpretation. The discrete scheme can also be represented by a Poisson random process model. Without loss of generality, we may assume that the total mass $\sum_{i\in V} u_i(0)\omega_i = 1$ and define the initial condition of the random process $X(t)$ by $P\{X(0) = i\} = u_i(0)\omega_i$. We choose the rate of the Poisson process as

$$\lambda_{i',i}(t) = \frac{a^+_{i',i}(t)}{\omega_i}.$$

Define now the stopping time and number of jumps through

$$\tau = \inf\{t : X(t) \notin V_v\};$$
$$N(t) = \sup\{N : 0 = s_0 < s_1 < \cdots < s_N \leq t;\ \forall 1 \leq i \leq N,\ X(s_i) \neq X(s_{i-1})\}.$$

Then we have a straightforward identity of the density for $t > 0$ by

$$u_i(t) = \frac{1}{|\Omega_i|} P\{\tau > t;\ X(t) = i\}.$$

By the fact that $\nabla\phi$ is bounded in L^∞ and (1.16), one can conclude that $b = \nabla\phi$ is also uniformly bounded in L^∞. Denote M_b the a priori bound of $\|b\|_{L^\infty}$, then

$$\lambda_{i_0;i}(t) = \frac{a_{i,i_0}(t)}{|\Omega_i|\,\delta x} \le \frac{M}{\delta x},$$

where $M = CM_b$.

Define $L_{\partial\Omega} = \mathrm{dist}(\mathrm{supp}\, u_0; \partial\Omega)$. By its definition, $\mathrm{dist}(X_0; \partial\Omega) \ge L_{\partial\Omega}$, so that it requires at least $L_{\partial\Omega}/\delta x$ jumps to reach the boundary, i.e.

$$\tau \le t \quad \text{only if} \quad N(t) \ge \frac{L_{\partial\Omega}}{\delta x}.$$

In particular, one can bound the probability of $\tau \le t$ by a homogeneous Poisson process, i.e.

$$P\{\tau \le t\} \le \mathcal{P}\left(\frac{M}{\delta x}; \frac{L_{\partial\Omega}}{\delta x} - 1\right);$$

where $\mathcal{P}(\lambda; \cdot)$ denotes the probability distribution of a homogeneous Poisson process with rate λ and starting from 0.

Consider $T > 0$ such that $L_{\partial\Omega} - MT > 0$, and let $\delta_T = (L_{\partial\Omega} - MT)/2$. Then one can deduce

$$P\{\tau \le T\} \le C\exp(-\delta_T/\delta x);$$

by standard estimates for homogeneous Poisson processes.

Hence choosing T s.t. $\delta_T > 0$, we obtain the mass leaking estimate, for any $t \le T$

$$0 \le 1 - \sum_{i\in V} |\Omega_i|\, u_i(t) \le C\exp(-\delta_T/\delta x) \to 0, \qquad \text{as } n \to \infty.$$

Step 3: Regularity of the continuous velocity field. We choose the continuous velocity field in Theorem 2.9 as $\tilde b = -\nabla\tilde\phi$, $-\Delta\tilde\phi = \tilde g$ (with Dirichlet BC on $\partial\tilde\Omega$). By our discussion in Step 1, we have uniform in n bounds on $\tilde b = -\nabla\tilde\phi$ in $L^\infty_t L^\infty_x \cap L^1_t H^1_x$.

Moreover, by our assumption (1.16) on the finite elements, one can bound the L^2 difference between $(a_{i,j})_{i,j\in V} = P_F b$ and $P_F \tilde b$ by

$$\begin{aligned}\|(a_{i,j})_{i,j\in V} - P_F\tilde b\|_{L^1_t L^2_x([0,T]\times F)} &= \|P_F b - P_F\tilde b\|_{L^1_t L^2_x([0,T]\times F)}\\ &\le C\|\nabla\phi - \nabla\tilde\phi\|_{L^1_t L^2_x} + C\,\delta x\,\|\tilde\phi\|_{L^1_t H^2_x};\end{aligned}$$

which converges to zero as $n \to \infty$ and $\delta x \to 0$.

We can also show that our choice of $\tilde b$ has Sobolev regularity in time, namely $\tilde b = -\nabla\tilde\phi \in W^{s,1}_t L^1_x([0,T]\times\tilde\Omega)$ for any $s < 1$. Notice that

$$-\Delta\tilde\phi = \tilde g = \sum_{i\in V} \chi_i\, g(u_i);$$

so that the issue is to show that $\sum_{i\in V} \chi_i\, g(u_i) \in W^{1,1}_t W^{-1,1}_x(\tilde\Omega)$.

Any solution u of the first-order scheme (1.12) satisfies the following identity for all i, ^{d}

$$\frac{d}{dt}u_i = \frac{1}{\lambda_i}\sum_{j\in V}\left(a_{i,j}u_j - a_{j,i}u_i\right) - 1_{V\setminus V_v}(i)\frac{1}{\lambda_i}\sum_{j\in V} a_{i,j}u_j.$$

When $i \in V_v$, the above equality is exactly the upwind scheme. When $i \in (V\setminus V_v)$, one has $u_i \equiv 0$ and the above equality reduce to $0 = 0$. Therefore,

$$\frac{d}{dt}g(u_i) = g'(u_i)\frac{du_i}{dt} = \left(g'(u_i)\frac{1}{\lambda_i}\sum_{j\in V}\left(a_{i,j}u_j - a_{j,i}u_i\right)\right) - \left(1_{V\setminus V_v}(i)g'(u_i)\frac{1}{\lambda_i}\sum_{j\in V}a_{i,j}u_j\right)$$
$$= G_i^{int} - G_i^{bd}.$$

The term G_i^{bd} measures the possible leaking at boundary. It is non-negative and from the previous step, it satisfies

$$\int_{\mathbb{R}^d}\sum_i G_i^{bd}(t)\,\psi_i(x)\,dxdt = \sum_{i\in V} G_i^{bd}(t)\lambda_i\,dt \le C L_g \exp(\,{}_T/x);$$

which implies that $\sum_{i\in V} G_i^{bd}\lambda_i \in L^1_t L^1_x$.

In addition, the term G_i^{int} can be reformulated as

$$G_i^{int} = g'(u_i)\frac{1}{\lambda_i}\sum_{j\in V}\left(a_{i,j}u_j - a_{j,i}u_i\right)$$
$$= \frac{1}{\lambda_i}\sum_{j\in V}\left(a_{i,j}g(u_j) - a_{j,i}g(u_i)\right) + \frac{1}{\lambda_i}\sum_{j\in V}\left(a_{i,j} - a_{j,i}\right)\left(g'(u_i)u_i - g(u_i)\right)$$
$$+ \frac{1}{\lambda_i}\sum_{j\in V} a_{i,j}\left(g'(u_i)(u_j - u_i) - [g(u_j) - g(u_i)]\right)$$
$$= G_i^A + G_i^B + G_i^M.$$

Since $|G_i^B| \le 2\sup_i |D_i| L_g M \le 2(L_g M)^2$, it is straightforward that $\sum_{i\in V} G_i^B\lambda_i \in L^\infty_t L^1_x$. Also, by the concavity of nonlinearity g, one has $G_i^M \ge 0$. Furthermore,

$$\int_0^T\int_\Omega \sum_i G_i^M(t)\,\psi_i(x)\,dxdt$$
$$= \int_0^T \sum_{i\in V} G_i^M(t)\lambda_i\,dt$$
$$= \int_0^T\sum_{i\in V}\left(\frac{d}{dt}g(u_i(t)) - G_i^A(t) - G_i^B(t) + G_i^{bd}(t)\right)\lambda_i\,dt$$
$$\le \sum_{i\in V}\left(g(u_i(T)) - g(u_i(0))\right)\lambda_i + 0 + \int_0^T\sum_{i\in V} G_i^B(t)\lambda_i\,dt + \int_0^T\sum_{i\in V} G_i^{bd}(t)\lambda_i\,dt$$
$$\le 2L_g M + 2T(L_g M)^2 + CTL_g\exp(\,{}_T/x);$$

where in the first inequality we use the observation that $\sum_{i\in V} G_i^A(t)\lambda_i = 0$. As a consequence, one has $\sum_{i\in V} G_i^M\lambda_i \in L^1_t L^1_x$.

Finally, for any test function $\varphi \in W^{1,\infty}$,

$$
\begin{aligned}
&\int \varphi(x) \sum_{i\in V} G^A_i(t)\chi_i(x)\,dx \\
&= \int_{\Omega_e} \varphi(x) \sum_{i;j\in V} \frac{1}{|\Omega_i|}\big(a_{i;j}g(u_j(t)) - a_{j;i}g(u_i(t))\big)\chi_i(x)\,dx \\
&= \sum_{e}\sum_{i;j\in V} \big(a_{i;j}g(u_j(t)) - a_{j;i}g(u_i(t))\big)\frac{1}{|\Omega_i|}\int \varphi(x)\chi_i(x)\,dx \\
&= \sum_{e}\sum_{i;j\in V} a_{i;j}g(u_j(t))\left(\frac{1}{|\Omega_i|}\int \varphi(x)\chi_i(x)\,dx - \frac{1}{|\Omega_j|}\int \varphi(y)\chi_j(y)\,dy\right) \\
&= \sum_{e}\sum_{(i;j)\in E_e} a_{i;j}g(u_j(t))\frac{1}{|\Omega_i|}\frac{1}{|\Omega_j|}\int\int \big(\varphi(x) - \varphi(y)\big)\chi_i(x)\chi_j(y)\,dxdy\ .
\end{aligned}
$$

By our structural assumptions, the number of terms in the last sum is at most $C(x)^{d}$ and each term is bounded by $C(x)^d L_g M^2 \|\varphi\|_{W^{1;\infty}}$. Hence

$$\left|\int_{\Omega_e} \varphi(x) \sum_{i\in V} G^A_i(t)\chi_i(x)\,dx\right| \le C L_g M^2 \|\varphi\|_{W^{1;\infty}};$$

which means $\sum_{i\in V} G^A_i\chi_i \in L^\infty_t W^{-1;1}_x$.

By combining all estimates above, we conclude that

$$\frac{d}{dt}g = \frac{d}{dt}\sum_{i\in V}\chi_i g(u_i) \in L^\infty_t W^{-1;1}_x:$$

It is also straightforward that $g \in L^\infty_t L^\infty_x$. Thus one indeed has that $g \in W^{1;1}_t W^{-1;1}_x$, which implies that $\theta = \nabla \varphi \in W^{s;1}_t L^1_x$ for any $s < 1$.

Step 4. (Compactness) Combine the previous results and apply them to $(C^{(n)}; F^{(n)})$ and $(P^{(n)}; N^{(n)})$ for all $n \in \mathbb{N}_+$. Since all functions are defined on a bounded domain, Sobolev embeddings also directly apply. We may then use Theorem 2.9 with $p = 1$, $q = 2$ and and $s < 1$, yielding the following asymptotic estimate for $\theta > 1/2$,

$$
\begin{aligned}
\limsup_{n\to\infty} \|u_{(n)}(t)\|_{h_0;1;\theta} \le{}& \limsup_{n\to\infty} \|u_{(n)}(0)\|_{h_0;1;\theta} \\
&+ C\int_0^t \Big(\|u_{(n)}(s)\|_{L^1(C)}\|u_{(n)}(s)\|_{h_0;1;\theta} + \|b_{(n)}(s)\|_{W^{1;1}}\|u_{(n)}(s)\|_{L^1(C)}\Big)\,ds\ ;
\end{aligned}
$$

where we used step 1 to bound the discrete divergence terms in Theorem 2.9 by the corresponding bound on $u_{(n)}$.

By Gronwall estimate we conclude that

$$C^1_{\log;\theta} = \sup_{0<h_0<1/2}\ \limsup_{n\to\infty} \|u_{(n)}(t)\|_{h_0;1;\theta} < \infty:$$

The last part is to argue that $t \to K_h \star_x u_{(n)}(t)$, the trajectory of mollied extended density, is equicontinuous on a compact subset of space $L^1(\mathbb{R}^d)$, which is performed in the same manner as in the proof of Theorem 1.3.

3. Proving Theorem 2.9. In this section we start the proof of Theorem 2.9, which is spread into Section 3, 4, 5 and 5. We introduce in this section three theorems that each corresponds to a specific step and explain why they together prove Theorem 2.9.

The first step in the proof is naturally an estimate of the time evolution of $\|u(t)\|_{h_0;1;}$, which we state as Theorem 3.1. Most terms in the estimate behave as one can expect from the continuous model. However, there is one additional term involving what we call a residue $r_i(t)$ which given by linear equation (3.2) and restated here,

$$\sum_{i'\in\partial\{i\}} (\tilde{x}_{i'}^{(k)} - \tilde{x}_i^{(k)})\, a_{i';i}(t) = b_i(t)|\Omega_i| + r_i(t)|\Omega_i|; \forall i \in V;\ t \in [t_{k-1}; t_k];\ 1 \le k \le m:$$

We need to control this residue to conclude the bound on $\|u(t)\|_{h_0;1;}$ through Gronwall lemma. The study of the residue is where our proof fully deviates from the continuous setting.

In essence the size of the residue follows from the choice of virtual coordinates $\tilde{x}_i$. We correspondingly introduce two theorems: The first one identifies some good assumptions for the virtual coordinates to make the residue small, which we state as Theorem 3.3. The second theorem 3.4 shows that virtual coordinates satisfying such assumptions actually exist, at least where the mesh has periodic patterns.

3.1. Step 1: Propagation of regularity in the discrete setting. Our first result reproduces the propagation of regularity in [4] for scheme (1.12) but with additional terms caused by the discretization. The proof of the theorem is postponed to Section 4.

Theorem 3.1. Consider the semi-discrete scheme (1.12) on a mesh $(C; F)$ over a bounded domain $\Omega \subset \mathbb{R}^d$ with piecewise smooth boundary as in Definition 1.5, having discretiza-tion size δx and satisfying the structural assumptions (1.14) and (2.6). Let $(a_{i;j}(t))_{i;j\in V}$ be the coefficients of scheme (1.12) and $D(t) = (D_i(t))_{i\in V}$ be the discrete divergence given by (2.3). Let $b(t; x)$ be a continuous velocity field on $\mathbb{R}^d$ and denote its discretization by $(b_i(t))_{i\in V} = P_F\, b(t;\cdot)$.

Choose $M; \tilde{M} > 0$, such that $\delta x \le M \le \tilde{M} < 1/32$. Divide the time interval $[0; T]$ as $0 = t_0 < t_1 < \cdots < t_m = T$. For each interval $[t_{k-1}; t_k]$, let $(\tilde{x}_i^{(k)})_{i\in V}$ be virtual coordinates on the mesh satisfying

$$|\tilde{x}_i^{(k)} - \tilde{x}_{i'}^{(k)}| \le 2M; \qquad \forall (i; i') \in E; \tag{3.1a}$$

$$|\tilde{x}_i^{(k)} - x_i| \le 2\tilde{M}; \qquad \forall i \in V: \tag{3.1b}$$

Let $(r_i(t))_{i\in V}; t \in [0; T]$ be the residue function given by

$$\sum_{i'\in\partial\{i\}} (\tilde{x}_{i'}^{(k)} - \tilde{x}_i^{(k)})\, a_{i';i}(t) = b_i(t)|\Omega_i| + r_i(t)|\Omega_i|; \quad \forall i \in V;\ t \in [t_{k-1}; t_k];\ 1 \le k \le m: \tag{3.2}$$

Let $k(t) = \min\{k : t < t_k\}; \forall t \in [0; T]$. Then any solution $u(t) = (u_i(t))_{i\in V}$, $t \in [0; T]$ of the semi-discrete scheme (1.12), satisfies for $0 < h_0 < 1/2$

$$\|u(t)\|_{h_0;1;} \le (L_0 + L_1 + L_2 + L_3); \tag{3.3}$$

where

$$L_0 = ku(0)k_{h_0;1;} \quad = \quad 1 + C(M=h_0)_k^{(t)+1};$$

$$L_1 = C \int_0^t k\,\mathrm{div}\,b(s)k_{L^1(C)} ku(s)k_{h_0;1;} + kb(s)k_{W^{1;q}} ku(s)k_{L^p(C)}$$

$$+ kD(s)k_{L^1(C)} ku(s)k_{h_0;1;} + ku(s)k_{L^p(C)} kD(s)k_{h_0;p;p(1=p)} \quad ds$$

$$L_2 = C \int_0^t j\log h_0 j \; M^2=h^2 x k(a_{i;j}(s))_{i;j2V} k_{L^q(F)} ku(s)k_{L^q(C)} + j\log h_0 j \; M=h^2 kb(s)k_{L^q} ku(s)k_{L^q(C)} \tag{3.4}$$

$$+ (j\log h_0 j \; x=h^2) kb(s)k_{W^{1;q}} ku(s)k_{L^q(C)} \quad ds$$

$$+ C(j\log h_0 j^1 \;) \; ku(0)k_{L^1(C)} \quad ku(t)k_{L^1(C)} \;;$$

$$L_3 = C \int_0^t (j\log h_0 j \; =h_0) k(r_i(s))_{i2V} k_{L^p(C)} ku(s)k_{L^p(C)} \, ds;$$

provided that $1 \; p < q \; 1$, $\max f1 \; 1=q; 1=2g$, $M \; h_0 < 1=2$. The constant C depends on
, the exponents $p; q$ and the constant in structural assumptions (1.14).

What we have exhibited in this subsection is a rather incomplete result. The terms ; L_2 and L_3 in Theorem 3.1 are due to the discretization error. The term L_2 would tend to zero as the discretization size $x \; ! \; 0$, provided that the virtual coordinates are chosen such that $\max fM=(x)^{1=2}; M^{1=2} g \; ! \; 0$ and h_0 are chosen decaying to zero with an even slower speed.

Unfortunately, L_3 will not vanish so easily. To eliminate it asymptotically, one should nd a sophisticated way to determine suitable division $0 = t_0 < \; < t_m = T$ and virtual coordinates $(e_i^{(k)})_{i2V}$ on each $[t_{k\;1}; t_k]$, making the norm of residue $krk_{L^1_t L^1_x([0;T]C)}$ decay. To control the term , one should also control the number m in the division of $[0; T]$.

Finally we emphasize that we do not need to assume that the coecients $a_{i;j}(t)$ be given from the discretization of the velocity eld $\tilde{b}$. The connection between $a_{i;j}$ and $\tilde{b}$ stems only from the denition of the residue in (3.2). And in fact, as we will see in the conclusion of the proof, the $a_{i;j}$ are typically derived from a slightly dierent eld $b = \tilde{b}$.

3.2. Step 2: Controlling the residue through virtual coordinates. It remains to investigate in what circumstance Theorem 3.1 give useful results, in the sense that all the additional terms due to the discretization error vanish asymptotically. The main question is how to control the so-called residue through a proper selection of virtual coordinates. We rst introduce the key notion of admissible family of virtual coordinates that works for any constant eld.

Definition 3.2. Consider a mesh $(C; F)$ over $\mathbb{R}^d$. Let

$$(\hat{x}_i(b_c))_{i2V} \; 2 \; (\mathbb{R}^d)^V; \quad 8 b_c \; 2 \; \mathbb{R}^d$$

be a family of virtual coordinates. We say that it is an admissible family of virtual coordinates on
for constant elds with relative drift M, absolute drift M and residue bound M in L^p if the following properties are true:

(1) For each n 2 N_+ and any vector b_c 2 R^d, one has

$$\hat{x}_i(b_c) = \hat{x}_i(b_c); \qquad 8 > 0; i\ 2\ V:$$

(2) The following bounds hold:

$$\sup_{jb_cj=1;(i;i^0)2E} \hat{x}_i(b_c) \quad \hat{x}_{i^0}(b_c) \ M;$$

$$\sup_{jb_cj=1;i2V} j\hat{x}_i(b_c) \quad x_i j \ M:$$

(3) For any vector b_c 2 R^d, let $(a_{i;j}(b_c))_{i;j2V}$ = $P_F b$ and $(b_i(b_c))_{i2V}$ = $P_C b$ be the discretization of the constant velocity eld $b(x)$ b_c. Dene $(\hat{r}(b_c)_i)_{i2V}$ as

$$\sum_{i^0 2V} \left(\hat{x}_{i^0}\ b_c \quad \hat{x}_i(b_c) \right) a_{i^0;i}(b_c) = b_i(b_c)_i + \hat{r}(b_c)_i; \qquad 8i\ 2\ V: \tag{3.5}$$

Dene the maximal residue function $((\hat{r}_{max})_i)_{i2V}$ by

$$(\hat{r}_{max})_i = \sup_{jb_cj=1} jr_i(b_c)j; \quad 8i\ 2\ V;$$

then its L^p norm is bounded by

$$k(\hat{r}_{max})_i k_{L^p(C)} \ j \ j^{1=p} M:$$

By assuming that an admissible family of virtual coordinates exist, we have the following theorem that controls the residue for any Sobolev velocity eld and whose proof is performed in Section 5.

Theorem 3.3. Consider the semi-discrete scheme (1.12) on a mesh (C; F) over a bounded domain
R^d with piecewise smooth boundary as in Denition 1.5, having discretiza-tion size x and satisfying the structural assumptions (1.14) and (2.6). Let $(a_{i;j}(t))_{i;j2V}$ be the coecients of scheme (1.12) and D(t) = $fD_i(t)g_{i2V}$ be the discrete divergence given by (2.3). Assume that u(t) = $(u_i(t))_{i2V}$, t 2 [0; T] is a solution of the semi-discrete scheme (1.12).

Moreover, with some 1 p < q 1 and 0 < s 1, assume that there exists a continuous velocity eld b(t; x) bounded in $L^q(W^{1;q}_t)_x \setminus L^p(W^{s;p}_x)_t([0;T]$), and an admissible family of virtual coordinates on
for constant elds, dened in Denition 3.2, with relative drift M, absolute drift M and residue bound M in $L^{1=(1=p\ \ 1=q)}$. In addition, let maxf1 1=q; 1=2g, and assume that M h_0 < 1=2.

Then one can choose a division of time interval [0; T] as 0 = t_0 < t_1 < < t_m = T, and on each interval $[t_{k\ 1}; t_k]$, virtual coordinates $(x^{(k)}_i)_{i2V}$, such that the terms and

L_3 in estimate (3.3) are bounded by

$$
\begin{aligned}
L_3 = C(|\log h_0|/h_0) \Big[& \big\| (M/\delta x)(a_{i;j})_{i;j\in V} - P_F b \big\|_{L^p([0;T]F)} \\
& + (M^s M/\delta x)^{1/(1+s)} \|b\|_{L^1_t(W^{s,p}_x)} \\
& + M(\delta x)^{1+(1/p-1/q)} \|(a_{i;j}(t))_{i;j\in V}\|_{L^q([0;T]F)} + \|b\|_{L^q_t(W^{1;q}_x)} \Big] \\
& + M\|b\|_{L^q_t(L^q_x)} \|u\|_{L^1_t L^p_x([0;t]C)}; \\
\exp & \big(C(1/h_0)(M^s M/\delta x)^{1/(1+s)} \big);
\end{aligned}
\tag{3.6}
$$

where the constant C depends on T; $\ldots$, the exponents $p; q; s$, and the constant in structural assumption (1.14).

3.3. Step 3: Constructing admissible virtual coordinates for periodic meshes. It remains to study how to find admissible virtual coordinates for constant fields as in Definition 3.2. At this moment, it is still unclear to us whether this is possible for any arbitrary mesh. Nevertheless, for a mesh with periodic pattern, we are able to ensure that one can find an admissible family of virtual coordinates such that the residue $(\hat{r}_{max})_i$ actually vanish on any inner cell of the mesh.

Theorem 3.4. Let $(C; F)$ be a periodic mesh over $\mathbb{R}^d$ as in Definition 1.6. Let $(a_{i;j}(b_c))_{i;j\in V} = P_F b$ be the discretization of the constant velocity field $b(x) \equiv b_c$. Then for any constant velocity field $b_c \in \mathbb{R}^d$, there exist virtual coordinates $(\hat{x}_i(b_c))_{i\in V} \in (\mathbb{R}^d)^V$ solving the linear system

$$
\sum_{i'\in V} \big(\hat{x}_{i'}(b_c) - \hat{x}_i(b_c)\big) a_{i';i}(b_c) = b_{ci}; \qquad \forall i \in V_0; \tag{3.7}
$$

and satisfying the following properties:

(1) The virtual coordinates are homogeneous in the sense that

$$
\hat{x}_i(\lambda b_c) = \hat{x}_i(b_c); \qquad \forall b_c \in \mathbb{R}^d;\ \lambda > 0;\ i \in V:
$$

(2) The virtual coordinates are uniformly bounded by

$$
\sup_{|b_c|=1;\, i\in V} |\hat{x}_i(b_c) - x_i| < C(|V_0|)\delta x;
$$

where δx is the discretization size and $C(|V_0|)$ depends only on the number of cell functions in a period.

(3) The virtual coordinates are periodic in the sense that

$$
\hat{x}_{[m](i)}(b_c) = \hat{x}_i(b_c) + \sum_{k=1}^{d} m_k L_k; \qquad \forall b_c \in \mathbb{R}^d;\ i; [m](i) \in V;
$$

where $L_1; \ldots; L_n \in \mathbb{R}^d$ are given as in Definition 1.6.

Notice that for constant velocity field, one has $a_{[m](i'),[m](i)} = a_{i',i}$ when $[m](i'), [m](i)$ are well-defined. Hence (3.7) can be naturally extend to all $i \in V$. That is, the maximal residue function $((\hat r_{max})_i)_{i\in V}$ in Definition 3.2 vanishes at all $i \in V$.

3.4. Proof of Theorem 2.9. We are now ready to complete the proof of Theorem 2.9.

Proof of Theorem 2.9. We apply Theorem 3.4, Theorem 3.1 and Theorem 3.3 successively:

Apply Theorem 3.4: Let (C, F) be one mesh in the sequence of meshes we consider in Theorem 2.9 and let δx be the discretization size. Since the meshes have periodic patterns over, one can choose $\hat x_i(b_c)$ as the periodic solution in Theorem 3.4 when $i \in V$, and choose $\hat x_i(b_c) = x_i$ as barycenter when $i \in \bar V \setminus V$. Then the maximal residue function $((\hat r_{max})_i)_{i\in V}$ in Definition 3.2 vanishes at all $i \in V$. Notice that

$$\sup_{|b_c|=1,\, i\in V} |\hat x_i(b_c) - x_i| < C(|V_0|)\,\delta x$$

where $C(|V_0|)$ is the constant in Theorem 3.4.

By our definition of discretization size and barycenter one has

$$\sup_{(i,i')\in E} |x_i - x_{i'}| \le 2\delta x.$$

Moreover, if $i \in \bar V \setminus V$, by definition, one has $(P_C b)_i = 0$ and $(P_F b)_{i',i} = 0$ for all $i' \in V$. Hence, one has $b_i \equiv 0$, $a_{i',i} \equiv 0$ in (3.5) and the residue $((\hat r_{max})_i)_{i\in V}$ in Definition 3.2 actually vanishes on all $i \in \bar V \setminus V$.

In summary, the residue can be non-zero only at $i \in \bar V \setminus V$. Thus one has that

$$
\begin{aligned}
|\Pi|^{-1/r}\|(\hat r_{max})_i\|_{L^r(V)} &\le \left(|\Pi|^{-1} \sum_{i\in \bar V\setminus V} \sup_{|b_c|=1} \left| \sum_{i'\in V} (\hat x_{i'}(b_c) - \hat x_i(b_c))\, a_{i',i}(b_c) - b_{c,i} \right|^r \right)^{1/r} \\
&\le \left(|\Pi|^{-1} \sum_{i\in \bar V\setminus V} C(\delta x)^d \right)^{1/r} \\
&\le C\, |\Pi|^{-1} |\partial \Pi|\, \delta x^{1/r}.
\end{aligned}
$$

Let $r = 1/(1/p - 1/q)$ as required in Theorem 3.3. One can see that such choice of $\hat x_i(b_c)$ forms an admissible family of virtual coordinates in Definition 3.2, with relative drift M, absolute drift M and residue bound M in $L^{1/(1/p-1/q)}$ given by

$$M = M = (2C(|V_0|) + 1)\delta x; \qquad M = C\, |\Pi|^{-1} |\partial \Pi|\, \delta x^{1/p - 1/q}.$$

The constant C here depends on, the exponents p, q, the constant in structural assumption (1.14), and, in particular, the number of cell functions in a period. This is why the constant bounding pattern size is part of the requirement of Theorem 2.9.

Moreover, to fulll the requirement in Theorem 3.1, the discretization size x needs to be chosen small so that $M \ \min\{h_0; 1=32\}$, which is why the inequality only holds asymptotically.

Apply Theorem 3.1 and Theorem 3.3: We have proved the existence of an admissible family of virtual coordinates for constant velocity elds. Recall that Theorem 3.3 shows how to reduce the residue term for non-constant velocity elds provided such family exists. Hence, we choose the coecients $a(t) = (a_{i;j})_{i;j\in V}$, $b(t)$ and the solution $u(t) = (u_i(t))_{i\in V}$ as in Theorem 2.9. Then the rest condi-tions required in Theorem 3.3, namely the boundedness of their Lebesgue and Sobolev norms and the boundedness of

$$(a_{i;j})_{i;j\in V} \quad P_F\, b \in L^q([0;T]F);$$

are directly guaranteed by the assumptions in Theorem 2.9. According to Theorem 3.3, one can choose virtual coordinates for specic coecients, such that the propagation of regularity (3.3) in Theorem 3.1, is bounded by (3.6).

For clarity, we recall the full result,

$$\|u(t)\|_{h_0;1;} \ (L_0 + L_1 + L_2 + L_3);$$

where the precise formulations of these terms are given by

$$\begin{aligned}
L_0 &= \|u(0)\|_{h_0;1;};\\
L_1 &= C\int_0^t \Big(\|\operatorname{div} b(s)\|_{L^1(C)}\|u(s)\|_{h_0;1;} + \|b(s)\|_{W^{1;q}}\|u(s)\|_{L^p(C)}\\
&\quad + \|D(s)\|_{L^1(C)}\|u(s)\|_{h_0;1;} + \|u(s)\|_{L^p(C)}\|D(s)\|_{h_0;p;p(1=p)}\Big)\, ds;
\end{aligned}$$

where the constant C depends on , the exponents $p;q$ and the constant in structural assumptions (1.14). Also,

$$= \exp\ C(1=h_0)(M^s M=x)^{1=(1+s)};$$

$$L_2 = C \int^t \quad j\log h_0 j \quad M^2=h_0^2 x k(a_{i;j}(s))_{i;j2V} k_{L^q(F)} ku(s)k_{L^q(C)}$$
$$+ \quad j\log h_0 j \quad M=h_0^2 \ k\tilde{b}(s)k_{L^q} ku(s)k_{L^q(C)}$$
$$+ \ (j\log h_0 j \quad x=h_0^2)k\tilde{b}(s)k_{W^{1;q}} ku(s)k_{L^q(C)} \qquad ds$$

$$+ \ C(j\log h_0 j^1 \ ) \ ku(t)k_{L^1(C)} \quad ku(0)k_{L^1(C)} \ ;$$

$$L_3 = C(j\log h_0 j \ =h_0) \ (M=x)(a_{i;j})_{i;j2V} \quad P_F \tilde{b}_{L^q([0;T]F)}$$
$$+ \ (M^s M=x)^{1=(1+s)} k\tilde{b}k_{L^p_t(W^{s;p}_x)}$$
$$+ \quad M(x)^{\frac{1}{1+(1=p \quad 1=q)}} \quad k(a_{i;j}(t))_{i;j2V} k_{L^q([0;T]F)} + k\tilde{b}k_{L^q_t(W^{1;q}_x)}$$
$$+ \ Mk\tilde{b}k_{L^q_t(L^q_x)} \quad kuk_{L^1_t L^1_x([0;t]C)}:$$

where the constant C depends on T; , the exponents $p;q;s$, the constant in structural assumption (1.14) and the constant bounding pattern size.

Conclude Theorem 2.9 We can now make all the constants explicit in the propagation of regularity with,

$$(M^s M=x)^{1=(1+s)} \ . \ (x)^{1=(1+s)}; \qquad M^2=x;\ M \ . \ x;$$
$$M(x)^{\frac{1}{1+(1=p \quad 1=q)}} \ . \ (x)^{\frac{1=p \quad 1=q}{1+(1=p \quad 1=q)}}; \qquad M \ . \ (x)^{1=p \quad 1=q} \ . \ (x)^{\frac{1=p \quad 1=q}{1+(1=p \quad 1=q)}}:$$

Thus ; L_2; L_3 can be expressed as

$$= \exp\ C(1=h_0)(x)^{s=(1+s)};$$
$$L_2 = C\int_0^t \quad j\log h_0 j \ =h_0^2(x) \quad ka(s)k_{L^q(F)} + k\tilde{b}(s)k_{W^{1;q}} \quad ku(s)k_{L^q(C)}\ ds$$

$$+ \ (j\log h_0 j^1 \ ) \ ku(t)k_{L^1(C)} \quad ku(0)k_{L^1(C)} \ ;$$

$$L_3 = C(j\log h_0 j \ =h_0) \ a \quad P_F \tilde{b}_{L^q([0;T]F)} + (x)^{s=(1+s)} k\tilde{b}k_{L^p_t(W^{s;p}_x)}$$
$$+ \ (x)^{\frac{1=p \quad 1=q}{1+(1=p \quad 1=q)}} \ ka(t)k_{L^q([0;T]F)} + k\tilde{b}k_{L^q_t(W^{1;q}_x)} \qquad kuk_{L^1_t L^p_x([0;t]C)}:$$

One can check that this exactly corresponds to (2.9) of Theorem 2.9.Applying the argument to the entire sequence of meshes $(C^{(n)}; F^{(n)})$ concludes the proof of Theorem 2.9.

4. Proof of Theorem 3.1. In this section we complete the proof of Theorem 3.1. For simplicity we omit the variable t in the derivation when there is no ambiguity. Throughout the calculation, C denotes a generic constant that only depends on , the exponents $p;q$ and the constant in the structural assumptions (1.14).

4.1. The Kruzkov's doubling of variables for the semi-discrete scheme. Notice that any solution u of the rst-order scheme (1.12) satises the following identity:

$$\frac{d}{dt}u_i = \frac{1}{\ i}\sum_{j\in V}\left(a_{i;j}u_j \quad a_{j;i}u_i\right) \quad 1_{V\setminus V}(i)\frac{1}{\ i}\sum_{j\in V}a_{i;j}u_j: \tag{4.1}$$

When $i \in V$ the above equality is the upwind scheme. When $i \in (V\setminus V)$, one has $u_i \ 0$ and the above equality reduce to $0 = 0$. In this section let us use the notation

$$R_i = \quad 1_{V\setminus V}(i)\frac{1}{\ i}\sum_{j\in V}a_{i;j}u_j:$$

The term R_i measures the possible leaking of mass at boundary. It is easy to verify that $R_i \ 0$ and

$$\frac{d}{dt}\sum_{i\in V}u_i\ i = \sum_{i\in V}R_i\ i: \tag{4.2}$$

The next proposition explains how to bound the time derivative of our semi-norms.

Proposition 4.1. For any solution u of the rst-order scheme (1.12) and any non-negative discrete kernel $\{K_{i;j}\}_{i;j\in V}$, the following inequality holds in the sense of distribution:

$$\begin{aligned}\frac{d}{dt}\sum_{i;j\in V}K_{i;j}|u_i \quad u_j|\ i\ j \quad & 2\sum_{i;j\in V}\sum_{i^0\in V}(K_{i^0;j} \quad K_{i;j})a_{i^0;i}|u_i \quad u_j|\ j\\ &+ (\quad 2)\sum_{i;j\in V}K_{i;j}\,\mathrm{sgn}(u_i \quad u_j)(D_iu_j)\ i\ j\\ &+ 2\sum_{i;j\in V}K_{i;j}\,\mathrm{sgn}(u_i \quad u_j)R_i\ i\ j\\ &= A_K + D_K + R_K:\end{aligned}$$

Proof of Proposition 4.1. The following equality holds in distributional sense:

$$\begin{aligned}&\frac{d}{dt}\sum_{i;j\in V}K_{i;j}|u_i \quad u_j|\ i\ j\\ &= \sum_{i;j\in V}K_{i;j}\,\mathrm{sgn}(u_i \quad u_j)\left(\frac{1}{\ i}\sum_{i^0\in V}\left(a_{i;i^0}u_{i^0} \quad a_{i^0;i}u_i\right) + R_i \quad \frac{1}{\ j}\sum_{j^0\in V}\left(a_{j;j^0}u_{j^0} \quad a_{j^0;j}u_j\right) \quad R_j\right)\ i\ j\\ &= 2\sum_{i;j\in V}K_{i;j}\,\mathrm{sgn}(u_i \quad u_j)\left(\frac{1}{\ i}\sum_{i^0\in V}\left(a_{i;i^0}u_{i^0} \quad a_{i^0;i}u_i\right) + R_i\right)\ i\ j:\end{aligned}$$

Proving the rst equality is nothing but the chain rule applied to the semi-discrete scheme. The second equality follows from symmetry, by switching the indexes i and j, i^0 and j^0.

The next step is to check that this can be further decomposed into our sum $A_K + D_K + R_K$ plus a non-positive term. Indeed,

$$
\begin{aligned}
&\sum_{i,j\in V} K_{i,j}\,\mathrm{sgn}(u_i-u_j)\left(\frac{1}{\pi_i}\sum_{i'\in V} a_{i,i'}u_{i'} - a_{i',i}u_i\right)\pi_i\pi_j\\
&=\sum_{i,j\in V}\sum_{i'\in V}\Bigg\{K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_{i'} - K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_i\Bigg\}\pi_j\\
&=\sum_{i,j\in V}\sum_{i'\in V}\Bigg\{K_{i',j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_i - K_{i',j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_j\\
&\qquad\qquad - K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_i + K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_j\Bigg\}\pi_j\\
&\quad+\sum_{i,j\in V}\sum_{i'\in V}\Bigg\{-K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i',i}u_j + K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_j\Bigg\}\pi_j\\
&\quad+\sum_{i,j\in V}\sum_{i'\in V}\Bigg\{-K_{i,j}\,\mathrm{sgn}(u_{i'}-u_j)a_{i,i'}u_{i'} + K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_{i'}\\
&\qquad\qquad + K_{i,j}\,\mathrm{sgn}(u_{i'}-u_j)a_{i,i'}u_j - K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_j\Bigg\}\pi_j\\
&= A_K^c + D_K^c + N_K.
\end{aligned}
$$

By our assumption, $K_{i,j}\ge 0$, $a_{i,j}\ge 0$ for all $i,j\in V$. It is easy to verify that the third term

$$
\begin{aligned}
N_K &= \sum_{i,j\in V}\sum_{i'\in V}\Bigg\{-K_{i,j}\,\mathrm{sgn}(u_{i'}-u_j)a_{i,i'}u_{i'} + K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_{i'}\\
&\qquad\qquad + K_{i,j}\,\mathrm{sgn}(u_{i'}-u_j)a_{i,i'}u_j - K_{i,j}\,\mathrm{sgn}(u_i-u_j)a_{i,i'}u_j\Bigg\}\pi_j\\
&= \sum_{i,j\in V}\sum_{i'\in V} K_{i,j}\big(\mathrm{sgn}(u_{i'}-u_j) - \mathrm{sgn}(u_i-u_j)\big)a_{i,i'}(u_j-u_{i'})\pi_j
\end{aligned}
$$

is always non-positive, and it does not vanish only at edges that $\mathrm{sgn}(u_{i'} - u_j) = -\mathrm{sgn}(u_i - u_j)$. In addition, one can reformulate A^0_K and D^0_K as

$$\begin{aligned}
A^0_K &= \sum_{i,j\in V}\sum_{i'\in V}\Big(K_{i',j}\,\mathrm{sgn}(u_i - u_j) a_{i',i} u_i - K_{i',j}\,\mathrm{sgn}(u_i - u_j) a_{i',i} u_j \\
&\qquad - K_{i,j}\,\mathrm{sgn}(u_i - u_j) a_{i',i} u_i + K_{i,j}\,\mathrm{sgn}(u_i - u_j) a_{i',i} u_j \Big)\mu_j \\
&= \sum_{i,j\in V}\sum_{i'\in V} (K_{i',j} - K_{i,j}) a_{i',i} |u_i - u_j|\mu_j; \\
D^0_K &= \sum_{i,j\in V}\sum_{i'\in V}\Big(-K_{i,j}\,\mathrm{sgn}(u_i - u_j) a_{i',i} u_j + K_{i,j}\,\mathrm{sgn}(u_i - u_j) a_{i,i'} u_j \Big)\mu_j \\
&= \sum_{i,j\in V} K_{i,j}\,\mathrm{sgn}(u_i - u_j)\Big(\sum_{i'\in V} a_{i',i} - a_{i,i'}\Big) u_j \mu_i\mu_j.
\end{aligned}$$

It is straightforward to see that $A^0_K + D^0_K = A_K/2 + D_K/2$. Hence

$$\begin{aligned}
&\sum_{i,j\in V} K_{i,j}\,\mathrm{sgn}(u_i - u_j)\Big(\frac{1}{\mu_i}\sum_{i'\in V} a_{i,i'} u_{i'} - a_{i',i} u_i + R_i\Big)\mu_i\mu_j \\
&= A^C_K + D^C_K + N_K + R_K/2 \\
&\le A_K/2 + D_K/2 + R_K/2.
\end{aligned}$$

Multiplying both sides by 2, one obtains the inequality in the proposition.

From now on we fix the kernel $K_{i,j}$ in the above proposition as $K^{\tilde{h}}_{i,j}$ in Definition 2.3 for $0 < h < 1/2$ and $\tilde{x} = (\tilde{x}_i)_{i\in V} \in (\mathbb{R}^d)^V$. Moreover, assume that $h \ge \max\{\delta x, \sup_{i\in V} |\tilde{x}_i - x_i|\}$. Then the term R_K can be bounded by

$$\begin{aligned}
|R_K| &= 2\Big|\sum_{i,j\in V} K^{\tilde{h}}_{i,j}\,\mathrm{sgn}(u_i - u_j) R_i \mu_i\mu_j\Big| \le 2\sum_{i\in V}\sum_{j\in V} K^{\tilde{h}}_{i,j} |R_i|\mu_i\mu_j \\
&\le C|\log h|\,\|(R_i)_{i\in V}\|_{L^1(C)}.
\end{aligned} \tag{4.3}$$

Moreover, the term D_K can then be estimated through

$$\begin{aligned}
|D_K| &= 2\Big|\sum_{i,j\in V} K^{\tilde{h}}_{i,j}\,\mathrm{sgn}(u_i - u_j)(D_i u_j)\mu_i\mu_j\Big| \\
&= \Big|\sum_{i,j\in V} K^{\tilde{h}}_{i,j}\,\mathrm{sgn}(u_i - u_j)(D_i u_j - D_j u_i)\mu_i\mu_j\Big| \\
&= \Big|\sum_{i,j\in V} K^{\tilde{h}}_{i,j}\frac{D_i + D_j}{2}|u_i - u_j|\mu_i\mu_j \\
&\quad + \sum_{i,j\in V} K^{\tilde{h}}_{i,j}(D_i - D_j)\frac{u_i + u_j}{2}\,\mathrm{sgn}(u_i - u_j)\mu_i\mu_j\Big| \\
&\le C|\log h|\Big(\|D\|_{L^1(C)}\|u\|_{h_0,1,\tilde{x}} + \|u\|_{L^p(C)}\|D\|_{h_0,p,p(-1/p),\tilde{x}}\Big).
\end{aligned} \tag{4.4}$$

The last inequality is a consequence of the following two estimations: Using the bound on the divergence, one has

$$\begin{aligned}\sum_{i;j\in V}\tilde K^h_{i;j}\frac{D_i+D_j}{2}|u_i-u_j|\,\mu_{ij} &\le \|D\|_{L^\infty(V)}\sum_{i;j\in V}\tilde K^h_{i;j}|u_i-u_j|\,\mu_{ij}\\ &\le |\log h|\,\|D\|_{L^\infty(C)}\|u\|_{;1;\tilde x}.\end{aligned}$$

Also, by Hölder estimate

$$\begin{aligned}\sum_{i;j\in V}K_{\tilde h}(D_i-D_j)\frac{u_i+u_j}{2}\,\mathrm{sgn}(u_i-u_j)\,\mu_{ij} &\le \Big(\sum_{i;j\in V}\tilde K^h_{i;j}|D_i-D_j|^p\mu_{ij}\Big)^{1/p}\Big(\sum_{i;j\in V}\tilde K^h_{i;j}\Big|\frac{u_i+u_j}{2}\Big|^p\mu_{ij}\Big)^{1/p}\\ &\le C|\log h|\,\|u\|_{L^p(C)}\|D\|_{p;p(1-1/p);\tilde x}.\end{aligned}$$

The above Hölder estimate is for $1<p<\infty$ but can be extended to $p=\infty$ in the obvious way.

4.2. Bounding the discrete commutator term. We now investigate the discrete commutator term A_K when $K_{i;j}$ is chosen as $\tilde K^h_{i;j}$ in Definition 2.3 for $\tilde x=(\tilde x_i)_{i\in V}\in(\mathbb R^d)^V$ and $\max\{\delta x;\sup_{i\in V}|\tilde x_i-x_i|\}\le h<1/2$. Recall that

$$A_K/2=\sum_{i;j\in V}\sum_{i'\in V}(\tilde K^h_{i';j}-\tilde K^h_{i;j})\,a_{i';i}|u_i-u_j|. \tag{4.5}$$

We begin by a short lemma about the scaling of the continuous kernel K^h.

Lemma 4.2. Take $x,y,s\in\mathbb R^d$ such that $0<h<1/2$ and $|s|<h$. Then

$$\big|K_h(x-y)-K_h(x-y+s)+\nabla K_h(x-y)\cdot s\big|\le\frac{C|s|^2}{(|x-y|+h)^{d+2}}. \tag{4.6}$$

Also,

$$\big|\nabla K_h(x-y+s)-\nabla K_h(x-y)\big|\le\frac{C|s|}{(|x-y|+h)^{d+2}}. \tag{4.7}$$

We are going to use this lemma to reduce a few terms to simpler forms, with a tolerable error. In particular, the following lemma mimics the continuous commutator estimate in [4], provided that one can find suitable auxiliary functions $(\tilde x_i)_{i\in V}$ and $(b_i)_{i\in V}$.

Lemma 4.3. Consider the semi-discrete scheme (1.12) on a mesh $(C;F)$ over $\mathbb R^d$ as in Definition 1.5, having discretization size δx and satisfying the structural assumptions (1.14). Let $(a_{i;j})_{(i;j)\in E}$ be the coefficients of the scheme and let $D=(D_i)_{i\in V}$ be the discrete divergence defined as in (2.3). Let $b(x)$ be a continuous velocity field on $\mathbb R^d$ and denote $(b_i)_{i\in V}=\tilde P_V b$. Choose virtual coordinates $(\tilde x_i)_{i\in V}$ on the mesh satisfying

$$\begin{aligned}|\tilde x_i-\tilde x_{i'}|&<M, &&\forall(i;i')\in E;\\ |\tilde x_i-x_i|&<M, &&\forall i\in V.\end{aligned} \tag{4.8}$$

Let $\tilde K^h_{i,j}$ be as in Definition 2.3 corresponding to $(\tilde x_i)_{i\in V}$ and let $(r_i(t))_{i\in V}$ be the residue function given by

$$\sum_{i'\in\partial\{i\}} (\tilde x_{i'} - \tilde x_i)\, a_{i',i} = \tilde b_i |i| + r_i |i|, \quad \forall i \in V. \tag{4.9}$$

Then the discrete commutator term A_K given through (4.5) can be bounded by

$$\begin{aligned} |A_K| \le{} & C|\log h| \, \|\operatorname{div} \tilde b\|_{L^1} \|u\|_{h_0,1,\infty} + \|\tilde b\|_{W^{1,q}} \|u\|_{L^p(C)} \\ & + C\, \bar M_2/h^2\, x\|(a_{i,j})_{i,j\in V}\|_{L^q(F)} \|u\|_{L^{q^*}} + C\, M/h^2 \|(\tilde b_i)_{i\in V} - \tilde b\|_{L^q(C)} \|u\|_{L^{q^*}} \\ & + C(1/h)\|r\|_{L^p}\|u\|_{L^{p^*}}; \end{aligned} \tag{4.10}$$

provided that $1 \le p < q \le \infty$, $\max\{1 - 1/q, 1/2\}$, and $M < \bar M < h_0 < h$.

Note that conditions (4.8) and (4.9) exactly correspond to (3.1) and (3.2) in Theorem 3.1 once the time-dependency is removed.

Proof. We have that

$$\begin{aligned} A_K/2 &= \sum_{i,j\in V} \sum_{i'\in V} (\tilde K^h_{i',j} - \tilde K^h_{i,j})\, a_{i',i} |u_i - u_j| \\ &= \sum_{i,j\in V} \sum_{i'\in V} \nabla K^h(\tilde x_i - \tilde x_j)\cdot(\tilde x_{i'} - \tilde x_i)\, a_{i',i} |u_i - u_j| \\ &\quad + \sum_{i,j\in V} \sum_{i'\in V} \left((\tilde K^h_{i',j} - \tilde K^h_{i,j}) - \nabla K^h(\tilde x_i - \tilde x_j)\cdot(\tilde x_{i'} - \tilde x_i)\right) a_{i',i} |u_i - u_j| \\ &= A_K^{(1)} + A_K^{(2)} \end{aligned}$$

By Lemma 4.2 and assumption (4.8), one has that

$$(\tilde K^h_{i',j} - \tilde K^h_{i,j}) - \nabla K^h(\tilde x_i - \tilde x_j)\cdot(\tilde x_{i'} - \tilde x_i) \le \frac{C M^2}{(|\tilde x_i - \tilde x_j| + h)^{d+2}} \le \frac{C M^2}{(|x_i - x_j| + h)^{d+2}}.$$

Therefore one can bound $A_K^{(2)}$ by

$$
\begin{aligned}
A_K^{(2)} &\le \sum_{i,j\in V}\sum_{i'\in V}\frac{CM^2}{(|x_i-x_j|+h)^{d+2}}\,a_{i',i}|u_i-u_j|\\
&\le \sum_{i,j\in V}\sum_{i'\in V}\frac{CM^2}{(|x_i-x_j|+h)^{d+2}}\,a_{i',i}|u_i| + \sum_{i,j\in V}\sum_{i'\in V}\frac{CM^2}{(|x_i-x_j|+h)^{d+2}}\,a_{i',i}|u_j|\\
&\le CM^2(\delta x)^{-1}\sum_{i,i':(i,i')\in E} a_{i',i}|u_i|\,\delta x\sum_{j\in V}\frac{1}{(|x_i-x_j|+h)^{d+2}}\\
&\quad + CM^2(\delta x)^{-1}\left(\sum_{i,i':(i,i')\in E} a_{i',i}^{\,q}(\delta x)^{d-q(d-1)}\right)^{1/q}\\
&\qquad\left(\sum_{i,i':(i,i')\in E}\left(\sum_{j\in V}\frac{1}{(|x_i-x_j|+h)^{d+2}}|u_j|\right)^{q}(\delta x)^d\right)^{1/q}\\
&\le C\,M^2/h^2\,\delta x\,\|(a_{i,j})_{(i,j)\in E}\|_{L^q(F)}\|u\|_{L^q(C)}.
\end{aligned}
$$

For $A_K^{(1)}$, one can apply the identity (4.9) to obtain

$$
\begin{aligned}
A_K^{(1)} &= \sum_{i,j\in V}\sum_{i'\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot(\tilde x_{i'}-\tilde x_i)a_{i',i}|u_i-u_j|\\
&= \sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot\tilde b_i\,|u_i-u_j|\\
&\quad+\sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot r_i\,|u_i-u_j|\\
&= A_K^{(1,1)}+A_K^{(1,2)}.
\end{aligned}
$$

Repeating the argument on $A_K^{(2)}$, we bound the residue term $A_K^{(1,2)}$ by

$$
\begin{aligned}
A_K^{(1,2)} &= \sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot r_i|u_i-u_j|\\
&\le \sum_{i,j\in V}\frac{C}{(|x-y|+h)^{d+1}}|r_i||u_i-u_j|\\
&\le C(1/h)\|r\|_{L^p(C)}\|u\|_{L^p(C)}.
\end{aligned}
$$

Finally, symmetrize the expression of $A_K^{(1,1)}$ to obtain

$$
\begin{aligned}
A_K^{(1,1)} &= \sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot\tilde b_i|u_i-u_j|\\
&= \frac{1}{2}\sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot(\tilde b_i-\tilde b_j)|u_i-u_j|.
\end{aligned}
$$

Choose measurable sets $(V_i)_{i\in V}\subset\mathbb{R}^d$ and a piecewise constant extension $u^V=\sum_{i\in V}u_i\mathbb{1}_{V_i}$ by Lemma 2.6. Those satisfy

$$|V_i| = \omega_i = \int_{\mathbb{R}^d}\psi_i;\qquad \sup_{x\in V_i}|x-x_i| < 2\Delta x;\qquad \forall i\in V;\qquad V_i\cap V_j=\emptyset;\qquad \forall i,j\in V:$$

and

$$\operatorname{supp} u^V \subset \bigcup_{i\in V} V_i \subset \Omega + B(0;1) \subset \Omega + B(0;3):$$

This leads us to introduce the continuous commutator term

$$A_K^{(1;1;1)} = \frac{1}{2}\int_{\mathbb{R}^{2d}}\nabla K^h(x-y)\cdot(\tilde b(x)-\tilde b(y))|u^V(x)-u^V(y)|\,dxdy:$$

Notice that $\operatorname{supp} u^V\subset\Omega+B(0;1)$ and $\operatorname{supp}\nabla K^h\in B(0;2)$. Then for $x\notin\Omega+B(0;3)$, either $y\notin\Omega+B(0;1)$, making $|u^V(x)-u^V(y)|=0$, or $y\in\Omega+B(0;1)$, making $\nabla K^h(x-y)=0$. The same argument applies to y. As a consequence, the integral formulating $A_K^{(1;1;1)}$ can be taken over any subset of $\mathbb{R}^{2d}$ including $\left(\Omega+B(0;3)\right)^2$. In particular,

$$A_K^{(1;1;1)} = \frac{1}{2}\int_{(\bigcup_{i\in V}V_i)^2}\nabla K^h(x-y)\cdot(\tilde b(x)-\tilde b(y))|u^V(x)-u^V(y)|\,dxdy:$$

Combine the above discussion with Lemma 4.2 and assumption (4.8), one has

$$\begin{aligned}
2A_K^{(1;1;1)}-A_K^{(1;1)} &= \int_{(\bigcup_{i\in V}V_i)^2}\nabla K^h(x-y)\cdot(\tilde b(x)-\tilde b(y))|u^V(x)-u^V(y)|\,dxdy\\
&\quad-\sum_{i,j\in V}\nabla K^h(\tilde x_i-\tilde x_j)\cdot(\tilde b_i-\tilde b_j)|u_i-u_j|\omega_i\omega_j\\
&\le \sum_{i,j\in V}\int_{V_iV_j}\left(\nabla K^h(x-y)-\nabla K^h(\tilde x_i-\tilde x_j)\right)\cdot(\tilde b(x)-\tilde b(y))|u_i-u_j|\,dxdy\\
&\quad+\sum_{i,j\in V}\int_{V_iV_j}\nabla K^h(\tilde x_i-\tilde x_j)\cdot\left((\tilde b(x)-\tilde b(y))-(\tilde b_i-\tilde b_j)\right)|u_i-u_j|\,dxdy\\
&\le \sum_{i,j\in V}\int_{V_iV_j}\frac{CM}{(|x-y|+h)^{d+2}}|\tilde b(x)-\tilde b(y)||u_i-u_j|\,dxdy\\
&\quad+\sum_{i,j\in V}\int_{V_iV_j}\frac{C}{(|x-y|+h)^{d+1}}\left|(\tilde b(x)-\tilde b(y))-(\tilde b_i-\tilde b_j)\right||u_i-u_j|\,dxdy\\
&\le C\Delta x M/h^2\,\|\tilde b\|_{L^q}\|u\|_{L^q(C)}+C(\Delta x/h)\|\tilde b\|_{W^{1,q}}\|u\|_{L^q(C)}:
\end{aligned}$$

Finally, the continuous commutator term $A_K^{(1;1;1)}$ can be estimated by Lemma 16 in [4]. The paper [4] also considered some non-linearity within the advection equation, which makes the formulations more complicated than what we need here. For the sake of completeness, we thus restate a simplied version of Lemma 16 in [4] with the notations of our paper.

Lemma 4.4. (Lemma 16 in [4], reformulated) *Assume that for some $1 < q, r < \infty$, we have $u \in L^{q,\infty}$ and b belonging to Besov space $B^1_{q,r}$. Then provided $1 - 1/r$,*

$$\left|\int_{\mathbb{R}^{2d}} \nabla K_h(x-y)\,(b(x)-b(y))\,dx\,dy\right|$$

$$\le C|\log h|\left(\|\operatorname{div} b\|_{L^1}\,|\log h| \int_{\mathbb{R}^{2d}} K_h(x-y)\,|u(x)-u(y)|\,dx\,dy + \|\nabla b\|_{B_{q,r}}\|u\|_{L^{q,\infty}}\right).$$

We are going to apply this lemma by taking $b = \tilde{b}$, $u = u^V$ and $r = \max\{q, 2\}$.

First we recall the classical bound $\|\nabla b\|_{B^0_{q,q\vee 2}} \le C\|\nabla b\|_{L^q} \le C\|b\|_{W^{1,q}}$. We also remark that $\tilde{b}$ is defined on $\tilde{\Omega}$, but one can nevertheless extend it to $\mathbb{R}^d$ with $\|\tilde{b}\|_{W^{1,q}(\mathbb{R}^d)} \le C\|\tilde{b}\|_{W^{1,q}(\tilde{\Omega})}$. We also recall that, since we consider a bounded domain Ω and assume $1 \le p < q \le \infty$, then $\|u\|_{L^{q,\infty}} \le C\|u\|_{L^p}$.

Therefore, for $1 - 1/r = \max\{1 - 1/q, 1/2\}$, one has

$$2A_K^{(1,1,1)} = \int_{\mathbb{R}^{2d}} \nabla K_h(x-y)\,(\tilde{b}(x) - \tilde{b}(y))\,|u^V(x) - u^V(y)|\,dx\,dy$$

$$\le C|\log h|\left(\|\operatorname{div} \tilde{b}\|_{L^1(\tilde{\Omega})}\,|\log h| \int_{\mathbb{R}^{2d}} K_h(x-y)\,|u^V(x) - u^V(y)|\,dx\,dy + \|\tilde{b}\|_{W^{1,q}}\|u^V\|_{L^p}\right).$$

The terms involving u^V can be further bounded by the discrete density $(u_i)_{i\in V}$. In particular, applying Lemma 2.5 by choosing $u = v = u^V$, $f_k, g_k : \mathbb{R}^d \to \mathbb{R}^d$, $k = 1, 2$ satisfying $f_1(x) = f_2(x) = x_i$ for all $x \in V_i$, and $g_1(x) = g_2(x) = x$ for all $x \in \mathbb{R}^d$, one has

$$\int_{\mathbb{R}^{2d}} K_h(x-y)\,|u^V(x) - u^V(y)|\,dx\,dy \le C \sum_{i,j} \tilde{K}^h_{i,j}\,|u_i - u_j|\,|i|\,|j| \le |\log h|\,\|u\|_{h_0,1,\infty}.$$

This finally leads to the estimate

$$A_K^{(1,1)} \le C|\log h|\,\|\operatorname{div} \tilde{b}\|_{L^1(\tilde{\Omega})}\|u\|_{h_0,1,\infty} + \|\tilde{b}\|_{W^{1,q}}\|u\|_{L^p(\Omega)} + C\sqrt{M}/h^2\,\|\tilde{b}\|_{L^q}\|\tilde{u}\|_{L^q(\Omega)} + C(x/h)\|\tilde{b}\|_{W^{1,q}}\|\tilde{u}\|_{L^q(\Omega)}.$$

Combine the estimate for $A_K^{(1,1)}$, $\tilde{A}_K^{(1,2)}$ and $\tilde{A}_K^{(2)}$, we conclude (4.10), which finishes the proof of Lemma 4.3.

We are now ready to conclude the proof of Theorem 3.1.

Proof of Theorem 3.1. Let us first consider the case $m = 1$, i.e. the $(x_i)_{i\in V}$ are time-independent instead of just piecewise constant along time.

Then by Definition 2.3 and Proposition 4.1, one has that

$$\|u(t)\|_{h_0,1,\infty} \le \|u(0)\|_{h_0,1,\infty} + \int_0^t \sup_{h_0 \le h \le 1/2} |\log h| \left(\frac{d}{ds}\sum_{i,j\in V} \tilde{K}^h_{i,j}\,|u_i(s) - u_j(s)|\,|i|\,|j|\right) ds$$

$$\le \|u(0)\|_{h_0,1,\infty} + \int_0^t \sup_{h_0 \le h \le 1/2} |\log h| \left(A_K(s) + D_K(s) + R_K(s)\right) ds.$$

Substituting $A_K(s) + D_K(s) + R_K(s)$ by (4.3), (4.4) and (4.10), and rearranging the terms, one deduces that

$$
\begin{aligned}
\|u(t)\|_{h_0;1;\infty} - \|u(0)\|_{h_0;1;\infty}
&\le C\int_0^t \Big(\|\operatorname{div}\tilde{b}(s)\|_{L^\infty(C)} \|u(s)\|_{h_0;1;\infty} + \|\tilde{b}(s)\|_{W^{1,q}} \|u(s)\|_{L^p(C)} \\
&\qquad + \|D(s)\|_{L^\infty(C)} \|u(s)\|_{h_0;1;\infty} + \|u(s)\|_{L^p(C)} \|D(s)\|_{h_0;p;p(1-1/p);\infty} \Big)\, ds \\
&\quad + C\int_0^t \Big(|\log h_0|\, M^2/h_0^2\, \|(a_{i,j}(s))_{i,j\in V}\|_{L^q(F)} \|u(s)\|_{L^q(C)} \\
&\qquad + |\log h_0|\, M/h_0^2\, \|\tilde{b}(s)\|_{L^q} \|u(s)\|_{L^q(C)} \\
&\qquad + (|\log h_0|\, x/h_0^2) \|\tilde{b}(s)\|_{W^{1,q}} \|u(s)\|_{L^q(C)} \\
&\qquad + (|\log h_0|^{-1}) \|(R_i(s))_{i\in V}\|_{L^1(C)} \Big)\, ds \\
&\quad + C\int_0^t (|\log h_0|/h_0) \|(r_i(s))_{i\in V}\|_{L^p(C)} \|u(s)\|_{L^p(C)}\, ds.
\end{aligned}
$$

We can change the $\|u(s)\|_{h_0;1;\infty}$ norms into $\|u(s)\|_{h_0;1}$ norms through Proposition 2.7 with $h_2 = M$,

$$
\begin{aligned}
\|u(t)\|_{h_0;1} &\le (L_0 + L_1 + L_2 + L_3) \\
&= \left(1 + C(M/h_0)^2\right) \Bigg[\|u(0)\|_{h_0;1} \\
&\quad + C\int_0^t \Big(\|\operatorname{div}\tilde{b}(s)\|_{L^\infty(C)} \|u(s)\|_{h_0;1} + \|\tilde{b}(s)\|_{W^{1,q}} \|u(s)\|_{L^p(C)} \\
&\qquad + \|D(s)\|_{L^\infty(C)} \|u(s)\|_{h_0;1} + \|u(s)\|_{L^p(C)} \|D(s)\|_{h_0;p;p(1-1/p)} \Big)\, ds \\
&\quad + C\int_0^t \Big(|\log h_0|\, M^2/h_0^2\, \|(a_{i,j}(s))_{i,j\in V}\|_{L^q(F)} \|u(s)\|_{L^q(C)} \\
&\qquad + |\log h_0|\, M/h_0^2\, \|\tilde{b}(s)\|_{L^q} \|u(s)\|_{L^q(C)} \\
&\qquad + (|\log h_0|\, x/h_0^2) \|\tilde{b}(s)\|_{W^{1,q}} \|u(s)\|_{L^q(C)} \\
&\qquad + (|\log h_0|^{-1}) \|(R_i(s))_{i\in V}\|_{L^1(C)} \Big)\, ds \\
&\quad + C\int_0^t (|\log h_0|/h_0) \|(r_i(s))_{i\in V}\|_{L^p(C)} \|u(s)\|_{L^p(C)}\, ds \Bigg].
\end{aligned}
$$

Here we rewrite the last term of L_2 by

$$
\begin{aligned}
\int_0^t \|(R_i(s))_{i\in V}\|_{L^1(C)}\, ds &= \int_0^t \sum_{i\in V} R_i(s)|i|\, ds = -\int_0^t \frac{d}{ds} \sum u_i(s)|i|\, ds \\
&= \|u(0)\|_{L^1(C)} - \|u(t)\|_{L^1(C)}.
\end{aligned}
$$

where we use $R_i \geq 0$ and identity (4.2). The coefficient $\left(1 + C(M/h_0)\right)$ is multiplied twice because Proposition 2.7 is actually applied to the left and right hand side separately. Since $k(t) = \min\{k : t < t_k\} = 1$ (as we assume $m = 1$) we have $k(t) + 1 = 2$, so all coefficients matches to (3.3), which finishes the proof for the case $m = 1$.

When $m > 1$, within each interval $[t_{k-1}, t_k]$ we still have constant $(\phi_i^{(k)})_{i\in V}$ constant, and the semi-norm $\|u(t)\|_{h_0,1,\phi^{(k)}}$ propagates exactly as above. However, at every end-point t_k one need to shift from $(\phi_i^{(k)})_{i\in V}$ to $(\phi_i^{(k+1)})_{i\in V}$, yielding an extra $C(M/h_0)$ factor.

Define

$$
\begin{aligned}
L_1(\tau;t) = C \int_\tau^t \Big(& \|\operatorname{div} b(s)\|_{L^1(C)} \|u(s)\|_{h_0,1} + \|b(s)\|_{W^{1,q}} \|u(s)\|_{L^p(C)} \\
& + \|D(s)\|_{L^1(C)} \|u(s)\|_{h_0,1} + \|u(s)\|_{L^p(C)} \|D(s)\|_{h_0,p,p(-1/p)} \Big)\, ds \\
L_2(\tau;t) = C \int_\tau^t \Big(& |\log h_0| \, M^2/h_0^2 x \, \|(a_{i,j}(s))_{i,j\in V}\|_{L^q(F)} \|u(s)\|_{L^q(C)} \\
& + |\log h_0| \, M/h_0^2 \, \|b(s)\|_{L^q} \|u(s)\|_{L^q(C)} + \\
& (|\log h_0| \, x/h_0^2) \|b(s)\|_{W^{1,q}} \|u(s)\|_{L^q(C)} \\
& + (|\log h_0|^{1}) \|(R_i(s))_{i\in V}\|_{L^1(C)} \Big)\, ds \\
L_3(\tau;t) = C \int_\tau^t & (|\log h_0| / h_0) \|(r_i(s))_{i\in V}\|_{L^p(C)} \|u(s)\|_{L^p(C)} \, ds.
\end{aligned}
$$

We now argue by induction that

$$
\|u(t)\|_{h_0,1,\phi^{(k)}} \leq \left(1 + C(M/h_0)\right)^{k(t)-1} \left(\|u(0)\|_{h_0,1,\phi^{(1)}} + L_1(0;t) + L_2(0;t) + L_3(0;t)\right)
$$

by induction. The base case $k = 1$ was obtained as before, and for $k > 1$, one has

$$
\begin{aligned}
\|u(t)\|_{h_0,1,\phi^{(k)}} &\leq \|u(t_k)\|_{h_0,1,\phi^{(k)}} + L_1(t_{k-1};t) + L_2(t_{k-1};t) + L_3(t_{k-1};t) \\
&\leq \left(1 + C(M/h_0)\right) \left(\|u(t_k)\|_{h_0,1,\phi^{(k-1)}} + L_1(t_{k-1};t) + L_2(t_{k-1};t) + L_3(t_{k-1};t)\right) \\
&\leq \left(1 + C(M/h_0)\right)^{k(t)-1} \left(\|u(0)\|_{h_0,1,\phi^{(1)}} + L_1(0;t) + L_2(0;t) + L_3(0;t)\right).
\end{aligned}
$$

Finally, multiplying by $\left(1 + C(M/h_0)\right)$ twice more, we are able to replace the discrete semi-norm on both sides to $\|u(t)\|_{h_0,1}$ or $\|u(0)\|_{h_0,1}$. This gives

$$
\begin{aligned}
\|u(t)\|_{h_0,1} &\leq \left(1 + C(M/h_0)\right)^{k(t)+1} \left(\|u(0)\|_{h_0,1} + L_1 + L_2 + L_3\right) \\
&= (L_0 + L_1 + L_2 + L_3),
\end{aligned}
$$

which finishes the proof of Theorem 3.1.

5. Proof of Theorem 3.3.

This section is devoted to the proof of Theorem 3.3. We first note that all of our estimates are on domains with bounded measure, which lets us immediately bound any L^p or $W^{s,p}$ norms by L^q or $W^{s,q}$ with any $q \geq p$.

Within this section, the generic constant C that we use depends on , the exponents $p; q$ and the constant in the structural assumptions (1.14), and also on T and the exponent s in the statement of Theorem 3.3.

Step 1: Constructing the space and time partitions. Choose m 2 N, and introduce the straightforward time partition

$$0 = t_0 < t_1 < \quad < t_m = T;$$

$$= \frac{T}{m}; \qquad t_l = (l=m)T \qquad 8l = 0; : : : ; m:$$

A very small choice of time step will lead to large terms in (3.3), while a larger value of allows the velocity eld to oscillate more in each time interval making controlling L_3 in (3.3) more dicult. Thus, determine the optimal choice of turns out to a key step of the proof.

the partition is more complicated in the spatial direction. We divide the mesh into large partitions roughly corresponds to large hypercubes of size with x that will be determined later. At this moment, it suces to assume that we have for example 8 x 1.

More precisely, we divide
into subdomains f
$_k g_{k2J}$, roughly centered around points f$y_k g_{k2J}$ 2 R^d such that

$$B(y_k;) \quad {}_k \quad B(y_k; C) \qquad \text{and} \qquad j@ \; {}_k j \quad C^{d \; 1}:$$

Next, choose a partition f$V_k g_{k2J}$ of the index set $\overline{V}$
, by assigning i 2 V
to any V_k such that

$$\text{supp}_i \setminus {}_k = ?:$$

By denition $\text{supp}_i \setminus$
= ? and one can nd at least one
$_k$ intersecting supp_i so that $\overline{S}_{k2J} V_k = V$
.

Then, for all k 2 J, dene

$$_k = \sum_{i2V_k} {}_i; \qquad U_k = \text{supp} \; {}_k; \qquad U_k^0 = fx \; 2 \; R^d : \; {}_k(x) = 1g:$$

By denition $U_k^c U_k$ and it is easy to verify that

$$U_k^c \quad {}_k + B_x; \qquad {}_k \quad U_k + B_x; \qquad \text{and} \qquad C \quad ^{1d} \; k \; {}_k k_{L^1(} \; {}_{)} \; k \; {}_k k_{L^1} \quad C^d:$$

Moreover, for all k 2 J, dene the \boundary" of V_k as those indices that do not intersect with any index in another part of the domain or

$$@V_k = V_k \; n f i \; 2 \; V_k \; j \;\; \text{if } j \; 2 \; V \text{ and } \text{supp}_i \setminus \text{supp}_j = ?; \text{ then } \; {}_k(\text{supp}_j) \;\; 1g:$$

Then this boundary has codimension 1 in the sense that by decomposing $_k = {}^{(0)}_k + {}^{(1)}_k$

$$^{(0)}_k = \sum_{i2(V_k n@V_k)} {}_i; \qquad ^{(1)}_k = \sum_{i2@V_k} {}_i; \qquad U_k^{00} = fx \; 2 \; R^d : \; {}^{(0)}_k(x) = 1g;$$

one has

$_k$ $U^{00} + {}^k B_{4x}$; $(U_k \; n \; U^{00}_k)$ @ ^{k}

$_k$ + B_{4x}; k $^{(1)}k_{L^1}$ C^d ^1x:

Observe that the number $|J|$ of domains $_k$ can be estimated by $|J| \lesssim |\,|/^d$. Hence the previous estimates yield

$$\Big|\bigcup_{j:\,k\in J}(U_k \setminus U_k^{00})\Big| \leq C(x/)\,|\,| \leq \sum_{k\in J} \|{}_k^{(1)}\|_{L^1} \leq C(x/)\,|\,|$$

For later discussion, we define $\partial V = \bigcup_{k\in J} \partial V_k$ as the set of all boundary indices, and $\partial E = \{(i;j) \in E : i \in \partial V \text{ or } j \in \partial V\}$ the set of all boundary edges.

Step 2: Constructing the virtual coordinates in Theorem 3.1. Introduce

$$(\tilde{b}_i(t))_{i\in V} = P_C \tilde{b}(t); \quad (\tilde{a}_{i;j}(t))_{i;j\in V} = P_F \tilde{b}(t):$$

the discretization of $\tilde{b}(t;x)$ on faces and cells as in (2.1) and (2.2) respectively.

We introduce another piecewise constant velocity field corresponding to the partition we just constructed: For all $1 \leq l \leq m$ and all $k \in J$, we take the average of $b(t;x)$ on $[t_{l-1};t_l] \times V_k$ in the following sense

$$b^{l;k} = \frac{1}{|t_l - t_{l-1}|\,\|{}_k\|_{L^1()}} \int_{[t_{l-1};t_l]} \tilde{b}(t;x)\,{}_k(x)\,dtdx;$$

and define the "piecewise" extension $b(t;x)$ by

$$b(t;x) = \sum_{:\,k\in J} {}_k(x)\,b^{l;k}; \qquad \forall (t;x) \in [t_{l-1};t_l) \times$$

We finally introduce as before the discretization on faces and cells of $b(t;x)$,

$$(b_i(t))_{i\in V} = P_C b(t); \quad (a_{i;j}(t))_{i;j\in V} = P_F b(t):$$

The extension $b(t;x)$ from $(b^{l;k})_{1\leq l\leq m;k\in J}$ is not exactly piecewise. Nevertheless, by our construction in Step 1, each $_k$ has compact support on U_k, with $_k = 1$ on the set U_k^0, with $U_k \setminus U_k^0$ small. In such sets U_k^0, one has that $b = b^{l;k}$. Moreover, for interior indices $i \in V_k \setminus \partial V_k$, $a_{;i}(t)$ and $b_i(t)$ are not only the discretization of $b(t;\cdot)$, but also coincide with the discretization of the constant velocity field $b^{l;k}$.

Theorem 3.3 assumes that Definition (3.5) applies for constant fields. This yields virtual coordinates $(\hat{x}_i(b_c))_{i\in V};\ b_c \in \mathbb{R}^d$,

$$\sum_{i'\in V} \left(\hat{x}_{i'}\big(b^{l;k}\big) - \hat{x}_i\big(b^{l;k}\big)\right) a_{;i}(t) = b_i(t)_i + \hat{r}\big(b^{l;k}\big)_{ji}; \qquad \forall t \in [t_{l-1};t_l];\ i \in V_k \setminus \partial V_k: \tag{5.1}$$

Inspired by (5.1), we choose piecewise constant in time virtual coordinates on the mesh $(\tilde{x}_i(t))_{i\in V}$ by

$$\tilde{x}_i(t) = \hat{x}_i\big(b_i(t)\big); \quad \forall t \in [0;T];\ i \in V:$$

Then in the interiors indices $i \in \bar{V} \setminus \partial V$, (5.1) implies that

$$\sum_{i'\in V} (\tilde{x}_{i'}(t) - \tilde{x}_i(t))\,a_{;i}(t) = b_i(t)_i + \hat{r}^i(b_i(t))_i; \qquad \forall t \in [0;T];\ i \in \bar{V} \setminus \partial V: \tag{5.2}$$

By our construction, we can see that on each time interval $(t_{k-1};t_k)$, the virtual coordinates $(\tilde{x}_i(t))_{i\in V}$ are time-independent and we may use the notation $(\tilde{x}_i^{(k)})_{i\in V}$ as in Theorem 3.1.

It is straightforward to deduce the uniform bounds

$$\sup_{(i,i')\in E,\,1\le k\le m} \left|\tilde x_i^{(k)} - \tilde x_{i'}^{(k)}\right| < 2M;$$

$$\sup_{i\in V,\,1\le k\le m} \left|\tilde x_i^{(k)} - x_i\right| < 2M;$$

because the $(\tilde x_i^{(k)})_{i\in V}$ are obtained through $(\hat X_i(b_C))_{i\in V}$. Therefore, those virtual coordinates satisfy the requirement of Theorem 3.1.

We reformulate the residue equation (3.2) as

$$\sum_{i'\in V} (\tilde x_{i'}(t) - \tilde x_i(t))\,a_{i',i}(t) = \tilde b_i(t)_i + r_i(t)_i, \qquad \forall t\in[0,T],\; i\in V. \tag{5.3}$$

Note that if $i\in V\setminus\tilde V$, one has $a_{i',i}\equiv 0$ for all $i'\in V$ and $\tilde b_i = (P_F\tilde b)_i = 0$. Hence the residue r_i vanishes for all $i\in V\setminus\tilde V$.

Subtracting (5.2) from (5.3), we obtain

$$\sum_{i'\in V} (\tilde x_{i'}(t) - \tilde x_i(t))\,\big(a_{i',i}(t) - \tilde a_{i',i}(t)\big) = b_i(t) - \tilde b_i(t)_i + r_i(t) - \Lambda(b_i(t))_i, \qquad \forall t\in[0,T],\; i\in\tilde V\setminus\partial V. \tag{5.4}$$

By definition, we have $|\Lambda_i(b_i(t))| \le (\Lambda_{max})_i|b_i(t)|$ with in addition,

$$\|(a_{i,j}-\tilde a_{i,j})_{i,j\in V}\|_{L^p([0,T]F)} \lesssim \|P_F(b-\tilde b)\|_{L^p([0,T]F)} \le C\|b-\tilde b\|_{L^p([0,T])};$$
$$\|(b_i-\tilde b_i)_{i\in V}\|_{L^p([0,T]C)} \lesssim \|P_C(b-\tilde b)\|_{L^p([0,T]F)} \le C\|b-\tilde b\|_{L^p([0,T])}.$$

Therefore, the main obstacle to bound $\|r\|_{L^p([0,T]C)}$ is to derive good estimates on $\|\tilde b - b\|_{L^p}$.

Step 3: Bounding $\|\tilde b - b\|_{L^p}$. We introduce the average in time of $b(t,x)$ by

$$b^l(x) = \frac{1}{\delta}\int_{[t_{l-1},t_l]} \tilde b(t,x)\,dt, \qquad \forall x\in\Omega,$$
$$b^0(t,x) = b^l(x), \qquad \forall (t,x)\in[t_{l-1},t_l)\times\Omega.$$

It is obvious that the two ways of averaging of velocity field

$$\tilde b \mapsto b; \qquad \tilde b \mapsto b^0$$

and the discretizations P_C, P_F are all linear mappings.

We can first quantify the oscillation in time by comparing $\tilde b$ and b^0. For fixed $x\in\mathbb{R}^d$, the function $b^0(\cdot,x)$ is constant on each time interval $[t_{l-1},t_l)$, $1\le l\le m$. Therefore,

$$\int_{[0,T]}\int \left|\tilde b(t,x) - b^0(t,x)\right|^p dx\,dt = \int_{[0,T]}\int \left|\tilde b(t,x) - \frac{1}{\delta}\int_{I(t)}\tilde b(s,x)\,ds\right|^p dx\,dt \le \frac{1}{\delta}\int_{[0,T]}\int\int_{I(t)} \left|\tilde b(t,x) - \tilde b(s,x)\right|^p ds\,dx\,dt,$$

where $I(t)$ denotes the interval $\ni t$, $t\in[t_{l-1},t_l)$.

Since $t_l - t_{l-1} = \varepsilon$, we have that

$$\left(\int_{[0,T]\times\Omega} |\tilde b(t,x) - b^0(t,x)|^p\,dx\,dt\right)^{1/p} \le C\,\varepsilon\,\|b\|_{L^p_x(W^{1,p}_t)([0,T]\times\Omega)}.$$

Through interpolation, this shows that for $0 \le s \le 1$,

$$\|\tilde b - b^0\|_{L^p([0,T]\times\Omega)} \le C\,\varepsilon^s\,\|b\|_{L^p_x(W^{s,p}_t)([0,T]\times\Omega)}.$$

We can also bound spatial oscillations on $\bar b$ thanks to b^0. For any t, denote l s.t. $t \in [t_{l-1}, t_l)$, and write

$$\bar b(t,x) = \sum_{k\in J} \psi_k(x)\, b^{l,k} = \sum_{k\in J} \psi_k(x) \frac{1}{|t_l - t_{l-1}|\,\|\psi_k\|_{L^1(\Omega)}} \int_{[t_{l-1},t_l]\times\Omega} \tilde b(t,y)\,\psi_k(y)\,dt\,dy$$

$$= \sum_{k\in J} \psi_k(x) \frac{1}{\|\psi_k\|_{L^1(\Omega)}} \int_\Omega b^0(t,y)\,\psi_k(y)\,dy.$$

Moreover, since $\sum_{k\in J} \psi_k(x) = \sum_{i\in \bar V} \phi_i(x) = 1$ for all $x \in \Omega$, one has

$$b^0(t,x) - \bar b(t,x) = \sum_{k\in J} \psi_k(x)\, b^0(t,x) - \sum_{k\in J} \psi_k(x) \frac{1}{\|\psi_k\|_{L^1(\Omega)}} \int_\Omega b^0(t,y)\,\psi_k(y)\,dy$$

$$= \sum_{k\in J} \psi_k(x) \frac{1}{\|\psi_k\|_{L^1(\Omega)}} \int_\Omega \left(b^0(t,x) - b^0(t,y)\right) \psi_k(y)\,dy.$$

Therefore,

$$\int_{[0,T]\times\Omega} |b^0(t,x) - \bar b(t,x)|^q\,dx\,dt$$

$$= \int_{[0,T]\times\Omega} \left|\sum_{k\in J} \psi_k(x) \frac{1}{\|\psi_k\|_{L^1(\Omega)}} \int_\Omega \left(b^0(t,x) - b^0(t,y)\right) \psi_k(y)\,dy\right|^q dx\,dt$$

$$\le \int_{[0,T]}\int_\Omega \sum_{k\in J} \psi_k(x) \frac{1}{\|\psi_k\|_{L^1(\Omega)}} \int_\Omega |b^0(t,x) - b^0(t,y)|^q\,\psi_k(y)\,dy\,dx\,dt$$

$$\le C \int_{[0,T]} \sum_{k\in J} \frac{1}{|U_k\cap\Omega|} \int_{U_k\cap\Omega}\int_{U_k\cap\Omega} |b^0(t,x) - b^0(t,y)|^q\,dy\,dx.$$

We recall that the last part of our assumption (1.14) states that $\#\{k \in V : (\mathrm{supp}\,\psi_k) \cap B(x,\varepsilon) \neq \emptyset\} \le C$ for all $x \in \mathbb{R}^d$. From their construction, any point $x \in$

also belongs to at most C sets U_k so that

$$\int_{[0,T]} \int b^0(t;x) \quad b(t;x)^q \, dxdt \quad C \int_{[0,T]} \int \frac{1}{|B_C|} \int 1_{B_C}(x \quad y) b^0(t;x) \quad b^0(t;y)^q \, dydx;$$

by which we conclude

$$\|b^0 \quad b\|_{L^q([0,T] \quad e \quad q \quad x)} \quad C\|b\|_{L_t(W^{1;q})([0,T]\quad)} :$$

Finally for any $1 \quad p < q \quad 1$, one obtains that

$$\|b \quad b\|_{L^p([0,T] \quad e \quad x \quad t \quad e \quad t \quad x)} \quad C \quad {}^s\|b\|_{L^p(W^{s;p})([0,T]\quad)} + \|b\|_{L^q(W^{1;q})([0,T]\quad)} : \tag{5.5}$$

As we mentioned in Step 2, one can also bound $(\tilde a_{i;j} - a_{i;j})_{i;j\in V}$ and $(\tilde b_i - b_i)_{i\in V}$ by the right hand side of (5.5).

Step 4: Optimizing all parameters. We finally combine all previous estimates to try to derive the best bound on the residue term $(r_i(t))_{i\in V}$.

On the interior set $\overline{V} \setminus \partial V$, by expanding (5.4) and using that $|\tilde x_{i_0}(t) - \tilde x_i(t)| \le 2M$, we have that

$$
\begin{aligned}
\|r\,1_{\overline{V}\setminus\partial V}\|_{L^p([0;T]C)} \le{}& C(M/x)\|(a_{i;j} - \tilde a_{i;j})_{i;j\in V}\|_{L^p([0;T]F)} + \|(b_i - \tilde b_i)_{i\in V}\|_{L^p([0;T]C)} + \|((\hat r_{\max})_i|\tilde b_i|)_{i\in V}\|_{L^p([0;T]C)} \\
\le{}& C(M/x)\|(a_{i;j} - \tilde a_{i;j})_{i;j\in V}\|_{L^p([0;T]F)} \\
&+ C(M/x)\|(\tilde a_{i;j} - a_{i;j})_{i;j\in V}\|_{L^p([0;T]F)} + \|(b_i - \tilde b_i)_{i\in V}\|_{L^p([0;T]C)} + \\
&C\|\hat r_{\max}\|_{L^{1/(1/p-1/q)}([0;T]C)}\|(b_i)_{i\in V}\|_{L^q([0;T]C)}.
\end{aligned}
$$

We recall that the admissible family of virtual coordinates has residue bound M in $L^{1/(1/p-1/q)}$. Applying (5.5) and using $1 < (M/x)$ leads to

$$
\begin{aligned}
\|r\,1_{\overline{V}\setminus\partial V}\|_{L^p([0;T]C)} \le{}& C(M/x)\|(a_{i;j})_{i;j\in V} - P_F\,\tilde b\|_{L^p([0;T]F)} \\
&+ C(M/x)^s\left(\|\tilde b\|_{L^p(W^{s;p}_x)([0;T])} + \|\tilde b\|_{L^q(W^{1;q}_t)([0;T])}\right) \\
&+ CM\|\tilde b\|_{L^q([0;T])}.
\end{aligned}
\tag{5.6}
$$

As for the boundary ∂V, we directly expand (5.3) to find that

$$
\begin{aligned}
\|r\,1_{\partial V}\|_{L^p([0;T]C)} \le{}& C(M/x)\|(a_{i;j})_{i;j\in V}1_{\partial E}\|_{L^p([0;T]F)} + \|(b_i)_{i\in V}1_{\partial V}\|_{L^p([0;T]C)} \\
\le{}& C(M/x)(x/\,)T\,|\;|^{1/p-1/q}\left(\|(a_{i;j})_{i;j\in V}\|_{L^q([0;T]F)} + \|\tilde b\|_{L^q([0;T])}\right);
\end{aligned}
\tag{5.7}
$$

by Hölder inequality.

Since the residue r_i vanishes for all $i \in V \setminus \overline{V}$, we have

$$
\begin{aligned}
\|r\|_{L^p([0;T]C)} \le{}& \|r\,1_{\overline{V}\setminus\partial V}\|_{L^p([0;T]C)} + \|r\,1_{\partial V}\|_{L^p([0;T]C)} \le C(M/x)\Bigg(\|(a_{i;j})_{i;j\in V} - P_F\,\tilde b\|_{L^p([0;T]F)} \\
&+ C(M/x)^s\left(\|\tilde b\|_{L^p(W^{s;p}_x)([0;T])} + \|\tilde b\|_{L^q(W^{1;q}_t)([0;T])}\right) \\
&+ (x/\,)^{1/p-1/q}\left(\|(a_{i;j}(t))_{i;j\in V}\|_{L^q([0;T]F)} + \|\tilde b\|_{L^q([0;T])}\right)\Bigg) \\
&+ CM\|\tilde b\|_{L^q([0;T])}
\end{aligned}
\tag{5.8}
$$

$);$

We are now ready to choose the parameters and . We also need to control the term $= 1 + C(M=h_0)$ $_k^{(t)+1}$ in Theorem 3.1, where $k(t)$ represents the number of times $(\mathbf{e}_i(t))_{i2V}$ jumps within $[0;t]$. Notice that by increasing the constant C we have exp $Ck(t)(M=h_0)$, and we can bound $k(t)$ by $k(T)$ $CT=$. To control L_3 and simultaneously, we use the following choice

$$= (x)^{\frac{1=p \quad 1=q}{1+(1=p \quad 1=q)}}; \qquad = (x\, M=M)^{1=(1+s)}:$$

It rst results the claimed bound on in (3.6), namely

$$= \exp \; C(1=h_0)(M^s M=x)^{1=(1+s)}:$$

This also yields a bound on the main residue term L_3 in Theorem 3.1 by

$$L_3 = C(j \log h_0 j \;=h_0)\, k r_t k_{L^1 L_x([0;t]C)}\, k u_t k_{L^1_x L^p([0;t]C)}$$
$$C(j \log h_0 j \;=h_0) k r_{t\,x} k_{L^p L^p([0;t]C)}\, k u k_{L_t^p L_x^1([0;t]C)}:$$

Inserting (5.8) on $krk_{L^p([0;T]C)}$ nally provides "

$$L_3 \;\; C\,(j \log h_0 j \;=h_0)\;\; (M=x)(a_{i;j})_{i;j2V} \quad P_F\, b_{L^p([0;T]F)}$$
$$+\;(M^s M=x)^{1=(1+s)}\, k b k_{L^p_t(W_x^{s;p})}$$
$$+\quad M(x)^{\frac{1}{1+(1=p \quad 1=q)}}\quad k(a_{i;j}(t))_{i;j2V}\, k_{L^q([0;T]F)} + k b k_{L^q_t(W_x^{1;q})}$$
$$\#$$
$$+\; M k b k_{L^q_t(L^q_x)}\;\; k u k_{L^1_t L^1_x([0;t]C)};$$

which nishes the proof of Theorem 3.3.

6. Proof of Theorem 3.4.

We rst reduce, in subsection 6.1, the innite linear system (3.7) to a nite linear system whose variables are $f\hat{x}_i g_{i2V_0}$, by making use of the periodic nature of the mesh. Due to the geometric nature of the meshes, both the matrix and the inhomogeneous term in the linear system have certain properties, which we focus on in subsections 6.2 and 6.3. Finally, in subsection 6.4 we conclude the uniform boundedness result.

6.1. The linear system for periodic meshes.

We rewrite (3.7) as

$$P_{j2V}\, a_{j;i}(b_c)\;\;\hat{x}_i(b_c) = P_{j2V}\, a_{j;i}(b_c)\hat{x}_j(b_c) \qquad b_{ci}; \qquad 8i\; 2\; V: \tag{6.1}$$

Since the mesh is periodic as in Denition 1.6, we are looking for periodic solutions as well, namely solutions satisfying

$$\hat{x}_{[m](i)}(b_c) = \hat{x}_i(b_c) + [m]L; \quad 8b_c\; 2\; R^d; [m]\; 2\; Z^d; i\; 2\; V; \tag{6.2}$$

where $[m]L = P_{k=1}^d m_k L_k \; 2\; R^d$.

The following lemma makes explicit the nite linear systems that the variables $\hat{x}_i$, reduced to $i\; 2\; V_0$, need to solve to be solutions to (6.1) over the full mesh.

Lemma 6.1. Let $(C;F) = f_i g_{i2V}; f n_{i;j} g_{(i;j)2E}$ be a periodic mesh as in Denition 1.6 and $b\;\; b_c$ be a constant velocity eld. Consider a function $(\hat{x}_i)_{i2V}$ dened on all cells, satisfying (6.2). It is a solution of (6.1) if and only if its restriction $(\hat{\chi}_i)_{i \in V_0}$ satisfying the following nite linear system: for all $i \in V_0$,

$$
\begin{aligned}
&\sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](j);i}(b_c) \quad \hat{\chi}_i(b_c) \\
= \; & \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](j);i}(b_c) \quad \hat{\chi}_j(b_c) \\
+ \; & \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](j);i}(b_c)[m]L \qquad b_{c_i}:
\end{aligned}
\tag{6.3}
$$

In the formulation above we let $a_{[m](j);i} = 0$ if $[m](j)$ lie outside the mesh V.

Proof. By Denition 1.6, for any $i \in V_0$ and any $l \in V$, such that $a_{l;i} = 0$, there exists unique $(j;[m]) \in V_0 \; \mathbb{Z}^d$ such that $l = [m](j)$. Therefore (6.1) is identical to

$$
\sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{j;i}(b_c) \quad \hat{\chi}_i(b_c) = \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{j;i}(b_c) \hat{\chi}_j(b_c) \qquad b_{c_i}; \qquad \forall i \in V_0:
$$

By the periodic condition (6.2), this is also identical to (6.3), which nishes the proof. We now introduce matrix notations on (6.3) to simplify the discussion in later subsections,

$$
A(b_c) = (A_{ij}(b_c))_{i;j \in V_0}; \quad A_{ij}(b_c) = \sum_{[m] \in \mathbb{Z}^d} a_{[m](i);j}(b_c):
$$

Let $A^T(b_c) = (A_{ji}(b_c))_{i;j \in V_0}$ denote the usual transpose matrix. In addition, let

$$
\begin{aligned}
&(b_c) = \mathrm{diag}\{_{ii}(b_c)\}_{i \in V_0}; \qquad _{ii}(b_c) = \sum_{l \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](l);i}(b_c); \\
&\varphi(b_c) = (\varphi_i(b_c))_{i \in V_0}; \quad \varphi_i(b_c) = \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](j);i}(b_c)[m]L \qquad b_{c_i}:
\end{aligned}
$$

Then the linear system (6.3) can be rewritten as

$$
_{ii}(b_c)\hat{\chi}_i(b_c) = \sum_{j \in V_0} A_{ji}(b_c)\hat{\chi}_j(b_c) + \varphi_i(b_c); \quad \forall i \in V_0;
$$

or more compactly,

$$
(\quad A^T)(b_c)\hat{\chi}(b_c) = \varphi(b_c); \tag{6.4}
$$

where $(\quad A^T)(b_c)$ is a square matrix in $\mathbb{R}^{V_0 V_0}$ and $\hat{\chi}(b_c); \varphi(b_c) \in \mathbb{R}^{V_0 d}$. We emphasize that (6.4) each d coordinate separately, namely it should be understood as

$$
(\quad A^T)(b_c)\hat{\chi}(b_c) = \varphi(b_c); \qquad = 1 ::: d;
$$

with the same matrix $(\quad A^T)(b_c)$ for each coordinate.

We may also omit the variable $b_c \in \mathbb{R}^d$ in the bracket when there is no ambiguity.

6.2. Recasting (6.4) into a discrete diusion operator. We can characterize the matrix $(\quad A^T)$ in the following manner.

Proposition 6.2. Dene the space of discrete diusion operators as

$$
M(n) = \Big\{ M \in \mathbb{R}^{n n} : M_{ii} \; 0; M_{ij} \; 0; \sum_{l=1}^{n} M_{il} = \sum_{l=1}^{n} M_{li} = 0; \; \forall i; j = 1; ::: ; n; \; i = j \Big\}:
$$

Then for all $b_c \in \mathbb{R}^d$, $(\quad A^T)(b_c) \in M(|V_0|)$.

Proof. The condition $M_{ii} \; 0$ in the denition of $M(n)$ is actually redundant. It can be easily derived by the conditions $M_{ij} \; 0$ and $\sum_{l=1}^{n} M_{il} = 0$. Thus, if suce to check that $(\quad A^T)$ satises the remaining properties in the denition.

Firstly, since A^T is non-negative and is diagonal, it is obvious that the non-diagonal entries of $(\quad A^T)$ are non-positive.

Secondly, the identity $\sum_{l\in V_0}(A-A^T)_{il}=0$ can be derived by an expansion

$$\begin{aligned}\sum_{l\in V_0}(A-A^T)_{il} &= \sum_{l\in V_0} A_{il} - \sum_{l\in V_0} A_{li}\\ &= \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i} - \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i} = 0.\end{aligned}$$

Now turning to the last property, as the velocity field is constant, we have that $a_{[m](i);l} = a_{i;[-m](l)}$. By taking the same expansion as before, we have

$$\begin{aligned}\sum_{l\in V_0}(A-A^T)_{li} &= \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i} - \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](i);l}\\ &= \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i} - \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{i;[-m](l)}.\end{aligned}$$

Hence to prove $\sum_{l\in V_0}(A-A^T)_{li}=0$, it suffices to show that

$$\sum_{l\in V} a_{i;l} - \sum_{l\in V} a_{l;i} = 0;$$

Since the constant velocity field is divergence-free, applying the divergence theorem to the extended cell C_i, we have

$$\begin{aligned}\sum_{l\in V} a_{i;l} - \sum_{l\in V} a_{l;i} &= \sum_{l\in V}\int \left(b_c\cdot n_{l;i}(x)\right)_- dx - \sum_{l\in V}\int \left(b_c\cdot n_{l;i}(x)\right)_+ dx\\ &= -\sum_{l\in V}\int b_c\cdot n_{l;i}(x)\,dx\\ &= -\int b_c\cdot\nabla\varphi_i(x)\,dx = \int_{C_i}(\operatorname{div} b_c)\varphi_i(x)\,dx = 0;\end{aligned}$$

which finish the proof.

Our next result is an inequality bounding the entries of $\varphi(b_c)$ by the entries of $(A-A^T)(b_c)$, which still relies on the divergence theorem but in a more intricate way.

Proposition 6.3. For all $b_c\in\mathbb{R}^d$ and $V_1\subset V_0$, the inhomogeneous term $\varphi(b_c)$ in the linear system (6.4) satisfies

$$\sum_{l\in V_1}\varphi(b_c) \le C(V_0)\sum_{i\in V_1;\,j\in V_0\setminus V_1} |A_{ij}(b_c)| + |A_{ji}(b_c)|;$$

where

$$C(V_0) = \sup_{x,y\in\bigcup_{i\in V_0}\operatorname{supp}\varphi_i} |x-y|.$$

Proof. By choosing an appropriate basis of $\mathbb{R}^d$, we may assume $0\in\bigcup_{i\in V_0}\operatorname{supp}\varphi_i$ and $b_c=(0,\dots,0,1)$ without loss of generality. For $r=1,\dots,(d-1)$, we then have

$$(\varphi_i)_r = \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i}\sum_{p=1}^n m_p(L_p)_r,\qquad i\in V_0;$$

while for $k=d$, the equation reads

$$(\varphi_i)_d = \sum_{l\in V_0}\sum_{[m]\in\mathbb{Z}^d} a_{[m](l);i}\sum_{p=1}^n m_p(L_p)_d - {}_i,\qquad i\in V_0.$$

For $r=1,\dots,d-1$, notice that $\operatorname{div}(x_r b_c) = \partial_d x_r = 0$, so that

$$\int x_r\, b_c\cdot\nabla\sum_{i\in V_1}\varphi_i(x)\,dx = -\int \operatorname{div}(x_r b_c)\sum_{i\in V_1}\varphi_i(x)\,dx = 0.$$

For $k = d$, we have instead that $\operatorname{div}(x_d b_c) = \partial_d x_d = 1$, leading

$$\int x_d b_c \cdot \nabla \Big(\sum_{i \in V_1} \psi_i(x)\Big)\, dx = -\int \operatorname{div}(x_d b_c) \sum_{i \in V_1} \psi_i(x)\, dx = -\sum_{i \in V_1} \psi_i.$$

We can summarize all those relation in vector form, as

$$\varphi_i = \sum_{l \in V_0} \sum_{[m] \in \mathbb{Z}^d} a_{[m](l);i} \Big(\sum_{p=1}^n m_p L_p\Big) \cdot b_c, \qquad i \in V_0;$$

and

$$\int x \cdot b_c \nabla \Big(\sum_{i \in V_1} \psi_i(x)\Big)\, dx = -\sum_{i \in V_1} b_c \psi_i.$$

For any i, j, denote

$$X^+_{j;i} = \int \big(x \cdot b_c\, n_{j;i}(x)\big)^+ dx;$$

$$X^-_{j;i} = \int \big(x \cdot b_c\, n_{j;i}(x)\big)^- dx;$$

$$X_{j;i} = X^+_{j;i} - X^-_{j;i} = \int x \cdot b_c\, n_{j;i}(x)\, dx.$$

Notice that $n_{[m](j);i}(x) = n_{j;[-m](i)}(x - \sum_{p=1}^n m_p L_p)$ by our periodic assumption, so that one obtains that

$$\begin{aligned}
& X^+_{[m](j);i} - X^-_{[-m](i);j} \\
&= \int \big(x \cdot b_c\, n_{[m](j);i}(x)\big)^+ dx - \int \big(x \cdot b_c\, n_{[-m](i);j}(x)\big)^- dx \\
&= \int \big(x \cdot b_c\, n_{[m](j);i}(x)\big)^+ dx - \int \big(x \cdot b_c\, n_{j;[-m](i)}(x)\big)^+ dx \\
&= \int \big(x \cdot b_c\, n_{[m](j);i}(x)\big)^+ dx - \int \Big(\big(x - \textstyle\sum_{p=1}^n m_p L_p\big) \cdot b_c\, n_{[m](j);i}(x)\Big)^+ dx \\
&= \int \big(b_c\, n_{[m](j);i}(x)\big)^+ dx \cdot \textstyle\sum_{p=1}^n m_p L_p \\
&= a_{[m](j);i} \cdot \textstyle\sum_{p=1}^n m_p L_p.
\end{aligned}$$

Hence φ_i can be reformulated as

$$\varphi_i = \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} \Big(X^+_{[m](j);i} - X^-_{[-m](i);j}\Big) \cdot b_c, \qquad i \in V_0;$$

and the summation of φ_i over any $V_1 \subset V$ reads

$$\sum_{i \in V_1} \varphi_i = \sum_{i \in V_1} \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} \Big(X^+_{[m](j);i} - X^-_{[-m](i);j}\Big) - \sum_{i \in V_1} b_c \psi_i.$$

Decompose

$$\begin{aligned}
& \sum_{i \in V_1} \sum_{j \in V_0} \sum_{[m] \in \mathbb{Z}^d} \Big(X^+_{[m](j);i} - X^-_{[-m](i);j}\Big) \\
&= \sum_{i \in V_1} \sum_{j \in V_1} \sum_{[m] \in \mathbb{Z}^d} \Big(X^+_{[m](j);i} - X^-_{[-m](i);j}\Big) \\
&\quad + \sum_{i \in V_1} \sum_{j \in V_0 \setminus V_1} \sum_{[m] \in \mathbb{Z}^d} \Big(X^+_{[m](j);i} - X^-_{[-m](i);j}\Big) \\
&= Y_1 + Y_2.
\end{aligned}$$

The term Y_1 can be further simplified. By switching i and j and take the inverse $-m$ to m in $\mathbb{Z}^d$, we have

$$\begin{aligned}
Y_1 &= \sum_{i\in V_1}\sum_{j\in V_1}\sum_{[m]\in \mathbb{Z}^d} X^+_{[m](j);i} - \sum_{i\in V_1}\sum_{j\in V_1}\sum_{[m]\in \mathbb{Z}^d} X^-_{[-m](i);j}\\
&= \sum_{i\in V_1}\sum_{j\in V_1}\sum_{[m]\in \mathbb{Z}^d} X_{[m](j);i}\,.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\sum_{i\in V_1} b_{ci} &= \int \nabla_x b_c \cdot \sum_{i\in V_1} r_i(x)\,dx\\
&= \int \nabla_x b_c \cdot \sum_{i\in V_1}\sum_{j\in V_0}\sum_{[m]\in \mathbb{Z}^d} n_{[m](j);i}(x)\,dx\\
&= \sum_{i\in V_1}\sum_{j\in V_0}\sum_{[m]\in \mathbb{Z}^d} X_{[m](j);i}\\
&= \sum_{i\in V_1}\sum_{j\in V_1}\sum_{[m]\in \mathbb{Z}^d} X_{[m](j);i} + \sum_{i\in V_1}\sum_{j\in V_0\setminus V_1}\sum_{[m]\in \mathbb{Z}^d} X_{[m](j);i}\\
&= Y_1 + Y_2 + \sum_{i\in V_1}\sum_{j\in V_0\setminus V_1}\sum_{[m]\in \mathbb{Z}^d}\left[X_{[m](j);i} - X^+_{[m](j);i} + X^-_{[-m](i);j}\right].
\end{aligned}$$

Therefore, if we define

$$\begin{aligned}
Y_3 &= \sum_{i\in V_1}\sum_{j\in V_0\setminus V_1}\sum_{[m]\in \mathbb{Z}^d}\left(X_{[m](j);i} - X^+_{[m](j);i} + X^-_{[-m](i);j}\right)\\
&= \sum_{i\in V_1}\sum_{j\in V_0\setminus V_1}\sum_{[m]\in \mathbb{Z}^d}\left(X^-_{[m](j);i} + X^-_{[-m](i);j}\right)
\end{aligned}$$

then we obtain that

$$\sum_{i\in V_1}\varphi_i = (Y_1 + Y_2) - (Y_1 + Y_2 + Y_3) = -Y_3.$$

Since

$$\left|X^-_{[m](j);i} + X^-_{[-m](i);j}\right| \le C(V_0)\left(a_{i;[m](j)} + a_{j;[-m](i)}\right) = C(V_0)\left(a_{[-m](i);j} + a_{[m](j);i}\right),$$

by taking the summation over $i\in V_1, j\in V_0$ and $[m]\in\mathbb{Z}^d$, it is straightforward that

$$|Y_3| \le C(V_0)\sum_{i\in V_1, j\in V_0\setminus V_1}\left(|A_{ij}(b_c)| + |A_{ji}(b_c)|\right),$$

which completes the proof.

6.3. Some properties of discrete diffusion operators. To begin with, let us recall the definition of irreducibility. A matrix is irreducible if it is not similar via a permutation of indices to a block upper triangular matrix with more than one block of strictly positive size. An equivalent definition is the following: Each matrix M can be associated to a directed graph G, with n vertices labeled with $1,\dots,n$. There is an edge from i to j in G if and only if $M_{ij}\neq 0$. Then M is irreducible if and only if G is strongly connected, i.e. one can reach any vertex starting from any vertex.

Let $I=\{1,\dots,n\}$, and consider any $J,\tilde J\subset I$. We denote by the submatrix $M(J,\tilde J)$, the matrix obtained by deleting from M all rows whose indexes are not in J and all columns whose indexes are not in $\tilde J$. More precisely, if $J=\{j_1,\dots,j_k\}$ and $\tilde J=\{\tilde j_1,\dots,\tilde j_{\tilde k}\}$, $M(J,\tilde J)$ it is defined by

$$M(J,\tilde J)_{i_1,i_2} = M_{j_{i_1},\tilde j_{i_2}}.$$

The following lemma shows that a discrete diffusion operator, if it is not irreducible, must be block-diagonal up to a permutation.

Lemma 6.4. For $M \in \mathcal{M}(n)$, there exists a decomposition $I_1 \cup \cdots \cup I_m = I = \{1,\ldots,n\}$, such that for all $k \in \{1,\ldots,m\}$, the diagonal square submatrix $M(I_k;I_k)$ is irreducible and for all $k,l \in \{1,\ldots,m\}, k \neq l$, the submatrix $M(I_k;I_l) = 0$. Moreover, the null space and image of M are

$$N(M) = \mathrm{span}\{1_{I_k}\}_{k=1}^{m},$$
$$\mathrm{Range}\, M = \{x \in \mathbb{R}^I : \textstyle\sum_{i \in I_k} x_i = 0, \forall 1 \le k \le m\}.$$

Proof. We are going to prove there exists a decomposition $I_1 \cup \cdots \cup I_m = I = \{1,\ldots,n\}$ by induction on the dimension n. When $n = 1$ there is nothing to prove. Let's consider $n \ge 2$ and assume the decomposition exists for $1,\ldots,(n-1)$. If M is an irreducible $n \times n$ matrix, again there is nothing to prove. If M is reducible, then it is similar to a block upper diagonal matrix via a permutation of indices, which means there exists $J \subsetneq I = \{1,\ldots,n\}$ s.t. $M_{ij} = 0$ for all $i \in I \setminus J$ and $j \in J$. Thus

$$0 = \sum_{i \in I}\sum_{j \in J} M_{ij} = \sum_{i \in J}\sum_{j \in J} M_{ij} + \sum_{i \in I\setminus J}\sum_{j \in J} M_{ij} = \sum_{i \in J}\sum_{j \in J} M_{ij}.$$

On the other hand,

$$\sum_{i \in J}\sum_{j \in I} M_{ij} = 0.$$

Hence

$$\sum_{i \in J}\sum_{j \in I\setminus J} M_{ij} = \sum_{i \in J}\sum_{j \in I} M_{ij} - \sum_{i \in J}\sum_{j \in J} M_{ij} = 0.$$

Since the off-diagonal entries are all non-negative, one must have $M_{ij} = 0$ for all $i \in J$ and $j \in I \setminus J$. Therefore $M(I; I\setminus J) = 0$ and $M(I\setminus J; I) = 0$.

Furthermore, $M(J;J) \in \mathcal{M}(|J|)$ and $M(I\setminus J; I\setminus J) \in \mathcal{M}(|I\setminus J|)$. Note that $|J|, |I\setminus J| < n$. Applying the induction argument on $M(J;J)$ and $M(I\setminus J; I\setminus J)$, we get decompositions $J = I_1^{(1)} \cup \cdots \cup I_{m^{(1)}}^{(1)}$ and $I \setminus J = I_1^{(2)} \cup \cdots \cup I_{m^{(2)}}^{(2)}$. It is easy to verify that the decomposition

$$I = I_1^{(1)} \cup \cdots \cup I_{m^{(1)}}^{(1)} \cup I_1^{(2)} \cup \cdots \cup I_{m^{(2)}}^{(2)}$$

satisfies the properties asserted in the lemma.

It remains to determine the null space and range of M. Assume first that M is irreducible. Let $x \in N(M)$ and

$$J = \{i \in I : x_i = \max_l x_l\}.$$

By construction, $J \neq \emptyset$ and we can argue by contradiction that $J = I$. If $J \neq I$ and $J \neq \emptyset$, by the irreducibility of M, one can find $i \in J$, $j \in I \setminus J$ such that $M_{ij} > 0$. However, for any i and so in particular any $i \in J$,

$$\textstyle\sum_j M_{ij} x_j = 0.$$

By our assumption of M,

$$\textstyle\sum_{j \in I\setminus\{i\}} (M_{ij} x_j) < \sum_{j \in I\setminus\{i\}} M_{ij} \max_l x_l = 0;$$

which implies that $x_j = \max_l x_l$ if $M_{ij} > 0$. Therefore, $j \in J$ whenever $i \in J$ and $M_{ij} > 0$, a contradiction. In conclusion, we have $J = I$ and $x_i =$ constant for all $i \in I$ or $N(M) = \mathrm{span}\{1\}$. As any column summation of M is zero, for any $x \in \mathbb{R}^I$, one has

$$\sum_i \sum_j M_{ij} x_j = \sum_j \Big(\sum_i M_{ij}\Big) x_j = 0.$$

Since $\operatorname{codim} \operatorname{Range} M = \dim N(M) = 1$, one has $\operatorname{Range} M = \{x \in \mathbb{R}^I : \sum_{i \in I} x_i = 0\}$. This concludes the proof when M is irreducible.

When M is reducible, we know, up to a permutation, that M is block diagonal and each diagonal block $M(I_k; I_k)$ is irreducible. Applying the previous result to each $M(I_k; I_k)$ completes the proof.

The next lemma allows us to estimate the ℓ^1 norm of M^{-1} when the non-negative entries of M are bounded from both above and below.

Lemma 6.5. Define

$$\mathcal{M}(n;\lambda_0;\lambda_1) = \{M \in \mathcal{M}(n) : M_{ij} = 0 \text{ or } \lambda_0 < |M_{ij}| < \lambda_1,\ \forall 1 \le i, j \le n, i \ne j\}.$$

Then there exists a constant $C(n;\lambda_0;\lambda_1) = C_2(n)\lambda_1^{(n-2)}\lambda_0^{-(n-1)}$ s.t. for any matrix $M \in \mathcal{M}(n;\lambda_0;\lambda_1)$ and $\psi \in \operatorname{Range} M$, there exists one $\hat\varphi$ satisfying $M\hat\varphi = \psi$ and $\|\hat\varphi\|_{\ell^1} \le C(n;\lambda_0;\lambda_1)\|\psi\|_{\ell^1}$. Moreover, for fixed $M \in \mathcal{M}(n;\lambda_0;\lambda_1)$ let us take the decomposition $I = I_1 \cup \cdots \cup I_m$ as in Lemma 6.4. The solution $\hat\varphi$ above is uniquely determined by imposing the conditions

$$\sum_{i \in I_k} \hat\varphi_i = 0, \quad \forall k = 1, \dots, m.$$

Proof. Let us first assume that M is irreducible. Let x be a solution of $Mx = \psi$. Since $N(M) = \mathrm{span}\{1\}$, we have that $M(x - \mu \mathbf{1}) = \psi$ for all $\mu \in \mathbb{R}$. By taking $\mu = \frac{1}{n}\sum_i x_i$ and $\hat\varphi = (x - \mu\mathbf{1})$ we have $M\hat\varphi = \psi$ and $\sum_i \hat\varphi_i = 0$. This solution is uniquely determined and it remains to show that there is a uniform bound $\|\hat\varphi\|_{\ell^1} \le C(n;\lambda_0;\lambda_1)\|\psi\|_{\ell^1}$.

Define

$$J_0 = \{1 \le i \le n : \hat\varphi_i = \max_{1\le j\le n} \hat\varphi_j\};$$
$$J_k = \{1 \le i \le n : \exists j \in J_{k-1} \text{ s.t. } |M_{ji}| > 0\} \cup J_{k-1}, \qquad \forall k \ge 1;$$
$$D_k = \max_{i \in J_k}\{(\max_{1\le j\le n} \hat\varphi_j) - \hat\varphi_i\}.$$

For $i \in J_k \setminus J_{k-1}$, take $j \in J_{k-1}$ s.t. $|M_{ji}| > 0$, the equality on the j-th entry now reads

$$|M_{ji}|\hat\varphi_i + \sum_{k\ne j} |M_{jk}|\hat\varphi_j - \sum_{k\ne i,j} |M_{jk}|\hat\varphi_k = \psi_j.$$

By the fact that the summation of each row of M is zero, one can rewritten the above equation as

$$|M_{ji}|(\max_l \hat\varphi_l - \hat\varphi_i) = \sum_{k\ne j} |M_{jk}|(\max_l \hat\varphi_l - \hat\varphi_j) + \psi_j - \sum_{k\ne i,j} |M_{jk}|(\max_l \hat\varphi_l - \hat\varphi_k);$$

by which we have

$$D_k \le \Big(1 + (n-2)\frac{\lambda_1}{\lambda_0}\Big)D_{k-1} + \frac{1}{\lambda_0}\|\psi\|_{\ell^1}.$$

By definition, $D_0 = 0$. Hence by induction we obtain

$$D_k \le \frac{\left[1 + (n-2)\frac{\gamma_1}{\gamma_0}\right]^k - 1}{(n-2)\gamma_1} \|\varphi\|_{\ell^1}.$$

By the irreducibility of M, unless $J_k = \{1 \dots n\}$ we have $J_k \subsetneq J_{k+1}$ (i.e. J_k is a proper subset of J_k). This implies that $|J_{k+1}| \ge |J_k| + 1$ and since $|J_0| \ge 1$, it proves that that $J_{n-1} = \{1 \dots n\}$.

Therefore, by taking $C(n; \gamma_0; \gamma_1)$ and $C_2(n)$ such that

$$C(n; \gamma_0; \gamma_1) = C_2(n)^{(n-2)} \gamma_1^{(n-1)} \gamma_0 D_{n-1};$$

we have $(\max_l \lambda_l - \min_l \lambda_l) \le C(n; \gamma_0; \gamma_1)\|\varphi\|_{\ell^1}$. By $\sum_i \lambda_i = 0$ we have $\min_l \lambda_l \le 0 \le \max_l \lambda_l$. Hence $\|\hat{x}\|_{\ell^1} \le C(n; \gamma_0; \gamma_1)\|\varphi\|_{\ell^1}$.

When M is reducible, we know, up to a permutation, that M is block diagonal and each diagonal block $M(I_k; I_k)$ is irreducible. Applying the previous result to each $M(I_k; I_k)$ completes the proof.

6.4. Uniform boundedness. We are finally able to show our uniform boundedness result.

Theorem 6.6. Consider $M \in \mathcal{M}(n)$ and $\varphi \in \mathbb{R}^n$ satisfying that $\exists C_0 > 0$, $\forall I^0 \subset I = \{1; \dots; n\}$,

$$\sum_{j \in I^0} \varphi \le C_0 \sum_{j \in I^0; i \in I \setminus I^0} |M_{ij}| + |M_{ji}| : \qquad (6.5)$$

Then there exists $x \in \mathbb{R}^n$ satisfying

$$Mx = \varphi; \quad \text{and} \quad \|x\|_{\ell^1} \le C_0 C_1(n);$$

for some constant $C_1(n)$ depending only on the dimension n.

Assuming for the moment that this theorem is correct. We can then immediately derive Theorem 3.4.

Proof of Theorem 3.4. Combining Proposition 6.2 and 6.3, we see that the linear system (6.4) satisfies the condition in Theorem 6.6 with $n = |V_0|$, $M = (A - A^T)$, $C_0 = C(V_0)$. Moreover, since we assume $\bigcup_{i \in V_0} \operatorname{supp}_i$ being connected in Definition 1.6, it is easy to verify that for some constant C

$$C(V_0) = \sup_{x;y \in \bigcup_{i \in V_0} \operatorname{supp}_i} |x - y| \le C\, |V_0| \Delta x;$$

where Δx is the discretization size of the mesh.

Hence by applying Theorem 6.6 to (6.4), we conclude that there are solutions $\lambda(b_c)$ for all possible b_c, with ℓ^1 norm uniformly bounded by

$$\|\hat{x}\|_{\ell^1} \le C(V_0)C_1(|V_0|) \le C(|V_0|)\Delta x;$$

where $C(|V_0|)$ depends only on $|V_0|$; the number of cell functions in a period. We can have the solutions satisfy that $\forall \alpha > 0; \lambda_i(\alpha b_c) = \alpha \lambda_i(b_c)$ simply by redefining $\lambda_i(b_c) = |b_c| \lambda_i(b_c/|b_c|)$ for all b_c s.t. $|b_c| \ne 0$. Finally, by Lemma 6.1 we can extend our solution $\lambda(b_c)$ to a periodic solution on the entire mesh, satisfying all properties claimed in Theorem 3.4.

We now conclude the section with the proof of Theorem 6.6:

Proof of Theorem 6.6. Let us begin with the solvability of $Mx=\varphi$. Take the decomposition $I = I_1\cup\cdots\cup I_m$ as in Lemma 6.4. By Lemma 6.4,

$$\mathrm{Range}\, M = \Big\{x\in\mathbb{R}^I : \sum_{i\in I_k} x_i = 0, \forall 1\le k\le m\Big\}.$$

For all $k=1,\dots,m$, we have $M(I_k, I\setminus I_k) = M(I\setminus I_k, I_k) = 0$. By (6.5) this implies that $\sum_{I_k}\varphi = 0$, $\forall k=1,\dots,m$. Therefore $\varphi\in \mathrm{Range}\, M$ and $Mx=\varphi$ is solvable.

We now turn to the proof of ℓ^1 bound of x. We can WLOG assume $C_0 = 1$ and $\max_{i\neq j}|M_{ij}| = \lambda_1 = 1$, because the general case can be reduced to it by a scaling $(\lambda_1^{-1}M)(C_0^{-1}x) = (\lambda_1^{-1}C_0^{-1}\varphi)$.

The result is immediate when $n=1$ and we prove the cases $n\geq 2$ by induction. Assume the theorem holds for any $p\times p$ matrices with $p\leq (n-1)$, we are going to show that the theorem holds for $n\times n$ matrices.

First, since

$$\mathcal{M}_0(n) = \mathcal{M}(n)\cap\{M\in\mathbb{R}^{n\times n} : \max_{i\neq j}|M_{ij}| = 1\}$$

is compact, it suffices to show that there is a local bound. More explicitly, we are going to show that for any $M^{(0)}\in\mathcal{M}_0(n)$, there is an open neighborhood $U\ni M^{(0)}$, s.t. for all $M\in U\cap\mathcal{M}_0(n)$ and φ satisfying (6.5) with $C_0=1$, there is a constant $C = C(n, U\cap\mathcal{M}_0(n))$ s.t. one can take $x\in\mathbb{R}^n$ satisfying $Mx=\varphi$ and $\|x\|_{\ell^1}\leq C(n, U\cap\mathcal{M}_0(n))$. Then by compactness, we conclude immediately that there is a uniform bound $C_1(n) = C(n,\mathcal{M}_0(n))$.

For arbitrary $M^{(0)}\in\mathcal{M}_0(n)$, introduce the irreducible decomposition $I = I_1\cup\cdots\cup I_m$ as in Lemma 6.4. For $1\leq k\leq m$, define $E_k = \mathrm{span}\{1_{\{i\}}\}_{i\in I_k}$, $F_k = \{x\in E_k : \sum_i x_i = 0\}$. In addition, define $E_0 = \mathrm{span}\{1_{I_k}\}_{k=1}^m$, $F_0 = \{x\in E_0 : \sum_i x_i = 0\}$. Note that we have

$$\mathbb{R}^n = (E_1\oplus\cdots\oplus E_m) = (F_1\oplus\cdots\oplus F_m)\oplus E_0 = (F_1\oplus\cdots\oplus F_m)\oplus F_0\oplus\mathrm{span}\{1_I\}.$$

Since there are non-zero non-diagonal entries in $M^{(0)}$, there exists k such that $|I_k|>1$. Hence $m<n$.

Let

$$\delta_0 = \min\Big\{|M^{(0)}_{ij}| : i\neq j, |M^{(0)}_{ij}|\neq 0\Big\},$$
$$r = \frac{1}{2nC(n,\delta_0,1)(1+2\max_{1\leq p<n}C_1(p))},$$

where the constant $C(n,\delta_0,1)$ is as in Lemma 6.5.

Define the open set

$$U_r(M^{(0)}) = \Big\{M\in\mathbb{R}^{n\times n} : \max_{1\leq i,j\leq n}|M - M^{(0)}|_{ij} < r\Big\}.$$

Define also the following linear mappings,

$$P: \mathbb{R}^n \to E_0 = \mathrm{span}\{1_{I_k}\}_{k=1}^m \cong \mathbb{R}^m$$
$$x \mapsto \sum_{k=1}^m \Big(\frac{1}{|I_k|}\sum_{i\in I_k} x_i\Big) 1_{I_k};$$

i.e. we take the average on each I_k,

$$\begin{aligned} Q : \mathbb{R}^m &\to E_0 = \mathrm{span}\{1_{I_k}\}_{k=1}^m \subset \mathbb{R}^n \quad y \\ &\mapsto \sum_{k=1}^m y_k 1_{I_k}; \end{aligned}$$

i.e. we project on the canonical basis of E_0, and finally

$$\begin{aligned} W : \mathbb{R}^m &\to \mathbb{R}^m \\ (y_1, \dots, y_m) &\mapsto (|I_1| y_1, \dots, |I_m| y_m). \end{aligned}$$

For $x \in \mathbb{R}^n$ solving $Mx = \varphi$, introduce the decomposition

$$x = x_1 + x_2 + \lambda 1_I; \qquad x_1 \in (F_1 \oplus \cdots \oplus F_m); \; x_2 \in F_0.$$

Without loss of generality, we can assume $\lambda = 0$. We are going to discussion two situations.

Case 1: $\|x_2\|_{\ell^1} \geq 2 \max_{1 \leq p < n} C_1(p) \|x_1\|_{\ell^1}$. Since $F_0 \subset E_0$, x_2 belongs to the image of Q and as Q is trivially one-to-one from $\mathbb{R}^m$ to E_0, we can define $y_2 = Q^{-1}(x_2) \in \mathbb{R}^m$. Since $M(x_1 + x_2) = \varphi$, we have

$$(WQ^{-1}PMQ)(y_2) = (WQ^{-1}PM)(x_2) = (WQ^{-1}P)(\varphi) - (WQ^{-1}PM)(x_1). \tag{6.6}$$

Since

$$(WQ^{-1}PMQ)_{kl} = \sum_{i \in I_k} \sum_{j \in I_l} M_{ij};$$

it is easy to verify that $(WQ^{-1}PMQ) \in \mathcal{M}(m)$.

Our goal is to apply the induction argument to the $m \times m$ linear system (6.6). To apply the induction argument, we need to provide proper estimates on the terms in the right-hand side. For any $J' \subset J = \{1, \dots, m\}$, we have

$$\begin{aligned} &\sum_{k \in J'} \left((WQ^{-1}P)(\varphi) - (WQ^{-1}PM)(x_1) \right)_k \\ = &\sum_{k \in J'} \left((WQ^{-1}P)(\varphi) \right)_k - \sum_{k \in J'} \left((WQ^{-1}PM)(x_1) \right)_k \\ = &\; L_1 + L_2. \end{aligned}$$

By our assumption on φ,

$$\begin{aligned} |L_1| = \left| \sum_{k \in J'} \sum_{i \in I_k} \varphi_i \right| &\leq \sum_{k \in J'; l \in J \setminus J'} \sum_{i \in I_k; j \in I_l} \left(|M_{ij}| + |M_{ji}| \right) \\ &= \sum_{k \in J'; l \in J \setminus J'} \left((WQ^{-1}PMQ)_{kl} + (WQ^{-1}PMQ)_{lk} \right). \end{aligned}$$

On the other hand,

$$\begin{aligned} |L_2| = &\left| \sum_{k \in J'} \sum_{i \in I_k} \sum_{j \in I} M_{ij} (x_1)_j \right| = \left| \sum_{j \in I} \sum_{k \in J'} \sum_{i \in I_k} M_{ij} (x_1)_j \right| \\ \leq &\left| \sum_{l \in J'} \sum_{j \in I_l} \sum_{k \in J'} \sum_{i \in I_k} M_{ij} (x_1)_j \right| + \left| \sum_{l \in J \setminus J'} \sum_{j \in I_l} \sum_{k \in J'} \sum_{i \in I_k} M_{ij} (x_1)_j \right| = \\ &\left| \sum_{l \in J'} \sum_{j \in I_l} \sum_{k \in J \setminus J'} \sum_{i \in I_k} M_{ij} (x_1)_j \right| + \left| \sum_{l \in J \setminus J'} \sum_{j \in I_l} \sum_{k \in J'} \sum_{i \in I_k} M_{ij} (x_1)_j \right| \\ \leq &\sum_{k \in J'; l \in J \setminus J'} \sum_{i \in I_k; j \in I_l} \left(|M_{ij}| + |M_{ji}| \right) \|x_1\|_{\ell^1} \\ = &\sum_{k \in J'; l \in J \setminus J'} \left((WQ^{-1}PMQ)_{kl} + (WQ^{-1}PMQ)_{lk} \right) \|x_1\|_{\ell^1}. \end{aligned}$$

In conclusion we have

$$|L_1 + L_2| \leq (1 + \|x_1\|_{\ell^1}) \sum_{k \in J'; l \in J \setminus J'} \left((WQ^{-1}PMQ)_{kl} + (WQ^{-1}PMQ)_{lk} \right);$$

which satisfies necessary assumption for the induction argument with $C_0 = (1+\|x_1\|_{\ell^1})$. As a consequence, we have

$$\|x_2\|_{\ell^1} = \|y_2\|_{\ell^1} \le (1+\|x_1\|_{\ell^1})C_1(m) \le C_1(m) + \|x_2\|_{\ell^1}/2;$$

by using the relation between $\|x_1\|_{\ell^1}$ and $\|x_2\|_{\ell^1}$ assumed at the beginning of the case.

Hence, as claimed

$$\|x_2\|_{\ell^1} \le 2C_1(m) \le 2\max_{1\le p<n} C_1(p);$$

$$\|x\|_{\ell^1} \le \|x_1\|_{\ell^1} + \|x_2\|_{\ell^1} \le 1 + 2\max_{1\le p<n} C_1(p).$$

Case 2: $\|x_2\|_{\ell^1} < 2\max_{1\le p<n} C_1(p)\|x_1\|_{\ell^1}$. Since $M(x_1+x_2)=\varphi$, we have

$$M^{(0)}(x_1) = M^{(0)}(x_1+x_2) = M(x_1+x_2) + (M^{(0)}-M)(x_1+x_2) = \varphi + (M^{(0)}-M)(x_1+x_2).$$

On each $I_k\times I_k$ block ($k=1,\dots,m''$) we can apply Lemma 6.5. For $M\in U_r(M^{(0)})\cap\mathcal{M}(n)$, this gives

$$\|x_1\|_{\ell^1} \le C(n;\delta_0;1)\|\varphi\|_{\ell^1} + rd\Big(1+2\max_{1\le p<n} C_1(p)\Big)\|x_1\|_{\ell^1} = C(n;\delta_0;1)\|\varphi\|_{\ell^1} + \|x_1\|_{\ell^1}/2.$$

By (6.5), for all $i\in I$,

$$|\varphi_i| \le \sum_{j\in I\setminus\{i\}} |M_{ij}| + |M_{ji}| \le 2(n-1).$$

Hence

$$\|x_1\|_{\ell^1} \le 4(n-1)C(n;\delta_0;1);$$

$$\|x\|_{\ell^1} \le \|x_1\|_{\ell^1} + \|x_2\|_{\ell^1} \le 4(n-1)C(n;\delta_0;1)\Big(1+2\max_{1\le p<n} C_1(p)\Big).$$

This finishes the study of Case 2.

Summarizing the results from Case 1 and Case 2, for arbitrary $M^{(0)}\in\mathcal{M}_0(n)$ we have $U_r(M^{(0)})$ such that for all $M\in U_r(M^{(0)})$ and φ satisfying (6.5) with $C_0=1$, one can take $x\in\mathbb{R}^d$ satisfying

$$Mx=\varphi; \quad \text{and} \quad \|x\|_{\ell^1} \le 4(n-1)C(n;\delta_0;1)\Big(1+2\max_{1\le p<n} C_1(p)\Big).$$

Recall that $U_r(M^{(0)})$ is give by

$$r = \frac{1}{2nC(n;\delta_0;1)(1+2\max_{1\le p<n} C_1(p))};$$

$$U_r(M^{(0)}) = \Big\{M\in\mathbb{R}^{n\times n} : \max_{1\le i,j\le n} |M-M^{(0)}| < r\Big\}.$$

In conclusion, we can take

$$C\big(n; U_r(M^{(0)})\cap\mathcal{M}_0(n)\big) = 4(n-1)C(n;\delta_0;1)\Big(1+2\max_{1\le p<n} C_1(p)\Big).$$

We can conclude the proof by compactness.

We finish this section by explaining how one may be able to find a (very rough) upper bound in the previous proof. By Lemma 6.5, we have $C(n;\epsilon;1) = C_2(n)\,\epsilon^{-(n-1)}$ where $C_2(n)$ only depends on the dimension n. For any $M \in \mathcal{M}(n)$ let us define

$$\delta(M) = \min\{|M_{ij}| : i \neq j,\ |M_{ij}| \neq 0\};$$

$$r(M) = \frac{1}{2nC_2(n)[\delta(M)]^{-(n-1)}(1 + 2\max_{1 \le p<n} C_1(p))};$$

We are going to argue that for sufficiently small $\delta_{\min} = \delta_{\min}(n) > 0$ and arbitrary $M^{(0)} \in \mathcal{M}_0(n)$, there exists $M \in \mathcal{M}(n)$ such that $\delta(M) \ge \delta_{\min}$ and $M^{(0)} \in U_{r(M)}(M)$. Once this argument is proved and $\delta_{\min}$ is given explicitly, by the proof of Theorem 6.6, we have

$$C\left(n; U_{r(M)}(M) \cap \mathcal{M}_0(n)\right) \le 4(n-1)C_2(n)[\delta_{\min}(n)]^{-(n-1)}\left(1 + 2^{-1}\max_{p<n} C_1(p)\right):$$

Therefore, we can take

$$C_1(n) = 4(n-1)C_2(n)[\delta_{\min}(n)]^{-(n-1)}\left(1 + 2^{-1}\max_{p<n} C_1(p)\right);$$

which gives an explicit induction relation of $C_1(n)$.

We now describe how to find such M. Let us construct a sequence $\{M^{(k)}\} \subset \mathcal{M}(n)$ (starting with $M^{(0)}$) by the following iterations,

(1) If $\delta(M^{(k)}) \ge \delta_{\min}$, we stop the sequence. Otherwise, take

$$(i^{(k)}, j^{(k)}) \in \arg\min_{(i,j):\, i \neq j,\, |M^{(k)}_{ij}| \neq 0} |M^{(k)}_{ij}|:$$

(2) For $M^{(k)} \neq 0$, take the decomposition $I = I_1^{(k)} \cup \cdots \cup I_m^{(k)}$ as in Lemma 6.4. Assume WLOG that $i^{(k)}, j^{(k)} \in I_1^{(k)}$. By the equivalent definition of irreducible matrix that the associated directed graph is strongly connected, one can take a path $j^{(k)} = j_0^{(k)}, j_1^{(k)}, \ldots, j_p^{(k)} = i^{(k)}$ such that $M^{(k)}_{j_l^{(k)}, j_{l+1}^{(k)}} \neq 0$ for all $l = 1, \ldots, (p-1)$.

Let $P^{(k)}$ be the permutation matrix given by

$$P^{(k)}_{i,j} = \begin{cases} 1 & \text{if } (i,j) = (i^{(k)}, j^{(k)}) \\ & \text{or } (i,j) = (j_l^{(k)}, j_{l+1}^{(k)}),\ l = 1, \ldots, (p-1) \\ & \text{or } i = j \in I \setminus \{j_0^{(k)}, j_1^{(k)}, \ldots, j_p^{(k)}\} \\ 0 & \text{else:} \end{cases}$$

(3) Take $M^{(k+1)} = \left(M^{(k)} + \delta(M^{(k)})(P^{(k)} - I)\right)$.

For any $(i,j) \in I^2$ such that $i \neq j$ and $(P^{(k)} - I)_{ij} = 1$, we have $M^{(k)}_{i,j} < 0$. By the definition of $\delta(M^{(k)})$ it is easy to verify that the non-diagonal entries of $M^{(k+1)}$ given in this way are again non-positive. Hence $M^{(k+1)} \in \mathcal{M}(n)$ if $M^k \in \mathcal{M}(n)$. By induction, the above procedure produces a sequence $\{M^{(k)}\} \subset \mathcal{M}(n)$. We are going to argue that for sufficiently small $\delta_{\min}$, there must be an adequate candidate for M in $\{M^{(k)}\}$ before the sequence ends.

First, recall the definition $\delta(M) = \min\{|M_{ij}| : i \neq j; |M_{ij}| \neq 0\}$, it is straightforward to see that $M^{(k+1)}_{i^{(k)}, j^{(k)}} = 0$. Since there are only $n^2 - n$ non-zero non-diagonal entries and our process eliminates at least one entry at a time, the process must terminate somewhere before step $n^2 - n$.

Secondly, observe that $\|M^{(k+1)} - M^{(k)}\|_{\ell^1} \leq \delta(M^{(k)})$ and $|M^{(k+1)}_{i,j}| \geq |M^{(k)}_{i,j}|$, $\forall i \neq j$. By induction,

$$\|M^{(k+1)} - M^{(0)}\|_{\ell^1} \leq \sum_{l=0}^{k} \delta(M^{(l)}); \qquad \max_{i \neq j} |M^{(k+1)}_{ij}| \geq 1.$$

Thus, if

$$\sum_{l=0}^{k} \delta(M^{(l)}) \leq r(M^{(k+1)}) = \frac{1}{2nC_2(n)(\delta(M^{(k+1)}))^{-(n-1)}(1 + 2\max_{1 \leq p < n} C_1(p))};$$

or equivalently

$$\left(2nC_2(n)\left(1 + 2\max_{1 \leq p < n} C_1(p)\right) \sum_{l=0}^{k} \delta(M^{(l)})\right)^{1/(n-1)} \leq \delta(M^{(k+1)});$$

then it is guaranteed that $M^{(0)} \in U_{r(M^{(k+1)})}(M^{(k+1)})$.

To ensure that $M^{(k+1)}$ is an adequate candidate of M, we also need that $\delta(M^{(k+1)}) \geq \delta_{\min}$. Let us define the function

$$\phi(r) = \max\left\{\delta_{\min};\ \left(2nC_2(n)\left(1 + 2\max_{1 \leq p < n} C_1(p)\right) r\right)^{1/(n-1)}\right\}; \tag{6.7}$$

and reformulate what we just discussed as the following: If $\delta(M^{(k+1)}) \geq \phi(\sum_{l=0}^{k} \delta(M^{(l)}))$, we have that $\sum_{l=0}^{k} \delta(M^{(l)}) \leq r(M^{(k+1)})$ and $\delta(M^{(k+1)}) \geq \delta_{\min}$, so we can take $M = M^{(k+1)}$.

Let us assume that no candidate of M appears until step m. Then we should have $\delta(M^{(k)}) < \phi(\sum_{l=0}^{k-1} \delta(M^{(l)}))$ for $k = 1, \dots, m$. Define $\epsilon^{(0)} = \delta_{\min}$ and define $\epsilon^{(k)}$ inductively by

$$\epsilon^{(k)} = \phi\left(\sum_{l=0}^{k-1} \epsilon^{(l)}\right).$$

Obviously we have $\delta(M^{(0)}) \leq \epsilon^{(0)}$. Note that ϕ is increasing for any fixed $\delta_{\min} > 0$. Therefore we have $\delta(M^{(k)}) \leq \phi(\sum_{l=0}^{k-1} \delta(M^{(l)})) \leq \phi(\sum_{l=0}^{k-1} \epsilon^{(l)}) = \epsilon^{(k)}$ provided that $\delta(M^{(l)}) \leq \epsilon^{(l)}$ for all $l = 0, 1, \dots, (k-1)$. Applying the induction argument, we conclude that $\delta(M^{(k)}) \leq \epsilon^{(k)}$ for all $k = 0, 1, \dots, m$.

Let us discuss the growth of $\epsilon^{(k)}$. Note that for $r \to 0^+$ and $\delta_{\min} \to 0^+$ we have $\phi(r) \to 0^+$. Therefore $\lim_{\delta_{\min} \to 0} \epsilon^{(k)} = \lim_{\delta_{\min} \to 0} \phi(\sum_{l=0}^{k-1} \epsilon^{(l)}) = 0$ provided that $\lim_{\delta_{\min} \to 0} \epsilon^{(l)} = 0$ for all $l = 1, \dots, (k-1)$. Applying the induction argument, we conclude that $\lim_{\delta_{\min} \to 0} \epsilon^{(k)} = 0$ for all $k \geq 0$. In particular, by taking sufficiently small $\delta_{\min}$, one would have $\sum_{l=0}^{n^2-n} \epsilon^{(l)} < \frac{1}{2}$.

Since $\max_{i \neq j} |M^{(0)}_{ij}| = 1$, let us take i, j s.t. $|M^{(0)}_{ij}| = 1$. Observing that

$$|M^{(m+1)}_{ij}| \geq 1 - \sum_{l=0}^{m} \delta(M^{(l)}) \geq 1 - \sum_{l=0}^{m} \epsilon^{(l)}.$$

For $m < n^2 - n$, this implies that $|M_{1j}^{(m+1)}| > 0$, hence the iteration does not terminate at step $(m+1)$ either. By induction, unless there is an adequate candidate of M found, the iteration does not terminate before step $n^2 - n$. Recall that the iteration must terminate somewhere before step $n^2 - n$ as the non-zero entries are reducing, we conclude that an adequate candidate of M must appear somewhere before step $n^2 - n$.

7. Proof of remaining lemmas and propositions. In this section we collect the remaining missing proofs of various technical lemmas. Let us begin with Lemma 2.5.

Proof of Lemma 2.5. The equation (2.8) is equivalent to

$$\int_{\mathbb{R}^{2d}} \left|K_g^h(x;y) - K_f^h(x;y)\right| |u(x) - v(y)|^p \,dxdy \le C(h_1/h_0) \int_{\mathbb{R}^{2d}} K_f^h(x;y)|u(x) - v(y)|^p \,dxdy.$$

Notice that $\phi(x)$, yet $K^h(x)$, has compact support in the ball $B(0;2)$. Also, notice that

$$|f_1(x) - f_2(y)| \le |x - y| + |x - f_1(x)| + |y - f_2(y)| \le |x - y| + 2h_1;$$

where $h_1 \le 1/4$ by our assumption. Hence, $K_f^h(x;y)$ and $K_g^h(x;y)$ are non-zero only if $|x - y| \le 5/2$.

We further take the decomposition

$$\{(x;y) \in \mathbb{R}^{2d} : |x - y| \le 5/2\} = \{|x - y| \le 1/2\} \cup \{1/2 < |x - y| \le 5/2\};$$

by which we can rewrite the double integral as

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} \left|K_g^h(x;y) - K_f^h(x;y)\right| |u(x) - v(y)|^p \,dxdy \\ &= \left(\int_{|x-y| \le 1/2} + \int_{1/2 < |x-y| \le 5/2} \right) \left|K_g^h(x;y) - K_f^h(x;y)\right| |u(x) - v(y)|^p \,dxdy. \end{aligned}$$

On the set $\{|x - y| \le 1/2\}$, we have $|f_1(x) - f_2(y)| \le |x - y| + 2h_1 \le 1$. Hence

$$\chi(|f_1(x) - f_2(y)|) = \chi(|g_1(x) - g_2(y)|) = 1, \qquad \forall |x - y| \le 1/2.$$

Thus we have

$$\begin{aligned} &\int_{|x-y| \le 1/2} \left|K_g^h(x;y) - K_f^h(x;y)\right| |u(x) - v(y)|^p \,dxdy \\ &\le \sup_{|x-y| \le 1/2} \frac{\left|K_g^h(x;y) - K_f^h(x;y)\right|}{K_f^h(x;y)} \int_{\mathbb{R}^{2d}} K_f^h(x;y)|u(x) - v(y)|^p \,dxdy; \end{aligned}$$

where the coefficient before the integral can be bounded by

$$\sup_{|x-y| \le 1/2} \frac{\left|K_g^h(x;y) - K_f^h(x;y)\right|}{K_f^h(x;y)} = \sup_{|x-y| \le 1/2} \left| \frac{\frac{1}{h^d}\phi\left(\frac{|g_1(x) - g_2(y)|}{h}\right) - \frac{1}{h^d}\phi\left(\frac{|f_1(x) - f_2(y)|}{h}\right)}{\frac{1}{h^d}\phi\left(\frac{|f_1(x) - f_2(y)|}{h}\right)} \right| \le \frac{4h_1 d/h^{d+1}}{1/h^d} \le Ch_1/h_0.$$

Moreover, on the set $\{1/2 < |x-y| \le 5/2\}$,

$$
\begin{aligned}
&\int_{1/2<|x-y|\le 5/2} \left|K^h_g(x;y) - K^h_f(x;y)\right| |u(x)-v(y)|^p \,dx\,dy \\
&\le \int_{1/2<|x-y|\le 5/2} \left|K^h_g(x;y) - K^h_f(x;y)\right| 5^{p-1} \Big(|u(x) - v(4x/5+y/5)|^p \\
&\qquad + |v(4x/5+y/5) - u(3x/5+2y/5)|^p + |u(3x/5+2y/5) - v(2x/5+3y/5)|^p \\
&\qquad + |v(2x/5+3y/5) - u(x/5+4y/5)|^p + |u(x/5+4y/5) - v(y)|^p \Big)\,dx\,dy.
\end{aligned}
$$

As an example, let us look at the second term. With a change of variable

$$
w = 3x/5 + 2y/5, \quad z = x/5 - y/5,
$$

one has that

$$
\begin{aligned}
&5^{p-1}\int_{1/2<|x-y|\le 5/2} \left|K^h_g(x;y) - K^h_f(x;y)\right| |v(4x/5+y/5) - u(3x/5+2y/5)|^p \,dx\,dy \\
&= 5^{d+p-1} \int_{1/10<|z|\le 1/2} \left|K^h_g(w+2z;w-3z) - K^h_f(w+2z;w-3z)\right| |u(w) - v(w+z)| \,dw\,dz \\
&\le 5^{d+p-1} \sup_{w,z\in\mathbb{R}^d,\,1/10<|z|\le 1/2} \frac{\left|K^h_g(w+2z;w-3z) - K^h_f(w+2z;w-3z)\right|}{K^h_f(w;w+z)} \\
&\qquad \int_{\mathbb{R}^{2d}} K^h_f(w;w+z)|u(w) - v(w+z)| \,dw\,dz,
\end{aligned}
$$

where the coefficient before the integral can be bounded by

$$
\begin{aligned}
&\sup_{w,z\in\mathbb{R}^d,\,1/10<|z|\le 1/2} \frac{\left|K^h_g(w+2z;w-3z) - K^h_f(w+2z;w-3z)\right|}{K^h_f(w;w+z)} \\
&= \sup_{w,z\in\mathbb{R}^d,\,1/10<|z|\le 1/2} \frac{\left|K^h_g(w+2z;w-3z) - K^h_f(w+2z;w-3z)\right|}{\frac{1}{(|f_1(w)-f_2(w+z)|+h)^d}} \\
&\le C h_1 \le C h_1/h_0.
\end{aligned}
$$

The other four terms can be bounded by the same approach.

In conclusion, we have

$$
\begin{aligned}
&\int_{\mathbb{R}^{2d}} K^h_g(x;y)|u(x)-v(y)|\,dx\,dy \\
&\le \int_{\mathbb{R}^{2d}} \left|K^h_g(x;y) - K^h_f(x;y)\right| |u(x)-v(y)|\,dx\,dy + \int_{\mathbb{R}^{2d}} K^h_f(x;y)|u(x)-v(y)|\,dx\,dy \\
&\le (1 + C h_1/h_0) \int K^h_f(x;y)|u(x)-v(y)|\,dx\,dy.
\end{aligned}
\tag{7.1}
$$

Next we prove Lemma 2.6.

Proof of Lemma 2.6. We rst choose measurable sets $(V_i)_{i 2 V}$ R^d by the following. Divide R^d into small hypercubes

$$Q_{[m]} = \prod_{k=1}^{d} h \; n_k = \sqrt{d}x; \; (n_k + 1) = \sqrt{d}x; \qquad 8[m] = (n_1; : : : ; n_d) \, 2 \, Z^d:$$

where each hypercube has diameter x.

Then, in each hypercube $Q_{[m]}$ choose measurable sets $V_{i;[m]}$, $i \; 2 \; V$ satisfying

$$V_{i;[m]} \;\; Q_{[m]}; \qquad jV_{i;[m]}j = \int_{Q_{[m]}} {}_i; \qquad 8i \; 2 \; V; \quad \text{and} \quad V_{i;[m]} \setminus V_{j;[m]} \quad 8i; j \; 2 \; V;$$

which is always possible since $\sum_{i 2 V} {}_i$ 1 by our denition.

Choose $V_i = \bigcup_{[m]} V_{i;[m]}$ so that

$$jV_ij = {}_i = \int_{R^d} {}_i; \qquad \sup_{x 2 V_i} jx \quad x_i j < 2x; \qquad 8i \; 2 \; V; \quad \text{and} \quad V_i \setminus V_j \quad 8i; j \; 2 \; V:$$

Moreover, recall our assumption (2.6) that $\sum_{i 2 V} {}_i(x) = 1$;$8x$ 2 + $B(0;4)$. For any hypercube $Q_{[m]}$ + $B(0;4)$, one has $\sum_{i 2 V} jV_{i;[m]}j = jQ_{[m]}j$. Since we have assumed x 1=16, it is easy to verify that $Q_{[m]} \setminus$ + $B(0;3)$ = ? implies $Q_{[m]}$ + $B(0;4)$.

Then up to modication on a negligible set, one has that

$$\bigcup_{i 2 V} V_i \setminus \bigcup + B(0;3) = \bigcup_{i 2 V} \bigcup_{[m] 2 Z^d} V_{i;[m]} \setminus + B(0;3)$$
$$= \bigcup_{[m] 2 Z^d} \bigcup_{i 2 V} V_{i;[m]} \setminus + B(0;3)$$
$$= + B(0;3):$$

Recall our choice of piecewise constant extension $u^V = \sum_{i 2 V} u_i 1_{V_i}$ and

$$f(x) = \begin{cases} x_i; & \text{for } x \; 2 \; V_i; \; i \; 2 \; V; \\ x; & \text{for } x \; 2 \; \bigcup_{i 2 V} V_i: \end{cases}$$

Notice that u_i = 0 only if $\operatorname{supp}{}_i$ and we have assumed x 1=16. Hence the extended function u^V satises $\operatorname{supp} u^V$ + $B(0;1)$. Also, by our assumption it is straightforward that

$$\sup_{x 2 R^d} jx \quad f(x)j \quad \sup_{i 2 V} \sup_{x 2 V_i} jx \quad x_i j + \sup_{i 2 V} jx_i \quad x_i j \quad 2x + h_2 < 1=4:$$

Finally, let us consider the integral

$$\int_{R^{2d}} K^h \; f(x) \quad f(y) j u^V(x) \quad u^V(y) j^p \, dxdy:$$

We have proved $\operatorname{supp} u^V$ + $B(0;1)$ and by denition $\operatorname{supp} K^h$ 2 $B(0;2)$. Therefore, for x 2 +$B(0;3)$, either y 2 +$B(0;1)$, making $ju^V(x) \quad u^V(y)j = 0$, or y 2 +$B(0;1)$, making $K^h(x \quad y) = 0$.

The same argument also applies to y and as a consequence, the above integral can be taken instead over any subset of R^{2d} including + $B(0;3)$ ^{2}. In particular, it can be reformulated as

$$\int_{(\bigcup_{i 2 V} {}_i)^2} K^h \; f(x) \quad f(y) j u^V(x) \quad u^V(y) j \, dxdy = \sum_{i;j 2 V} K^h_{i;j} j u_i \quad u_j j^p {}_{ij};$$

which completes the proof.

We are now ready to prove Lemma 2.4, Proposition 2.7 and Proposition 2.8. Let us start with Proposition 2.7, which is immediate:

Proof of Proposition 2.7. By Lemma 2.6, choose measurable sets $(V_i)_{i\in V}\subset\mathbb{R}^d$, take piecewise constant extension $u^V=\sum_{i\in V}u_i 1_{V_i}$ and take $f_1^{(1)},f_2^{(1)},f_1^{(2)},f_2^{(2)}:\mathbb{R}^d\to\mathbb{R}^d$ by

$$\text{for } k=1,2;\quad f_1^{(k)}(x)=f_2^{(k)}(x)=\begin{cases}\tilde{x}_i^{(k)}, & \text{for } x\in V_i,\ i\in V;\\ x, & \text{for } x\notin\bigcup_{i\in V}V_i.\end{cases}$$

Then it is straightforward that

$$\begin{aligned}\|u\|_{h_0,p,\chi,\tilde{x}^{(k)}} &= \sup_{h_0\le h\le 1/2}|\log h|^{-1}\sum_{i,j\in V}\tilde{K}^h\big(\tilde{x}_i^{(k)}-\tilde{x}_j^{(k)}\big)|u_i-u_j|^p\chi_{ij}\\ &= \sup_{h_0\le h\le 1/2}|\log h|^{-1}\int_{\mathbb{R}^{2d}}K^h\big(f_1^{(k)}(x)-f_2^{(k)}(y)\big)|u^V(x)-u^V(y)|^p\,dx\,dy;\end{aligned}$$

and

$$\text{for } k,l=1,2;\quad \sup_{x\in\mathbb{R}^d}|x-f_l^{(k)}(x)|\le 2\delta x+h_2\le 3h_2.$$

We now apply Lemma 2.5 with $h_1=3h_2<1/4$, which gives that

$$\|u\|^p_{h_0,p,\chi,\tilde{x}^{(2)}}\le(1+Ch_2/h_0)\,\|u\|^p_{h_0,p,\chi,\tilde{x}^{(1)}}.$$

Noticing $(1+x)^{1/p}\le 1+x$ for all $x\ge 0$, we conclude that

$$\|u\|_{h_0,p,\chi,\tilde{x}^{(2)}}\le(1+Ch_2/h_0)\,\|u\|_{h_0,p,\chi,\tilde{x}^{(1)}},$$

which finishes the proof.

Next, let us prove Proposition 2.8.

Proof of Proposition 2.8. Choose any labeling of the index set V, and define

$$\begin{aligned} J:\ \mathbb{R}^d\times[0,1]&\to V\\ (y,\omega)&\mapsto\min\{i\in V:\textstyle\sum_{j\le i}\psi_j(y)\ge\omega\}.\end{aligned}$$

Notice that there are only a bounded number of nonzero $\psi_i(y)$ at any point y by our assumption.

Define also

$$\begin{aligned}\tilde{F}:\ V&\to\mathbb{R}^d\\ i&\mapsto x_i\end{aligned}$$

Then for all $y\in\mathbb{R}^d$, $\omega\in[0,1]$,

$$|(\tilde{F}\circ J)(y,\omega)-y|\le 2\delta x;\qquad u(y)=\sum_{i\in V}u_i\psi_i(y)=\int_0^1 u_{J(y,\omega)}\,d\omega.$$

By Lemma 2.5, for all $\omega_1,\omega_2\in[0,1]$

$$\begin{aligned}&\int_{\mathbb{R}^{2d}}K^h(x,y)|u_{J(x,\omega_1)}-u_{J(y,\omega_2)}|^p\,dx\,dy\\ &\le(1+C\delta x/h)\int_{\mathbb{R}^{2d}}K^h\big((\tilde{F}\circ J)(x,\omega_1),(\tilde{F}\circ J)(y,\omega_2)\big)|u_{J(x,\omega_1)}-u_{J(y,\omega_2)}|^p\,dx\,dy.\end{aligned}$$

Therefore, notice that $K^h\left(F(i);F(j)\right) = K^h\left(x_i;x_j\right) = K_{i;j}^h$, which implies that

$$
\begin{aligned}
&\int_{R^{2d}} K^h(x;y)|u(x)-u(y)|^p dxdy \\
&= \int_{R^{2d}} K^h(x;y)\left|\int_0^1 u_{J(x;\omega_1)}\, d\omega_1 - \int_0^1 u_{J(y;\omega_2)}\, d\omega_2\right|^p dxdy \\
&\le \int_{[0;1]^2}\int_{R^{2d}} K^h(x;y)\left|u_{J(x;\omega_1)} - u_{J(y;\omega_2)}\right|^p dxdyd\omega_1 d\omega_2 \\
&\le (1+Cx/h)\int_{[0;1]^2}\int_{R^{2d}} K^h\left((F\circ J)(x;\omega_1);(F\circ J)(y;\omega_2)\right)\left|u_{J(x;\omega_1)} - u_{J(y;\omega_2)}\right|^p dxdyd\omega_1 d\omega_2 \\
&= (1+Cx/h)\sum_{i;j\in V} K_{i;j}^h|u_i-u_j|^p{}_{ij}:
\end{aligned}
$$

Again by Lemma 2.5,

$$
\begin{aligned}
&\sum_{i;j\in V} K_{i;j}^h\left|(P_C u)_i - (P_C u)_j\right|^p{}_{ij} \\
&= \sum_{i;j\in V} K_{i;j}^h\left|\frac{1}{{}_i}\int_{R^d} u(x){}_i(x)\, dx - \frac{1}{{}_j}\int_{R^d} u(y){}_j(y)\, dy\right|^p{}_{ij} \\
&\le \sum_{i;j\in V} \int_{R^{2d}} K_{i;j}^h\left|u(x)-u(y)\right|^p{}_i(x){}_j(y)dxdy \\
&= \int_{R^{2d}} K^h\left((F\circ J)(x;\omega_1);(F\circ J)(y;\omega_2)\right)\left|u(x)-u(y)\right|^p dxdyd\omega_1 d\omega_2 \\
&\le (1+Cx/h)\int_{R^{2d}} K^h(x;y)|u(x)-u(y)|\, dxdy;
\end{aligned}
$$

concluding the proof.

Lemma 2.4 can then be derived from Proposition 2.8.

Proof of Lemma 2.4. Since we have assumed $h > h_0 > x$ and $\|(u_i)_{i\in V}\|_{h_0;p;} \le L$, by Proposition 2.8,

$$
\int_{R^{2d}} K^h(x;y)|u(x)-u(y)|^p dxdy \le (1+Cx/h)\sum_{i;j\in V} K^h_{i;j}|u_i-u_j|^p{}_{ij} \le C|\log h|;
$$

where the constant C may depend on L.

Introduce the renormalization factor

$$C_h = |\log|/\|K^h\|_{L^1}:$$

Then C_h is bounded form above and below uniformly with respect to h, and the renormalized kernel K^h reads

$$K^h(x) = K^h(x)/\|K^h\|_{L^1} = C_h|\log|^{-1}K^h(x):$$

Hence

$$\|u-K_h\star u\|_{L^p}^p = (C_h|\log h|^{-1})^p\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}K_h(x-y)\,(u(x)-u(y))\,dy\right|^p dx$$
$$\le C|\log h|^{-p}\|K_h\|_{L^1}^{p-1}\int_{\mathbb{R}^{2d}}K_h(x-y)|u(x)-u(y)|^p\,dy\,dx$$
$$\le C|\log h|^{-1}\int_{\mathbb{R}^{2d}}K_h(x-y)|u(x)-u(y)|^p\,dy\,dx.$$

This implies

$$\|u-K_h\star u\|_{L^p}^p\le C|\log h|^{-1} \tag{7.2}$$

We have finished the proofs of all lemmas and propositions in Section 2.2 but it remains to prove Proposition 2.10 as claimed in the proof of Theorem 1.3.

Proof of Proposition 2.10. Let $(C;F)$ be a mesh as in Definition 1.5 over $\mathbb{R}^d$ such that (1.14) hold. Assume that each face function $n_{i,j}\in F$ is of form $n_{i,j}(x)=N_{i,j}w_{i,j}(x)$, $\forall x\in\mathbb{R}^d$, where $N_{i,j}$ is a unit vector and $w_{i,j}$ is a scalar function.

Then for $1\le p\le\infty$,

$$\|P_F^0 b-P_F b\|_{L^p([0,T]\times F)}\le C\|b\|_{L^p_t(W^{1,p}_x)},$$

where the constant only depends on p and the constant in the structural assumption (1.14).

We are going to first prove the inequality for any fixed time t and To simplify the notation we omit t in all the calculations.

By definition, we have

$$P_F b-P_F^0 b_{i,j}=\int_{\mathbb{R}^d}\left(b(x)\cdot N_{i,j}\right)^+ w_{i,j}(x)\,dx-\left(\int_{\mathbb{R}^d}b(x)\cdot N_{i,j}w_{i,j}(x)\,dx\right)^+.$$

We introduce the more general function

$$I(b;N;w)=\int_{\mathbb{R}^d}\left(b(x)\cdot N\right)^+w(x)\,dx-\left(\int_{\mathbb{R}^d}b(x)\cdot N\,w(x)\,dx\right)^+$$

where $N\in\mathbb{R}^d$ is a unit vector and w is a non-negative, bounded function with compact support. It is straightforward that $I(b;N;w)\ge 0$ and for any two functions $v,w\ge 0$,

$$\begin{aligned}
&I(b;N;w)+I(b;N;v)\\
&=\int_{\mathbb{R}^d}\left(b(x)\cdot N\right)^+\left(w(x)+v(x)\right)dx-\left(\int_{\mathbb{R}^d}b(x)\cdot N\,w(x)\,dx\right)^+-\left(\int_{\mathbb{R}^d}b(x)\cdot N\,v(x)\,dx\right)^+\\
&\le\int_{\mathbb{R}^d}\left(b(x)\cdot N\right)^+\left(w(x)+v(x)\right)dx-\left(\int_{\mathbb{R}^d}b(x)\cdot N\left(w(x)+v(x)\right)dx\right)^+\\
&=I(b;N;w+v).
\end{aligned}$$

Hence if $0\le w\le u$, then $I(b;N;w)\le I(b;N;u)$.

Moreover, $I(b;N;w)$ is directly bounded by the following inequality

$$I(b;N;w)\le\sup\left\{\int_{\mathbb{R}^d}\left(b(x)\cdot N\right)^+w(x)\,dx,\ \int_{\mathbb{R}^d}\left(b(x)\cdot N\right)w(x)\,dx\right\}^+.$$

And it is easy to verify that (in distributional sense)

$$\frac{\partial}{\partial\lambda}\left\{\int_{\mathbb{R}^d}\big(b(x)-N-\lambda\big)^+ w(x)\,dx-\left(\int_{\mathbb{R}^d}\big(b(x)-N-\lambda\big)\,w(x)\,dx\right)^+\right\}$$

$$=\int_{\mathbb{R}^d}-\mathbb{1}_{\{b(x)-N-\lambda\ge 0\}}\,w(x)\,dx+\mathbb{1}_{\left\{\int_{\mathbb{R}^d}(b(x)-N-\lambda)\,w(x)\,dx\ge 0\right\}}\int_{\mathbb{R}^d}w(x)\,dx.$$

Thus the maximum is attained at $\lambda=$

$$\int_{\mathbb{R}^d}\big(b(x)-N\big)\,w(x)\,dx\Big/\int_{\mathbb{R}^d}w(x)\,dx\,;$$

thus

$$\begin{aligned} I(b;N;w)&\le\int_{\mathbb{R}^d}\big(b(x)-N-\lambda\big)^+ w(x)\,dx\\ &\le\frac{1}{\|w\|_{L^1}}\int_{\mathbb{R}^{2d}}|b(x)-b(y)|\,w(x)w(y)\,dxdy. \end{aligned}$$

Notice that $w_{i,j}\ge C\,\delta x^{-1}\mathbb{1}_{B(x_i,\delta x)}$ by our structural assumptions (1.14). Therefore,

$$\begin{aligned} \left|P_F b-P_F^0 b\right|_{i,j}&=I(b;N_{i,j};w_{i,j})\le I\big(b;N_{i,j};C(\delta x)^{-1}\mathbb{1}_{B(x_i,\delta x)}\big)\\ &\le\frac{C(\delta x)^{-1}}{\|B(x_i,\delta x)\|_{L^1}}\int_{\mathbb{R}^{2d}}|b(x)-b(y)|\,\mathbb{1}_{B(x_i,\delta x)}(x)\mathbb{1}_{B(x_i,\delta x)}(y)\,dxdy. \end{aligned}$$

When $p=\infty$, we have

$$\begin{aligned} \left\|P_F b-P_F^0 b\right\|_{L^\infty(F)}&=\sup_{i,j\in V}\left|P_F b-P_F^0 b\right|_{i,j}(\delta x)^{-(d-1)}\\ &\le C(\delta x)^{d-1}\,\delta x\,\|b\|_{W^{1,\infty}}(\delta x)^{-(d-1)}=\\ &\quad C\delta x\|b\|_{W^{1,\infty}}. \end{aligned}$$

When $p=1$, we have

$$\begin{aligned} \left\|P_F b-P_F^0 b\right\|_{L^1(F)}&=\sum_{i,j\in V}\left|P_F b-P_F^0 b\right|_{i,j}\delta x=\sum_{(i,j)\in E}\left|P_F b-P_F^0 b\right|_{i,j}\delta x\\ &=\sum_{(i,j)\in E}\frac{C}{\|B(x_i,\delta x)\|_{L^1}}\int_{\mathbb{R}^{2d}}|b(x)-b(y)|\,\mathbb{1}_{B(x_i,\delta x)}(x)\mathbb{1}_{B(x_i,\delta x)}(y)\,dxdy\\ &\le\sum_{(i,j)\in E}\frac{C}{\|B(x_i,\delta x)\|_{L^1}}\int_{x\in B(x_i,\delta x)}\int_{|z|\le 2\delta x}|b(x)-b(x+z)|\,dxdy\\ &\le\frac{C}{\|B(x_i,\delta x)\|_{L^1}}\int_{x\in\mathbb{R}^d}\int_{|z|\le 2\delta x}|b(x)-b(x+z)|\,dxdy\\ &\le C\delta x\|b\|_{W^{1,1}}. \end{aligned}$$

An interpolation completes the case $1\le p\le\infty$, i.e.

$$\left\|P_F b-P_F^0 b\right\|_{L^p(F)}\le C\delta x\|b\|_{W^{1,p}}.$$

Integrating now over time, we conclude that

$$\left\|P_F\bar b-P_F b\right\|_{L^p([0,T]\times F)}\le C\delta x\|b\|_{L^p_t(W^{1,p}_x)}.$$

References

[1] Giovanni Alberti, Gianluca Crippa, and Anna L Mazzucato, Exponential self-similar mixing and loss of regularity for continuity equations, Comptes rendus mathematique 352 (2014), no. 11, 901{906.

[2] Luigi Ambrosio, Transport equation and Cauchy problem for BV vector elds, Inventiones mathematicae 158 (2004), no. 2, 227{260.

[3] Luigi Ambrosio, Myriam Lecumberry, and Stefania Maniglia, Lipschitz regularity and approximate dierentiability of the DiPerna-Lions ow, Rendiconti del Seminario Matematico della Universita di Padova 114 (2005), 29{50.

[4] F Ben Belgacem and P-E Jabin, Convergence of numerical approximations to non-linear continuity equations with rough force elds, Archive for Rational Mechanics and Analysis 234 (2019), no. 2, 509{547.

[5] Fethi Ben Belgacem and Pierre-Emmanuel Jabin, Compactness for nonlinear continuity equations, Journal of Functional Analysis 264 (2013), no. 1, 139{168.

[6] Anna Bohun, Francois Bouchut, and Gianluca Crippa, Lagrangian ows for vector elds with anisotropic regularity, Annales de l'Institut Henri Poincare C, Analyse non lineaire 33 (2016), no. 6, 1409{1429.

[7] Franois Bouchut, Renormalized solutions to the Vlasov equation with coecients of bounded variation, Archive for Rational Mechanics and Analysis 157 (2001), no. 1, 75{90.

[8] Franois Bouchut and Gianluca Crippa, Lagrangian ows for vector elds with gradient given by a singular integral, Journal of hyperbolic dierential equations 10 (2013), no. 02, 235{282.

[9] Jean Bourgain, Haim Brezis, and Petru Mironescu, Another look at Sobolev spaces, Optimal Control and Partial Dierential Equations. In honour of Professor Alain Bensoussan's 60th birthday, 2000, pp. 439{455.

[10] Franck Boyer, Analysis of the upwind nite volume method for general initial- and boundary-value transport problems, IMA Journal of Numerical Analysis 32 (2012), no. 4, 1404{1439.

[11] Yann Brenier, Felix Otto, and Christian Seis, Upper bounds on coarsening rates in demixing binary viscous liquids, SIAM journal on mathematical analysis 43 (2011), no. 1, 114{134.

[12] Susanne C Brenner, L Ridgway Scott, and L Ridgway Scott, The mathematical theory of nite element methods, vol. 3, Springer, 2008.

[13] Didier Bresch and Pierre-Emmanuel Jabin, Global weak solutions of PDEs for compressible media: A compactness criterion to cover new physical situations, Shocks, singularities and oscillations in nonlinear optics and uid mechanics, Springer, 2017, pp. 33{54.

[14] Didier Bresch and Pierre-Emmanuel Jabin, Global existence of weak solutions for compressible Navier{Stokes equations: thermodynamically unstable pressure and anisotropic viscous stress tensor, Annals of Mathematics 188 (2018), no. 2, 577{684.

[15] Elia Brue and Quoc-Hung Nguyen, Sharp regularity estimates for solutions of the continuity equation drifted by Sobolev vector elds, arXiv preprint arXiv:1806.03466 (2018).

[16] Martin Burger, Yasmin Dolak-Struss, and Christian Schmeiser, Asymptotic analysis of an

advection-dominated chemotaxis model in multiple spatial dimensions, Communications in Mathematical Sci-ences 6 (2008), no. 1, 1{28.

[17] Nicolas Champagnat and Pierre-Emmanuel Jabin, Well posedness in any dimension for Hamilton-ian ows with non BV force terms, Communications in Partial Dierential Equations 35 (2010), no. 5, 786{816.

[18] ______, Strong solutions to stochastic dierential equations with rough coecients, The Annals of Probability 46 (2018), no. 3, 1498{1541.

[19] Maria Colombo, Gianluca Crippa, and Stefano Spirito, Renormalized solutions to the continuity equation with an integrable damping term, Calculus of Variations and Partial Dierential Equations 54 (2015), no. 2, 1831{1845.

[20] G Crippa and C De Lellis, Estimates and regularity results for the DiPerna-Lions ow, Journal fur die Reine und Angewandte Mathematik 616 (2008), 15{46.

[21] Franois Delarue and Frederic Lagoutiere, Probabilistic analysis of the upwind scheme for transport equations, Archive for rational mechanics and analysis 199 (2011), no. 1, 229{268.

[22] Franois Delarue, Frederic Lagoutiere, and Nicolas Vauchelet, Convergence order of upwind type schemes for transport equations with discontinuous coecients, Journal de Mathematiques Pures et Appliquees 108 (2017), no. 6, 918{951.

[23] Ronald J DiPerna and Pierre-Louis Lions, Ordinary dierential equations, transport theory and Sobolev spaces, Inventiones mathematicae 98 (1989), no. 3, 511{547.
[24] Robert Eymard, Thierry Gallouet, Raphaele Herbin, and J Latche, A convergent nite element-nite volume scheme for the compressible Stokes problem. Part II: the isentropic case, Mathematics of Computation 79 (2010), no. 270, 649{675.
[25] Robert Eymard, Thierry Gallouet, Raphaele Herbin, and J-C Latche, Convergence of the MAC scheme for the compressible Stokes equations, SIAM Journal on Numerical Analysis 48 (2010), no. 6, 2218{2246.
[26] Shizan Fang, Dejun Luo, and Anton Thalmaier, Stochastic dierential equations with coecients in Sobolev spaces, Journal of Functional Analysis 259 (2010), no. 5, 1129{1168.
[27] T Gallouet, R Herbin, and J-C Latche, Kinetic energy control in explicit nite volume discretiza-tions of the incompressible and compressible Navier-Stokes equations, International Journal on Finite Volumes 7 (2010), no. 2, 1{6.
[28] Thierry Gallouet, Laura Gastaldo, Raphaele Herbin, and Jean-Claude Latche, An unconditionally stable pressure correction scheme for the compressible barotropic Navier-Stokes equations, ESAIM: Mathematical Modelling and Numerical Analysis 42 (2008), no. 2, 303{331.
[29] Thierry Gallouet, Raphaele Herbin, and J-C Latche, A convergent nite element-nite volume scheme for the compressible Stokes problem. Part I: The isothermal case, Mathematics of Compu-tation 78 (2009), no. 267, 1333{1352.
[30] Thierry Goudon, Julie Llobell, and Sebastian Minjeaud, An asymptotic preserving scheme on stag-gered grids for the barotropic Euler system in low Mach regimes, Numerical Methods for Partial Dierential Equations 36 (2020), no. 5, 1098{1128.
[31] Mahir Hadzic, Andreas Seeger, Charles K Smart, and Brian Street, Singular integrals and a problem on mixing ows, Annales de l'Institut Henri Poincare C 35 (2018), no. 4, 921{943.
[32] Gautam Iyer, Alexander Kiselev, and Xiaoqian Xu, Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained ows, Nonlinearity 27 (2014), no. 5, 973.
[33] P-E Jabin and Nader Masmoudi, DiPerna{Lions ow for relativistic particles in an electromagnetic eld, Archive for Rational Mechanics and Analysis 217 (2015), no. 3, 1029{1067.
[34] Pierre-Emmanuel Jabin, Critical non-Sobolev regularity for continuity equations with rough velocity elds, Journal of Dierential Equations 260 (2016), no. 5, 4739{4757.
[35] Flavien Leger, A new approach to bounds on mixing, Mathematical Models and Methods in Applied Sciences 28 (2018), no. 05, 829{849.
[36] Benoit Merlet, L^1- and L^2-error estimates for a nite volume approximation of linear advection, SIAM journal on numerical analysis 46 (2008), no. 1, 124{150.
[37] Benoit Merlet and Julien Vovelle, Error estimate for nite volume scheme, Numerische Mathematik 106 (2007), no. 1, 129{155.
[38] Benoit Perthame and Anne-Laure Dalibard, Existence of solutions of the hyperbolic Keller-Segel model, Transactions of the American Mathematical Society 361 (2009), no. 5, 2319{2335.
[39] Andre Schlichting and Christian Seis, Convergence rates for upwind schemes with rough coecients, SIAM Journal on Numerical Analysis 55 (2017), no. 2, 812{840.
[40] Christian Seis, Maximal mixing by incompressible uid ows, Nonlinearity 26 (2013), no. 12, 3279.
[41] ______, A quantitative theory for the continuity equation, Annales de l'Institut Henri

Poincare C, Analyse non lineaire, vol. 34, Elsevier, 2017, pp. 1837–1850.

[42] , Optimal stability estimates for continuity equations, Proceedings of the Royal Society of Edinburgh Section A: Mathematics 148 (2018), no. 6, 1279–1296.

[43] Chad M Topaz and Andrea L Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, SIAM Journal on Applied Mathematics 65 (2004), no. 1, 152–174.

[44] Xicheng Zhang, Stochastic flows of SDEs with irregular coefficients and stochastic transport equations, Bulletin des sciences mathematiques 134 (2010), no. 4, 340–378.