



---

*Research article*

## On regularity and existence of weak solutions to nonlinear Kolmogorov-Fokker-Planck type equations with rough coefficients<sup>†</sup>

Prashanta Garain and Kaj Nyström\*

Department of Mathematics, Uppsala University, S-751 06 Uppsala, Sweden

<sup>†</sup> **This contribution is part of the Special Issue: PDEs and Calculus of Variations—Dedicated to Giuseppe Mingione, on the occasion of his 50th birthday**  
Guest Editors: Giampiero Palatucci; Paolo Baroni  
Link: [www.aimspress.com/mine/article/6240/special-articles](http://www.aimspress.com/mine/article/6240/special-articles)

\* **Correspondence:** Email: [kaj.nystrom@math.uu.se](mailto:kaj.nystrom@math.uu.se).

**Abstract:** We consider nonlinear Kolmogorov-Fokker-Planck type equations of the form

$$(\partial_t + X \cdot \nabla_Y)u = \nabla_X \cdot (A(\nabla_X u, X, Y, t)).$$

The function  $A = A(\xi, X, Y, t) : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$  is assumed to be continuous with respect to  $\xi$ , and measurable with respect to  $X, Y$  and  $t$ .  $A = A(\xi, X, Y, t)$  is allowed to be nonlinear but with linear growth. We establish higher integrability and local boundedness of weak sub-solutions, weak Harnack and Harnack inequalities, and Hölder continuity with quantitative estimates. In addition we establish existence and uniqueness of weak solutions to a Dirichlet problem in certain bounded  $X, Y$  and  $t$  dependent domains.

**Keywords:** Kolmogorov equation; parabolic; ultraparabolic; hypoelliptic; nonlinear Kolmogorov-Fokker-Planck equations; existence; uniqueness; regularity

**Mathematics Subject Classification:** 35H20, 35K65, 35K70, 35R03

---

### 1. Introduction and statement of main results

Several important evolution equations arising in kinetic theory, mathematical physics and probability can be written in the form

$$(\partial_t + X \cdot \nabla_Y)f = Q(f, \nabla_X f, X, Y, t), \tag{1.1}$$

where  $(X, Y, t) := (x_1, \dots, x_m, y_1, \dots, y_m, t) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} = \mathbb{R}^{N+1}$ ,  $N = 2m$ ,  $m \geq 1$ , and the coordinates  $X = (x_1, \dots, x_m)$  and  $Y = (y_1, \dots, y_m)$  are, respectively, the velocity and the position of the system. In its simplest form,

$$Q(f, \nabla_X f, X, Y, t) = \nabla_X \cdot \nabla_X f = \Delta_X f,$$

the equation in (1.1) was introduced and studied by Kolmogorov in a famous note published in 1934 in *Annals of Mathematics*, see [25]. In this case Kolmogorov noted that the equation in (1.1) is an example of a degenerate parabolic operator having strong regularity properties and he proved that the equation has a fundamental solution which is smooth off its diagonal. In fact, in this case the equation in (1.1) is hypoelliptic, see [24].

In kinetic theory,  $f$  represents the evolution of a particle distribution

$$f(X, Y, t) : U_X \times U_Y \times \mathbb{R}_+ \rightarrow \mathbb{R}, \quad U_X, U_Y \subset \mathbb{R}^m,$$

subject to geometric restrictions and models for the interactions and collisions between particles. In this case the left-hand side in (1.1) describes the evolution of  $f$  under the action of transport, with the free streaming operator. The right-hand side describes elastic collisions through the nonlinear Boltzmann collision operator. The Boltzmann equation is an integro- (partial)-differential equation with nonlocal operator in the kinetic variable  $X$ . The Boltzmann equation is a fundamental equation in kinetic theory in the sense that it has been derived rigorously, at least in some settings, from microscopic first principles. In the case of so called Coulomb interactions the Boltzmann collision operator is ill-defined and Landau proposed an alternative operator for these interactions, this operator is now called the Landau or the Landau-Coulomb operator. This operator can be stated as in (1.1) with

$$Q(f, \nabla_X f, X, Y, t) = \nabla_X \cdot (A(f)\nabla_X f + B(f)f), \quad (1.2)$$

where again  $A(f) = A(f)(X, Y, t)$  and  $B(f) = B(f)(X, Y, t)$  are nonlocal operators in the variable  $X$ . In this case the equation in (1.1) is a nonlinear, or rather quasilinear, drift-diffusion equation with coefficients given by convolution like averages of the unknown. As mentioned above the Landau equation is considered fundamental because of its close link to the Boltzmann equation for Coulomb interactions.

In the case of long-range interactions, the Boltzmann and Landau-Coulomb operators show local ellipticity provided the solution enjoys some pointwise bounds on the associated hydrodynamic fields and the local entropy. Indeed, assuming certain uniform in  $(Y, t) \in U_Y \times I$  bounds on local mass, energy, and entropy, see [30, 33], one can prove that

$$0 < \Lambda^{-1}I \leq A(f)(X, Y, t) \leq \Lambda I, \quad |B(f)(X, Y, t)| \leq \Lambda,$$

for some constant  $\Lambda \geq 1$  and for  $(X, Y, t) \in U_X \times U_Y \times I$ , i.e., under these assumptions the operator  $Q$  in (1.2) and in the Landau equation becomes locally uniformly elliptic. As a consequence, and as global well posedness for the Boltzmann equation and the construction of solutions in the large is an outstanding open problem, the study of conditional regularity for the Boltzmann and Landau equations has become a way to make progress on the regularity issues for these equations. We refer to [11, 13–15, 28, 33, 36, 37] for more on the connections between Kolmogorov-Fokker-Planck equations, the Boltzmann and Landau equation, statistical physics and conditional regularity.

Based on the idea of conditional regularity one is lead to study the local regularity of weak solutions to the equation in (1.1) with

$$Q(f, \nabla_X f, X, Y, t) = \nabla_X \cdot (A(X, Y, t) \nabla_X f) + B(X, Y, t) \nabla_X f, \quad (1.3)$$

assuming that  $A$  is measurable, bounded and uniformly elliptic, and that  $B$  is bounded. In [20], see also [21–23] for subsequent developments, the authors extended, for equations as in (1.1) assuming (1.3), the De Giorgi-Nash-Moser (DGNM) theory, which in its original form only considers elliptic or parabolic equations in divergence form, to hypoelliptic equations with rough coefficients including the one in (1.1) assuming (1.3). [20] has spurred considerable activity in the field, see below for a literature review, as the results proved give the correct scale- and translation-invariant estimates for local Hölder continuity and the Harnack inequality for weak solutions.

In this paper we consider equations as in (1.1) with

$$Q(f, \nabla_X f, X, Y, t) = \nabla_X \cdot (A(\nabla_X f, X, Y, t)), \quad (1.4)$$

subject to conditions on  $A$  which allow  $A$  to be a nonlinear function of  $\nabla_X f$ . In this case we refer to the equations in (1.1) as nonlinear Kolmogorov-Fokker-Planck type equations with rough coefficients. Our contributions is twofold. First, we establish higher integrability (Theorem 1.1) and local boundedness (Theorem 1.2) of weak sub-solutions, weak Harnack and Harnack inequalities (Theorem 1.3), and Hölder continuity with quantitative estimates (Theorem 1.4), for the equation

$$(\partial_t + X \cdot \nabla_Y)u = \nabla_X \cdot (A(\nabla_X u, X, Y, t)). \quad (1.5)$$

Second, we establish existence and uniqueness, in certain bounded  $X$ ,  $Y$  and  $t$  dependent domains, for a Dirichlet problem involving the equation in (1.5) also allowing for boundary data and a right hand side (Theorem 1.5). In the linear case, if  $A(X, Y, t)$  is a uniformly elliptic positive definite matrix with bounded measurable coefficients, then  $A(\xi, X, Y, t) = A(X, Y, t)\xi$  satisfies the hypothesis we impose on the symbol  $A$ , and in this case the equation in (1.5) reduces to the equation

$$(\partial_t + X \cdot \nabla_Y)u = \nabla_X \cdot (A(X, Y, t) \nabla_X u). \quad (1.6)$$

Concerning regularity, our results therefore generalize [20, 22, 23], to nonlinear Kolmogorov-Fokker-Planck type equations with rough coefficients.

To the best of our knowledge, nonlinear equations of the form in (1.5) have so far not been investigated in the literature, and the purpose of this paper is to contribute to the regularity and existence theory for these equations. We believe that generalizations of the De Giorgi-Nash-Moser (DGNM) theory to nonlinear Kolmogorov-Fokker-Planck type equations with rough coefficients are relevant and interesting. We also believe that our treatment of the Dirichlet problem is new and enlightening.

### 1.1. The symbol $A$

We consider equations as in (1.5) subject to conditions on  $A$ . Concerning the symbol  $A$  our baseline assumption is that  $A$  belongs to the class  $M(\Lambda)$ , where  $\Lambda \in [1, \infty)$  is a constant. In our treatment of the Dirichlet problem we will need to impose stronger conditions on  $A$  and we will assume that  $A$  belongs to the class  $R(\Lambda)$ . In the following  $\cdot$  denotes the standard Euclidean scalar product in  $\mathbb{R}^m$ .

**Definition 1.** Let  $\Lambda \in [1, \infty)$ . Then  $A$  is said to belong to the class  $M(\Lambda)$  if  $A = A(\xi, X, Y, t) : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$  is continuous with respect to  $\xi$ , measurable with respect to  $X, Y$  and  $t$ , and

$$\begin{aligned} (i) \quad & |A(\xi, X, Y, t)| \leq \Lambda|\xi|, \\ (ii) \quad & A(\xi, X, Y, t) \cdot \xi \geq \Lambda^{-1}|\xi|^2, \\ (iii) \quad & A(\lambda\xi, X, Y, t) = \lambda A(\xi, X, Y, t) \quad \forall \lambda \in \mathbb{R} \setminus \{0\}, \end{aligned} \quad (1.7)$$

for almost every  $(X, Y, t) \in \mathbb{R}^{N+1}$  and for all  $\xi \in \mathbb{R}^m$ .

**Definition 2.** Let  $\Lambda \in [1, \infty)$ . Then  $A$  is said to belong to the class  $R(\Lambda)$  if  $A = A(\xi, X, Y, t) : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$  is continuous with respect to  $\xi$ , measurable with respect to  $X, Y$  and  $t$ , and

$$\begin{aligned} (i) \quad & |A(\xi_1, X, Y, t) - A(\xi_2, X, Y, t)| \leq \Lambda|\xi_1 - \xi_2|, \\ (ii) \quad & (A(\xi_1, X, Y, t) - A(\xi_2, X, Y, t)) \cdot (\xi_1 - \xi_2) \geq \Lambda^{-1}|\xi_1 - \xi_2|^2, \\ (iii) \quad & A(\lambda\xi, X, Y, t) = \lambda A(\xi, X, Y, t) \quad \forall \lambda \in \mathbb{R} \setminus \{0\}, \end{aligned} \quad (1.8)$$

for almost every  $(X, Y, t) \in \mathbb{R}^{N+1}$  and for all  $\xi_1, \xi_2, \xi \in \mathbb{R}^m$ .

**Remark 1.1.** Note that (1.8)-(iii) implies that  $A(0, X, Y, t) = 0$  for a.e.  $(X, Y, t) \in \mathbb{R}^{N+1}$ . Hence we deduce from (1.8)-(i), (ii) and (iii) that  $R(\Lambda) \subset M(\Lambda)$ .

## 1.2. Dilations and group law

We will often use the notation  $(Z, t) = (X, Y, t) \in \mathbb{R}^{N+1}$  to denote points. The natural family of dilations for our operators and equations,  $(\delta_r)_{r>0}$ , on  $\mathbb{R}^{N+1}$ , is defined by

$$\delta_r(X, Y, t) = (rX, r^3Y, r^2t), \quad (1.9)$$

for  $(X, Y, t) \in \mathbb{R}^{N+1}$ ,  $r > 0$ . Our classes of operators are closed under the group law

$$(\tilde{Z}, \tilde{t}) \circ (Z, t) = (\tilde{X}, \tilde{Y}, \tilde{t}) \circ (X, Y, t) = (\tilde{X} + X, \tilde{Y} + Y + t\tilde{X}, \tilde{t} + t), \quad (1.10)$$

where  $(Z, t), (\tilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$ . Note that

$$(Z, t)^{-1} = (X, Y, t)^{-1} = (-X, -Y + tX, -t), \quad (1.11)$$

and hence

$$(\tilde{Z}, \tilde{t})^{-1} \circ (Z, t) = (\tilde{X}, \tilde{Y}, \tilde{t})^{-1} \circ (X, Y, t) = (X - \tilde{X}, Y - \tilde{Y} - (t - \tilde{t})\tilde{X}, t - \tilde{t}), \quad (1.12)$$

whenever  $(Z, t), (\tilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$ . Given  $(Z, t) = (X, Y, t) \in \mathbb{R}^{N+1}$  we let

$$\|(Z, t)\| = \|(X, Y, t)\| := |X, Y| + |t|^{\frac{1}{2}}, \quad |(X, Y)| = |X| + |Y|^{1/3}. \quad (1.13)$$

Given  $r > 0$  and  $(\tilde{Z}, \tilde{t}) = (\tilde{X}, \tilde{Y}, \tilde{t}) \in \mathbb{R}^{N+1}$ , we let

$$Q_r := \{(X, Y, t) : |X| < r, |Y| < r^3, -r^2 < t < 0\}, \quad Q_r(\tilde{Z}, \tilde{t}) := (\tilde{Z}, \tilde{t}) \circ Q_r. \quad (1.14)$$

We refer to  $Q_r(\tilde{Z}, \tilde{t})$  as a cylinder centered at  $(\tilde{Z}, \tilde{t})$  and of radius  $r$ .

### 1.3. Statement of main results: regularity of weak solutions

We here state the regularity part of our results, Theorem 1.1–Theorem 1.4. These theorems are derived under the assumption that the symbol  $A$  belongs to the class  $M(\Lambda)$  introduced in Definition 1. For the notions of weak sub-solutions, super-solutions and solutions, we refer to Definition 3 below. For the definitions of function spaces used we refer to the bulk of the paper.

**Theorem 1.1** (Higher integrability). *Let  $(Z_0, t_0) = (X_0, Y_0, t_0) \in \mathbb{R}^{N+1}$ ,  $0 < r_1 < r_0 \leq 1$  and let  $u$  be a non-negative weak sub-solution to (1.5) in an open set of  $\mathbb{R}^{N+1}$  containing  $Q_{r_0}(Z_0, t_0)$  in the sense of Definition 3 below. Then for any  $q \in [2, 2 + 1/m]$  and  $s \in [0, 1/3]$ , we have\**

$$\|u\|_{L^q(Q_{r_1}(Z_0, t_0))} \leq c_1 \left(2 + \frac{1}{m} - q\right)^{-1} \|u\|_{L^2(Q_{r_0}(Z_0, t_0))}, \quad (1.15)$$

$$\|u\|_{L_{t,X}^1 W_Y^{s,1}(Q_{r_1}(Z_0, t_0))} \leq c_2 \left(\frac{1}{3} - s\right)^{-1} \|u\|_{L^2(Q_{r_0}(Z_0, t_0))}. \quad (1.16)$$

Here

$$c_1 = \left(1 + \frac{1}{r_0 - r_1}\right) c, \quad c_2 = r_0^{1+2m} \left(1 + \frac{1}{r_0 - r_1}\right) c,$$

where

$$c = c(m, \Lambda) \left(1 + \frac{1}{(r_0 - r_1)^2} + \frac{|X_0| + r_0}{(r_0 - r_1)r_1^2} + \frac{1}{(r_0 - r_1)r_1}\right),$$

for some constant  $c(m, \Lambda) \geq 1$ .

**Theorem 1.2** (Local boundedness). *Let  $(Z_0, t_0) = (X_0, Y_0, t_0) \in \mathbb{R}^{N+1}$ ,  $0 < r_\infty < r_0 \leq 1$  and let  $u$  be a non-negative weak sub-solution to (1.5) in an open set of  $\mathbb{R}^{N+1}$  containing  $Q_{r_0}(Z_0, t_0)$  in the sense of Definition 3 below. Then for any  $p > 0$ , there exists a constant  $c = c(m, \Lambda) \geq 1$  and  $\theta = \theta(m) > 1$  such that*

$$\sup_{Q_{r_\infty}(Z_0, t_0)} u \leq c \left(\frac{1 + |X_0|}{r_\infty^2 (r_0 - r_\infty)^3}\right)^{\frac{\theta}{p}} \|u\|_{L^p(Q_{r_0}(Z_0, t_0))}. \quad (1.17)$$

**Theorem 1.3** (Harnack inequalities). *Let  $u$  be a non-negative weak super-solution to (1.5) in an open set of  $\mathbb{R}^{N+1}$  containing  $Q_1$  in the sense of Definition 3 below. Then there exists  $\zeta > 0$  and  $c \geq 1$ , both depending only on  $m$  and  $\Lambda$  such that*

$$\left(\iiint_{\tilde{Q}_{r_0/2}^-} u^\zeta(X, Y, t) \, dX \, dY \, dt\right)^{1/\zeta} \leq c \inf_{Q_{r_0/2}} u, \quad (1.18)$$

where  $r_0 = 1/20$  and  $\tilde{Q}_{r_0/2}^- := Q_{r_0/2}(0, 0, -19r_0^2/8)$ . Furthermore, if  $u$  is a non-negative weak solution to (1.5) in an open set of  $\mathbb{R}^{N+1}$  containing  $Q_1$ , then

$$\sup_{\tilde{Q}_{r_0/4}^-} u \leq c \inf_{Q_{r_0/4}} u, \quad (1.19)$$

where  $\tilde{Q}_{r_0/4}^- := Q_{r_0/4}(0, 0, -19r_0^2/8)$ .

---

\* $W_Y^{s,1}$  denotes the fractional Sobolev space.

**Theorem 1.4** (Hölder continuity). *Let  $u$  be a weak solution to (1.5) in an open set of  $\mathbb{R}^{N+1}$  containing  $Q_2$  in the sense of Definition 3 below. Then there exists  $\alpha \in (0, 1)$  and  $c \geq 1$ , both depending only on  $m$  and  $\Lambda$  such that*

$$\frac{|u(X_1, Y_1, t_1) - u(X_2, Y_2, t_2)|}{\|(X_2, Y_2, t_2)^{-1} \circ (X_1, Y_1, t_1)\|^\alpha} \leq c \|u\|_{L^2(Q_2)}, \quad (1.20)$$

whenever  $(X_1, Y_1, t_1), (X_2, Y_2, t_2) \in Q_1$ ,  $(X_1, Y_1, t_1) \neq (X_2, Y_2, t_2)$ .

#### 1.4. Statement of main results: existence and uniqueness for a Dirichlet problem

We here state the existence and uniqueness part of our results, Theorem 1.5. Throughout the paper we let  $U_X \subset \mathbb{R}^m$  be a bounded Lipschitz domain and let  $V_{Y,t} \subset \mathbb{R}^m \times \mathbb{R}$  be a bounded domain with boundary which is  $C^{1,1}$ -smooth, i.e.,  $C^1$  with respect to  $Y$  as well as  $t$ . Let  $N_{Y,t}$  denote the outer unit normal to  $V_{Y,t}$ . We establish existence and uniqueness of weak solutions to a formulation of the Dirichlet problem

$$\begin{cases} \nabla_X \cdot (A(\nabla_X u, X, Y, t)) - (\partial_t + X \cdot \nabla_Y)u = g^* & \text{in } U_X \times V_{Y,t}, \\ u = g & \text{on } \partial_{\mathcal{K}}(U_X \times V_{Y,t}). \end{cases} \quad (1.21)$$

Here

$$\partial_{\mathcal{K}}(U_X \times V_{Y,t}) := (\partial U_X \times V_{Y,t}) \cup \{(X, Y, t) \in \overline{U_X} \times \partial V_{Y,t} \mid (X, 1) \cdot N_{Y,t} < 0\}. \quad (1.22)$$

$\partial_{\mathcal{K}}(U_X \times V_{Y,t})$  will be referred to as the Kolmogorov boundary of  $U_X \times V_{Y,t}$ , and the Kolmogorov boundary serves, in our context, as the natural substitute for the parabolic boundary used in the context of the Cauchy-Dirichlet problem for uniformly elliptic parabolic equations. In particular, we study weak solutions in the sense of Definition 4. For the definition of the functional setting we refer to Section 2. We believe that the following result is of independent interest in particular as we allow the symbol  $A$  to depend nonlinearly on  $\nabla_X u$ .

**Theorem 1.5** (Existence and uniqueness). *Let  $(g, g^*) \in W(U_X \times V_{Y,t}) \times L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))$  and assume that  $A$  belongs to the class  $R(\Lambda)$  introduced in Definition 2. Then there exists a unique weak solution  $u$  to the problem in (1.21) in the sense of Definition 4 below. Furthermore, there exists a constant  $c$ , depending only on  $m$ ,  $\Lambda$  and  $U_X \times V_{Y,t}$ , such that*

$$\|u\|_{W(U_X \times V_{Y,t})} \leq c (\|g\|_{W(U_X \times V_{Y,t})} + \|g^*\|_{L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))}). \quad (1.23)$$

#### 1.5. Known regularity results

As mentioned, the equation in (1.6), possibly also allowing for lower order terms, has attracted considerable attention in recent years. Anceschi-Cinti-Pascucci-Polidoro-Ragusa [2, 12, 34] proved local boundedness of weak sub-solutions of (1.6) and some versions thereof. Their approach is based on the Moser's iteration technique, the use of fundamental solutions and a Sobolev type inequality is crucial. It is worth noting that while the results in these papers are stated assuming only bounded and measurable coefficients, an implicit regularity assumption on the coefficients is imposed as the authors use a stronger notion of weak solutions assuming also  $(\partial_t + X \cdot \nabla_Y)u \in L^2_{\text{loc}}$ . It is unclear for what assumptions on the coefficients such weak solutions can be constructed. Bramanti-Cerutti-

Manfredini-Polidoro-Ragusa [8, 32, 35] proved  $L^p$  estimates, interior Sobolev regularity and local Hölder continuity of weak solutions of (1.6) imposing additional assumptions on the coefficients beyond bounded, measurable and elliptic. In fact it was only recently that Golse-Imbert-Mouhot-Vasseur [20] proved local boundedness, Harnack inequality and local Hölder continuity of (true) weak solutions of (1.6) based on De-Giorgi and Moser's iteration technique. Still, it seems unclear to us how the authors actually resolve questions concerning the existence of weak solutions unless smooth coefficients are assumed qualitatively. However, subsequent developments have appeared in [21–23]. A weak Harnack inequality for weak super-solutions of (1.6) has been obtained by Guerand-Imbert [22] and this has been generalized by Anceschi-Rebucci [3]. In [23], Guerand-Mouhot revisited the theory for the linear equation in (1.6), also allowing for lower order terms, and gave lucid, novel and short proofs of the De Giorgi intermediate-value Lemma, weak Harnack and Harnack inequalities, and the Hölder continuity with quantitative estimates. [23] is an essentially self-contained account of the linear theory. Local Hölder continuity results are also proved in Wang-Zhang [38–40] for various linear analogues of (1.6). We emphasize that all results mentioned concern linear equations. Zhu [41] proved local boundedness and local Hölder continuity of weak solutions of (1.6) when the drift term  $\partial_t + X \cdot \nabla_Y$  is replaced by  $\partial_t + b(X) \cdot \nabla_Y$  for some nonlinear function  $b$ .

### 1.6. Known existence results

Boundary value problems for equations as in (1.6) but in non-divergence form were studied by Manfredini [31] who proved existence of strong solutions for the Dirichlet problem assuming Hölder continuous coefficients. Lanconelli-Lascialfari-Morbidelli [26, 27] considered a quasilinear case, still in non-divergence form, allowing the coefficients to depend not only on  $(X, Y, t)$  but also the solution  $u$ , and as a function of  $(X, Y, t)$  the coefficients are assumed to be with Hölder continuous. In fact, functional analytic approaches to weak solutions to Kramers equation and Kolmogorov-Fokker-Planck equations have only recently been developed. Albritton-Armstrong-Mourrat-Novack [1] have developed a functional analytic approach to study well-posedness of Kramers equation, and its parabolic analogue

$$\partial_t u - \Delta_X u + X \cdot \nabla_X u + X \cdot \nabla_Y u + b \cdot \nabla_X u = g^*, \quad (1.24)$$

for suitable  $g^*$ . Equation (1.24) is often referred to as the kinetic Fokker-Planck equation. Litsgård-Nyström [29] studied existence and uniqueness results for the (linear) Dirichlet problem associated with (1.6), with rough coefficients  $A$ . In particular, in [29] Theorem 1.5 is proved in the case when  $A(\xi, X, Y, t) = A(X, Y, t)\xi$ . However, existence and uniqueness for (1.5) do not seem to have been studied in the literature so far. It is important to note that Theorem 1.5 states, similar to [29], the existence of a unique weak solution  $u$  to the problem in (1.21) in the sense of Definition 4 below. The latter is, as it assumes no knowledge of underlying traces, trace spaces and extension operators in the functional setting considered, a weaker formulation of the Dirichlet problem compared to what one usually aims for. Indeed, this is one way to formulate a weak form of the Dirichlet problem which circumvents a largely open problem in the context of kinetic Fokker-Planck equations, linear as well as non-linear, and that is the problem of a well defined trace operator and trace inequality. We refer to Section 6 for more.

### 1.7. Proofs

The regularity part of our results is modelled on the approach of Golse-Imbert-Mouhot-Vasseur [20] and the work of Guerand-Mouhot [23]. In fact, as can be seen from the very formulations of our regularity results, this part of our work is strongly influenced by [23] and armed with Theorem 1.1 and Theorem 1.2 we can to large extent refer to the corresponding arguments in [23] for the proofs of Theorem 1.3 and Theorem 1.4. The new difficulties in our case stem from the nonlinearity of  $A$  in  $\nabla_X u$ . However, as we learn from the regularity theory for quasi-linear parabolic PDEs, see [16] for example, a careful development of the De Giorgi-Nash-Moser theory tends to be robust enough to handle the type of non-linearities considered in this paper. The higher integrability result in Theorem 1.1 is proved by combining the energy estimate in Lemma 3.1 with a Sobolev regularity estimate and here it is important that  $A$  has linear growth in  $\nabla_X u$ . In particular, in the proof of Theorem 1.1 one is lead, after preliminaries and the use of an appropriate cut-off function, to conduct estimates for a (global) weak sub-solution  $u_1$  to the equation

$$(\partial_t + X \cdot \nabla_Y)u_1 \leq \nabla_X \cdot A(\nabla_X u_1, X, Y, t) + g^*, \quad g^* := -(\nabla_X \cdot F_1 + F_0) \text{ in } \mathbb{R}^{N+1}, \quad (1.25)$$

where  $F_1, F_0$  are in  $L^2(\mathbb{R}^{N+1})$  and  $u_1, F_1, F_0$  are supported in  $Q_{r_0}(0, 0, 0)$ . To close the argument, as  $u_1$  is only a weak sub-solution, it seems important to replace it by a function which actually solves an equation. In particular, to make this operational one needs to construct a weak solution  $v$  to

$$(\partial_t + X \cdot \nabla_Y)v = \nabla_X \cdot A(\nabla_X v, X, Y, t) + g^*, \quad (1.26)$$

such that  $v$  bounds  $u_1$  from above. One approach to Sobolev regularity estimates is then attempt to use an approach based on Bouchut [7] which implies a Sobolev embedding

$$H_{X,Y,t}^{1/3}(\mathbb{R}^{N+1}) \rightarrow L_{X,Y,t}^q(\mathbb{R}^{N+1}), \quad q := \frac{6(2m+1)}{6m+1} > 2. \quad (1.27)$$

To get hold of the  $H_{X,Y,t}^{1/3}(\mathbb{R}^{N+1})$  norm of  $v$  one uses a result of Bouchut [7] which gives control of  $D_Y^{1/3}v, D_t^{1/3}v$  given energy estimates. To be able to bound  $u_1$  from above by  $v$  as in (1.26) one seems to need Theorem 1.5 and the comparison principle that we prove in Theorem 5.1 below. As the result of Bouchut [7] requires a solution which exists globally in time one can make this approach operational using Theorem 1.5 to prove Theorem 1.1 with the cylinders in (1.14) replaced by centered cylinders. An alternative approach to Sobolev regularity estimates, which in the end gives Theorem 1.1 as stated, is to first observe that if  $u_1$  satisfies (1.25), then one deduces that the weak formulation of (1.25) induces a positive distribution. One is therefore lead to prove estimates for  $v$  satisfying

$$(\partial_t + X \cdot \nabla_Y)v = \nabla_X \cdot A(\nabla_X v, X, Y, t) + g^* - \mu, \quad (1.28)$$

where  $\mu$  is now a positive measure. Due to the structure of  $g^*$ , Sobolev regularity estimates can then be deduced using a semi-classical approach via the fundamental solution associated to the linear equation  $(\partial_t + X \cdot \nabla_Y)f = \Delta_X f$  originally studied by Kolmogorov, see Lemma 10 in [23]. In the end, we follow this approach and here it is again important that  $A$  has linear growth in  $\nabla_X u$ . Armed with the Sobolev regularity estimates the proofs of Theorem 1.1–Theorem 1.4 can be completed along the lines of the corresponding arguments in the linear case. Finally, to prove the existence and uniqueness result in Theorem 1.5 we use a variational approach and proceed along the lines of [1, 4, 29]. In particular, our argument is similar to the proof of Theorem 1.1 in [29].



## 1.8. Organization of the paper

In Section 2 we introduce the functional setting and the notion of weak solutions. Section 3 is devoted to a number of preliminary technical results to be used in the proofs of Theorem 1.1–Theorem 1.4. Theorem 1.1–Theorem 1.4 are proved in Section 4, and in the proof of Theorem 1.3 and Theorem 1.4 we for brevity mainly refer to the corresponding arguments in [23]. Theorem 1.5 is proved in Section 5. In Section 6 we mention a number of challenging problems for future research which we hope will inspire the community to look further into the topic of nonlinear Kolmogorov-Fokker-Planck type equations.

## 2. The functional setting and weak solutions

### 2.1. Function spaces

We denote by  $H_X^1(U_X)$  the Sobolev space of functions  $g \in L^2(U_X)$  whose distributional gradient in  $U_X$  lies in  $(L^2(U_X))^m$ , i.e.,

$$H_X^1(U_X) := \{g \in L_X^2(U_X) \mid \nabla_X g \in (L^2(U_X))^m\},$$

and we set

$$\|g\|_{H_X^1(U_X)} := \left( \|g\|_{L^2(U_X)}^2 + \|\nabla_X g\|_{L^2(U_X)}^2 \right)^{1/2}, \quad g \in H_X^1(U_X).$$

We let  $H_{X,0}^1(U_X)$  denote the closure of  $C_0^\infty(U_X)$  in the norm of  $H_X^1(U_X)$  and we recall, as  $U_X$  is a bounded Lipschitz domain, that  $C^\infty(\overline{U_X})$  is dense in  $H_X^1(U_X)$ . In particular, equivalently we could define  $H_X^1(U_X)$  as the closure of  $C^\infty(\overline{U_X})$  in the norm  $\|\cdot\|_{H_X^1(U_X)}$ . Note that as  $H_{X,0}^1(U_X)$  is a Hilbert space it is reflexive, hence  $(H_{X,0}^1(U_X))^* = H_X^{-1}(U_X)$  and  $(H_X^{-1}(U_X))^* = H_{X,0}^1(U_X)$ , where  $(\cdot)^*$  denotes the dual. Based on this we let  $H_X^{-1}(U_X)$  denote the dual to  $H_{X,0}^1(U_X)$  acting on functions in  $H_{X,0}^1(U_X)$  through the duality pairing  $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{H_X^{-1}(U_X), H_{X,0}^1(U_X)}$ . We let  $L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))$  be the space of measurable function  $u : V_{Y,t} \rightarrow H_{X,0}^1(U_X)$  equipped with the norm

$$\|u\|_{L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))}^2 := \iint_{V_{Y,t}} \|u(\cdot, Y, t)\|_{H_X^1(U_X)}^2 \, dY \, dt.$$

$L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))$  is defined analogously. In analogy with the definition of  $H_X^1(U_X)$ , we let  $W(U_X \times V_{Y,t})$  be the closure of  $C^\infty(\overline{U_X} \times \overline{V_{Y,t}})$  in the norm

$$\|u\|_{W(U_X \times V_{Y,t})} := \left( \|u\|_{L_{Y,t}^2(V_{Y,t}, H_X^1(U_X))}^2 + \|(\partial_t + X \cdot \nabla_Y)u\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))}^2 \right)^{1/2}. \quad (2.1)$$

In particular,  $W(U_X \times V_{Y,t})$  is a Banach space and  $u \in W(U_X \times V_{Y,t})$  if and only if

$$u \in L_{Y,t}^2(V_{Y,t}, H_X^1(U_X)) \quad \text{and} \quad (\partial_t + X \cdot \nabla_Y)u \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)). \quad (2.2)$$

Note that the dual of  $L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))$ , denoted by  $(L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X)))^*$ , satisfies

$$(L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X)))^* = L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)),$$

and, as mentioned above,

$$(L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)))^* = L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X)).$$

Finally, the spaces  $L^2_{Y,t,\text{loc}}(V_{Y,t}, H^1_{X,\text{loc}}(U_X))$ ,  $L^2_{Y,t,\text{loc}}(V_{Y,t}, H^{-1}_{X,\text{loc}}(U_X))$ , and  $W_{\text{loc}}(U_X \times V_{Y,t})$  are defined in the natural way. The topological boundary of  $U_X \times V_{Y,t}$  is denoted by  $\partial(U_X \times V_{Y,t})$ . Let  $N_{Y,t}$  denote the outer unit normal to  $V_{Y,t}$ . We define a subset  $\partial_{\mathcal{K}}(U_X \times V_{Y,t}) \subset \partial(U_X \times V_{Y,t})$ , the Kolmogorov boundary of  $U_X \times V_{Y,t}$ , as in (1.22). We let  $C^{\infty}_{\mathcal{K},0}(\overline{U_X \times V_{Y,t}})$  and  $C^{\infty}_{X,0}(\overline{U_X \times V_{Y,t}})$  be the set of functions in  $C^{\infty}(\overline{U_X \times V_{Y,t}})$  which vanish on  $\partial_{\mathcal{K}}(U_X \times V_{Y,t})$  and  $\{(X, Y, t) \in \partial U_X \times \overline{V_{Y,t}}\}$ , respectively. We let  $W_0(U_X \times V_{Y,t})$  and  $W_{X,0}(U_X \times V_{Y,t})$  denote the closure in the norm of  $W(U_X \times V_{Y,t})$  of  $C^{\infty}_{\mathcal{K},0}(\overline{U_X \times V_{Y,t}})$  and  $C^{\infty}_{X,0}(\overline{U_X \times V_{Y,t}})$ , respectively.

## 2.2. Weak solutions

We here introduce the notion of weak solutions.

**Definition 3.** Let  $g^* \in L^2_{Y,t}(V_{Y,t}, H^{-1}_X(U_X))$ . A function  $u \in W_{\text{loc}}(U_X \times V_{Y,t})$  is said to be a weak sub-solution (or super-solution) to the equation

$$(\partial_t + X \cdot \nabla_Y)u - \nabla_X \cdot (A(\nabla_X u, X, Y, t)) + g^* = 0 \text{ in } U_X \times V_{Y,t}, \quad (2.3)$$

if for every  $V_X \times V_Y \times J \Subset U_X \times V_{Y,t}$ , and for all non-negative  $\phi \in L^2_{Y,t}(V_Y \times J, H^1_{X,0}(V_X))$ , we have

$$\begin{aligned} & \iiint_{V_X \times V_Y \times J} A(\nabla_X u, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & + \iint_{V_Y \times J} \langle g^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)u(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \leq 0 \quad (\text{or } \geq). \end{aligned} \quad (2.4)$$

We say that  $u \in W_{\text{loc}}(U_X \times V_{Y,t})$  is a weak solution to the Eq (2.3) if equality holds in (2.4) without a sign restriction on  $\phi$ .

Note that if  $u$  is a weak sub-solution (or super-solution) of (2.3) in the sense of Definition 3 above, with  $g^* \equiv 0$ , then

$$\begin{aligned} & \iint_{V_X \times V_Y} u(X, Y, t_2) \phi(X, Y, t_2) \, dX \, dY - \iint_{V_X \times V_Y} u(X, Y, t_1) \phi(X, Y, t_1) \, dX \, dY \\ & - \int_{t_1}^{t_2} \iint u(\partial_t + X \cdot \nabla_Y) \phi \, dX \, dY \, dt + \int_{t_1}^{t_2} \iint A(\nabla_X u, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & \leq 0 \quad (\text{or } \geq), \end{aligned} \quad (2.5)$$

whenever  $\phi \in C^{\infty}((t_1, t_2), C^{\infty}_0(V_X \times V_Y))$ , is non-negative function. Furthermore, equality holds in (2.5) for every weak solution  $u$  of (2.3) without a sign restriction on  $\phi$ .

**Remark 2.1.** Assume  $g^* \equiv 0$ . (i) From Definition 3, it is clear that, if  $u$  is a weak sub-solution (resp. super-solution or solution) of (2.3) in  $U_X \times V_{Y,t}$ , then for any  $k \in \mathbb{R}$ , the function  $v = (u - k)$  is also weak sub-solution (resp. super-solution or solution) of (2.3) in  $U_X \times V_{Y,t}$ .

(ii) Using the homogeneity property (iii) of  $A$ , it follows that, (a) for any  $c \geq 0$ ,  $cu$  is a weak sub-solution (resp. super-solution or solution) of (2.3) in  $U_X \times V_{Y,t}$ , provided  $u$  is a weak sub-solution (resp. super-solution or solution) of (2.3) in  $U_X \times V_{Y,t}$  and (b)  $u$  is a weak solution of (2.3) in  $U_X \times V_{Y,t}$  if and only if  $-u$  is a weak solution of (2.3) in  $U_X \times V_{Y,t}$ .

### 2.3. The Dirichlet problem

Theorem 1.5 is a statement concerning existence and uniqueness of weak solutions to a formulation of the Dirichlet problem in (1.21). In particular, we study weak solutions in the following sense.

**Definition 4.** Consider  $(g, g^*) \in W(U_X \times V_{Y,t}) \times L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))$ . Given  $(g, g^*)$ ,  $u$  is said to be a weak solution to the problem in (1.21) if

$$u \in W(U_X \times V_{Y,t}), \quad (u - g) \in W_0(U_X \times V_{Y,t}), \quad (2.6)$$

and if

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} A(\nabla_X u, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & + \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)u(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt = 0, \end{aligned} \quad (2.7)$$

for all  $\phi \in L^2_{Y,t}(V_{Y,t}, H_{X,0}^1(U_X))$  and where  $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{H_X^{-1}(U_X), H_{X,0}^1(U_X)}$  is the duality pairing in  $H_X^{-1}(U_X)$ . If in (2.7),  $=$  is replaced by  $\leq$  ( $\geq$ ) whenever  $\phi \geq 0$ , then  $u$  is said to be a weak sub- (super-) solution of (1.21) respectively.

### 3. Technical lemmas

In this section we prove a number of technical results to be used in the proof of Theorem 1.1–Theorem 1.4. Throughout the rest of the paper, we use the notation  $s^+ := \max\{s, 0\}$  for  $s \in \mathbb{R}$ . Moreover, from Sections 3 and 4, we assume that the symbol  $A$  belongs to the class  $M(\Lambda)$  introduced in Definition 1.

**Lemma 3.1.** Let  $Z_0 = (X_0, Y_0, t_0) \in \mathbb{R}^{N+1}$ ,  $0 < r_1 < r_0$ , be such that  $Q_{r_0}(Z_0, t_0) \Subset U_X \times V_{Y,t}$ . Let  $u$  be a weak sub-solution of the Eq (1.5) in  $U_X \times V_{Y,t}$  in the sense of Definition 3. Then

$$\begin{aligned} & \sup_{t_0 - r_1^2 < t < t_0} \iint_{Q'(Z_0, r_1)} u^2(X, Y, t) \, dX \, dY + \Lambda^{-1} \iiint_{Q_{r_1}(Z_0, r_1)} |\nabla_X u|^2 \, dX \, dY \, dt \\ & \leq c c_{0,1} \iiint_{Q_{r_0}(Z_0, t_0)} u(X, Y, t)^2 \, dX \, dY \, dt, \end{aligned} \quad (3.1)$$

where  $Q'(Z_0, r) := \{(X, Y) : (X, Y, t) \in Q_r(Z_0, t_0)\}$  for  $r > 0$ ,  $c = c(m, \Lambda) \geq 1$  and

$$c_{0,1} := \frac{1}{(r_0 - r_1)^2} + \frac{r_0 + |X_0|}{(r_0 - r_1)r_1^2} + \frac{1}{(r_0 - r_1)r_1} + 1.$$

*Proof.* Let  $t_1 := t_0 - r_0^2$  and  $t_2 := t_0$ . Considering  $l_1, l_2$ , such that  $t_1 < l_1 < l_2 < t_2$ , we introduce for  $\epsilon > 0$  the function  $\theta_\epsilon \in W^{1,\infty}((t_1, t_2))$  by

$$\theta_\epsilon(t) := \begin{cases} 0 & \text{if } t_1 \leq t \leq l_1 - \epsilon, \\ 1 + \frac{t-l_1}{\epsilon}, & \text{if } l_1 - \epsilon < t \leq l_1, \\ 1 & \text{if } l_1 < t \leq l_2, \\ 1 - \frac{t-l_2}{\epsilon} & \text{if } l_2 \leq t \leq l_2 + \epsilon, \\ 0 & \text{if } l_2 + \epsilon < t \leq t_2. \end{cases} \quad (3.2)$$

Let  $\psi \in [0, 1]$  be smooth in  $Q_{r_0}(Z_0, t_0)$  such that  $\psi \equiv 1$  on  $Q_{r_1}(Z_0, t_0)$  and  $\psi \equiv 0$  outside  $Q_{r_0}(Z_0, t_0)$  satisfying

$$|\nabla_X \psi| \leq \frac{c}{r_0 - r_1}, \quad |\nabla_Y \psi| \leq \frac{c}{(r_0 - r_1)r_1^2}, \quad |\partial_t \psi| \leq \frac{c}{(r_0 - r_1)r_1},$$

for some constant  $c = c(m) \geq 1$ .

Consider the function  $\phi(X, Y, t) = 2u(X, Y, t)\psi^2(X, Y, t)\theta_\epsilon(t)$ . We intend to test (2.5) with  $\phi$  and the following deductions are formal. However, as  $u$  is a weak sub-solution of the Eq (1.5) in  $U_X \times V_{Y,t}$  in the sense of Definition 3, we know that  $u \in W_{\text{loc}}(U_X \times V_{Y,t})$  and as  $W(U_X \times V_{Y,t})$  is defined as the closure of  $C^\infty(\overline{U_X} \times \overline{V_{Y,t}})$  in the norm introduced in (2.1) our deduction can be made rigorous a posteriori. Testing (2.5) with  $\phi(X, Y, t)$ , letting  $\epsilon \rightarrow 0$ , and then adding

$$\iiint u^2 \partial_t (\psi^2) \, dX \, dY \, dt$$

on both sides of the resulting inequality, we deduce that

$$\begin{aligned} I(l_2) - I(l_1) + 2 \iiint A(\nabla_X u, X, Y, t) \cdot \nabla_X (u\psi^2) \, dX \, dY \, dt \\ \leq \iiint u^2 (\partial_t + X \cdot \nabla_Y) \psi^2 \, dX \, dY \, dt, \end{aligned} \quad (3.3)$$

where

$$I(t) := \iint \psi^2(X, Y, t) u^2(X, Y, t) \, dX \, dY.$$

Using (1.7), (3.3) yields

$$\begin{aligned} I(l_2) - I(l_1) + 2 \iiint \psi^2 A(\nabla_X u, X, Y, t) \cdot \nabla_X u \, dX \, dY \, dt \\ \leq \iiint u^2 (\partial_t + X \cdot \nabla_Y) \psi^2 \, dX \, dY \, dt \\ - 4 \iiint u \psi A(\nabla_X u, X, Y, t) \cdot \nabla_X \psi \, dX \, dY \, dt \\ \leq \iiint u^2 \{(\partial_t + X \cdot \nabla_Y) \psi^2 + 4\Lambda^3 (\psi + |\nabla_X \psi|)^2\} \, dX \, dY \, dt \\ + \iiint \Lambda^{-1} \psi^2 |\nabla_X u|^2 \, dX \, dY \, dt. \end{aligned} \quad (3.4)$$

Furthermore, using (1.7)-(i), (ii) we can continue the above estimate and conclude that

$$\begin{aligned} I(l_2) - I(l_1) + \Lambda^{-1} \iiint_{Q'(Z_0, r_0) \times (l_1, l_2)} \psi^2 |\nabla_X u|^2 \, dX \, dY \, dt \\ \leq \iiint u^2 \{(\partial_t + X \cdot \nabla_Y) \psi^2 + 4\Lambda^3 (\psi + |\nabla_X \psi|)^2\} \, dX \, dY \, dt. \end{aligned} \quad (3.5)$$

Using the properties of  $\psi$  and first letting  $l_1 \rightarrow t_1$ , and then letting  $l_2 \rightarrow t_2$  in (3.5), we obtain

$$\Lambda^{-1} \iiint_{Q_{r_1}(Z_0, t_0)} |\nabla_X u|^2 \, dX \, dY \, dt$$

$$\begin{aligned}
&\leq \iiint_{Q_{r_0}(Z_0, t_0)} u^2 \{(\partial_t + X \cdot \nabla_Y) \psi^2 + 4\Lambda^3(\psi + |\nabla_X \psi|)^2\} \, dX \, dY \, dt \\
&\leq cc_{0,1} \iiint_{Q_{r_0}(Z_0, t_0)} u(X, Y, t)^2 \, dX \, dY \, dt,
\end{aligned} \tag{3.6}$$

where  $c = c(m, \Lambda) \geq 1$  and

$$c_{0,1} := \frac{1}{(r_0 - r_1)^2} + \frac{r_0 + |X_0|}{(r_0 - r_1)r_1^2} + \frac{1}{(r_0 - r_1)r_1} + 1.$$

Again using the properties of  $\psi$  and first letting  $l_1 \rightarrow t_1$  in (3.5), then taking supremum over  $l_2 \in [t_0 - r_1^2, t_0)$  and noting that for such  $l_2$ ,  $\psi \equiv 1$ , we also have

$$\sup_{t_0 - r_1^2 < t < t_0} \iint_{Q^t(Z_0, r_0)} u^2(X, Y, t) \, dX \, dY \tag{3.7}$$

$$\begin{aligned}
&\leq \iiint_{Q_{r_0}(Z_0, r_0)} u^2 \{(\partial_t + X \cdot \nabla_Y) \psi^2 + 4\Lambda^3(\psi + |\nabla_X \psi|)^2\} \, dX \, dY \, dt \\
&\leq cc_{0,1} \iiint_{Q_{r_0}(Z_0, t_0)} u(X, Y, t)^2 \, dX \, dY \, dt.
\end{aligned} \tag{3.8}$$

This completes the proof.  $\square$

**Lemma 3.2.** *Let  $u$  be a weak sub-solution of the Eq (1.5) in  $U_X \times V_{Y,t}$  in the sense of Definition 3. Let  $k \in \mathbb{R}$ . Then  $(u - k)^+$  is also a weak sub-solution of the Eq (1.5) in  $U_X \times V_{Y,t}$  in the sense of Definition 3.*

*Proof.* By Remark 2.1, it is enough to prove that  $u^+$  is a weak sub-solution of (1.5). Let  $\epsilon > 0$  and  $\phi \in L^2_{Y,t}(V_Y \times J, H^1_{X,0}(V_X))$  be a non-negative test function in (2.4). Then  $\frac{u^+}{(u^+ + \epsilon)} \phi \in L^2_{Y,t}(V_Y \times J, H^1_{X,0}(V_X))$  is also a non-negative test function in (2.4). Using  $\frac{u^+}{(u^+ + \epsilon)} \phi$  as a test function in (2.4), we obtain

$$\begin{aligned}
&\iiint_{V_X \times V_Y \times J} A(\nabla_X u, X, Y, t) \cdot \nabla_X \phi \frac{u^+}{(u^+ + \epsilon)} \, dX \, dY \, dt \\
&+ \epsilon \iiint_{V_X \times V_Y \times J} A(\nabla_X u, X, Y, t) \cdot \frac{\nabla_X u^+}{(u^+ + \epsilon)^2} \phi \, dX \, dY \, dt \\
&+ \iint_{V_Y \times J} \langle (\partial_t + X \cdot \nabla_Y) u(\cdot, Y, t), \frac{u^+}{(u^+ + \epsilon)} \phi(\cdot, Y, t) \rangle \, dY \, dt \leq 0.
\end{aligned} \tag{3.9}$$

Letting  $\epsilon \rightarrow 0$ , we obtain

$$\begin{aligned}
&\iint_{V_Y \times J} \langle (\partial_t + X \cdot \nabla_Y) u^+(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \\
&+ \iiint_{V_X \times V_Y \times J} A(\nabla_X u^+, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\
&+ \liminf_{\epsilon \rightarrow 0} \epsilon \iiint_{V_X \times V_Y \times J} \frac{A(\nabla_X u^+, X, Y, t) \cdot \nabla_X u^+}{(u^+ + \epsilon)^2} \phi \, dX \, dY \, dt \leq 0.
\end{aligned} \tag{3.10}$$

However, by (1.7)-(ii)

$$\iiint_{V_X \times V_Y \times J} \frac{A(\nabla_X u^+, X, Y, t) \cdot \nabla_X u^+}{(u^+ + \epsilon)^2} \phi \, dX \, dY \, dt \geq 0. \quad (3.11)$$

Hence,

$$\begin{aligned} & \iint_{V_Y \times J} \langle (\partial_t + X \cdot \nabla_Y) u^+(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \\ & + \iiint_{V_X \times V_Y \times J} A(\nabla_X u^+, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \leq 0. \end{aligned} \quad (3.12)$$

This proves that  $u^+$  is a weak sub-solution.  $\square$

The following result follows from [23, Lemma 10].

**Lemma 3.3.** *Let  $f \geq 0$  be locally integrable such that*

$$(\partial_t + X \cdot \nabla_Y - \Delta_X) f = \nabla_X \cdot F_1 + F_2 - \mu, \quad (3.13)$$

where  $F_1, F_2 \in L^1 \cap L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)$  and  $\mu \in M^1(\mathbb{R}^{2m} \times \mathbb{R}_-)$  is a non-negative measure with finite mass in  $\mathbb{R}^{2m} \times \mathbb{R}_-$  such that  $F_1, F_2$  and  $\mu$  have compact support, in the time variable, included in  $(-\tau, 0]$ . Then for any  $p \in [2, 2 + 1/m)$  and  $\sigma \in [0, 1/3)$  we have

$$\|f\|_{L^p(\mathbb{R}^{2m} \times \mathbb{R}_-)} \leq c \left(2 + \frac{1}{m} - p\right)^{-1} (\|F_1\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|F_2\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)}) \quad (3.14)$$

and

$$\begin{aligned} \|f\|_{L^1_{t,X} W_Y^{\sigma,1}(\mathbb{R}^{2m} \times \mathbb{R}_-)} & \leq c \left(\frac{1}{3} - \sigma\right)^{-1} (\|F_1\|_{L^1(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|F_2\|_{L^1(\mathbb{R}^{2m} \times \mathbb{R}_-)}) \\ & + c \left(\frac{1}{3} - \sigma\right)^{-1} \|\mu\|_{M^1(\mathbb{R}^{2m} \times \mathbb{R}_-)}, \end{aligned} \quad (3.15)$$

for some constant  $c = c(\tau)$ .

The lemmas stated so far will be sufficient for our proof of Theorem 1.1 and Theorem 1.2.

### 3.1. Additional lemmas for the proofs of Theorem 1.3 and Theorem 1.4

**Lemma 3.4** (Weak Poincaré inequality). *Let  $\epsilon \in (0, 1)$  and  $\sigma \in (0, \frac{1}{3})$ . Then every non-negative weak sub-solution  $u$  of (1.5) in  $Q_5$  in the sense of Definition 3 satisfies*

$$\left\| (u - u_{Q_1^-})^+ \right\|_{L^1(Q_1^+)} \leq c \left( \frac{1}{\epsilon^{m+2}} \|\nabla_X u\|_{L^1(Q_5)} + \epsilon^\sigma \left(\frac{1}{3} - \sigma\right)^{-1} \|u\|_{L^2(Q_5)} \right), \quad (3.16)$$

for some constant  $c = c(m, \Lambda) \geq 1$ , where  $Q_1^- := Q_1(0, 0, -1)$  and  $u_{Q_1^-} := \frac{1}{|Q_1^-|} \int_{Q_1^-} u$ .

*Proof.* Using that  $u$  is a non-negative weak sub-solution of (1.5), Theorem 1.1 and the property 1.7-(i), the conclusion of the lemma follows from the lines of the proof of [23, Proposition 13, pages 8–10].  $\square$

**Lemma 3.5** (Intermediate value lemma). *Let  $\delta_1, \delta_2 \in (0, 1)$  be given. Then there exists constants  $\theta = c(m, \Lambda)(\delta_1\delta_2)^{10m+15}$ ,  $r_0 = \frac{1}{20}$ , and  $v \geq c(m, \Lambda)(\delta_1\delta_2)^{5m+8}$ , such that the following holds. Let  $u : Q_1 \rightarrow \mathbb{R}$  be a weak sub-solution of (1.5) in  $Q_5$  in the sense of Definition 3, assume that  $u \leq 1$  in  $Q_{\frac{1}{2}}$ , and that*

$$|\{u \leq 0\} \cap Q_{r_0}^-| \geq \delta_1 |Q_{r_0}^-| \quad \text{and} \quad |\{u \geq 1 - \theta\} \cap Q_{r_0}| \geq \delta_2 |Q_{r_0}|, \quad (3.17)$$

where  $Q_{r_0}^- := Q_{r_0}(0, 0, -2r_0^2)$ . Then

$$\left| \{0 < u < 1 - \theta\} \cap Q_{\frac{1}{2}} \right| \geq v |Q_{\frac{1}{2}}|. \quad (3.18)$$

*Proof.* Using Lemma 3.1, Lemma 3.2 and Lemma 3.4, the result follows from the lines of the proof of [23, Theorem 3, pages 11–12].  $\square$

**Lemma 3.6** (Measure to pointwise upper bound). *Given  $\delta \in (0, 1)$  and  $r_0 = \frac{1}{20}$ , there exists a positive constant  $\gamma := \gamma(\delta) = c(m, \Lambda)\delta^{2(1+\delta^{-10m-16})} > 0$  such that the following holds. Let  $u$  be a weak sub-solution of (1.5) in  $Q_1$  in the sense of Definition 3, assume that  $u \leq 1$  in  $Q_{\frac{1}{2}}$  and that*

$$|\{u \leq 0\} \cap Q_{r_0}^-| \geq \delta |Q_{r_0}^-|, \quad (3.19)$$

where  $Q_{r_0}^- := Q_{r_0}(0, 0, -2r_0^2)$ . Then

$$u \leq 1 - \gamma \text{ in } Q_{\frac{r_0}{2}}.$$

*Proof.* Using Remark 2.1, Theorem 1.2 and Lemma 3.5, the result follows by proceeding along the lines of the proof of [23, Lemma 16, page 12].  $\square$

#### 4. Proof of Theorem 1.1–Theorem 1.4

In this section we prove Theorem 1.1–Theorem 1.4. We first note that since our class of operators is closed under the group law defined in (1.10), and by our definition of  $Q_{r_0}(Z_0, t_0)$ , we can throughout the second without loss of generality assume that  $(Z_0, t_0) = 0$ . Note that  $Q_{r_0} = Q_{r_0}(0, 0) = V_X \times V_Y \times J$  where  $V_X = B(0, r_0)$ ,  $V_Y = B(0, r_0^3)$ ,  $J = (-r_0^2, 0)$ , and where  $B(0, \rho)$  denotes the standard Euclidean ball with center at 0 and radius  $\rho$  in  $\mathbb{R}^m$ .

##### 4.1. Proof of Theorem 1.1

As discussed in subsection 1.7, since  $u$  is a weak sub-solution of (1.5), there exists a non-negative measure  $\bar{\mu}$  such that

$$(\partial_t + X \cdot \nabla_Y)u = \nabla_X \cdot (A(\nabla_X u, X, Y, t)) - \bar{\mu}.$$

We define  $r_2 := \frac{r_0+r_1}{2}$ . Let  $\phi_1 \in [0, 1]$  be smooth such that  $\phi_1 \equiv 1$  in  $Q_{r_1}(Z_0, t_0)$  and  $\phi_1 \equiv 0$  outside  $Q_{r_2}(Z_0, t_0)$  satisfying

$$|\nabla_X \phi_1| \leq \frac{c}{r_0 - r_2}, \quad |\nabla_Y \phi_1| \leq \frac{c}{(r_0 - r_2)r_2^2}, \quad |\partial_t \phi_1| \leq \frac{c}{(r_0 - r_2)r_2}, \quad (4.1)$$

for some constant  $c = c(m) \geq 1$ . Then we observe that  $v = u\phi_1$  is a weak solution of

$$(\partial_t + X \cdot \nabla_Y - \Delta_X)v = \nabla_X \cdot F_1 + F_2 - \mu \quad \text{in} \quad \mathbb{R}^{N+1}, \quad (4.2)$$

where

$$F_1 = A(\nabla_X u)\phi_1 - \phi_1 \nabla_X u - u \nabla_X \phi_1,$$

$$F_2 = -A(\nabla_X u) \cdot \nabla_X \phi_1 + u(\partial_t + X \cdot \nabla_Y)\phi_1 \quad \text{and} \quad \mu = \bar{\mu}\phi_1.$$

By Lemma 3.3, we have

$$\|v\|_{L^q(\mathbb{R}^{2m} \times \mathbb{R}_-)} \leq c \left(2 + \frac{1}{m} - q\right)^{-1} (\|F_1\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|F_2\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)}), \quad (4.3)$$

and

$$\|v\|_{L^1_{t,X} W^{s,1}_Y(\mathbb{R}^{2m} \times \mathbb{R}_-)} \leq c \left(\frac{1}{3} - s\right)^{-1} (\|F_1\|_{L^1(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|F_2\|_{L^1(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|\mu\|_{M^1(\mathbb{R}^{2m} \times \mathbb{R}_-)}), \quad (4.4)$$

for some uniform constant  $c$  and for every  $q \in [2, 2 + \frac{1}{m})$  and  $s \in [0, \frac{1}{3})$ . Using (4.1), (1.7)-(i), Lemma 3.1 and that  $0 < r_1 < r_0 \leq 1$ , it follows that

$$\|F_1\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)} + \|F_2\|_{L^2(\mathbb{R}^{2m} \times \mathbb{R}_-)} \leq c_1, \quad (4.5)$$

where

$$c_1 = c(m, \Lambda) \left(1 + \frac{1}{r_0 - r_1}\right) \left(1 + \frac{1}{(r_0 - r_1)^2} + \frac{|X_0| + r_0}{(r_0 - r_1)r_1^2} + \frac{1}{(r_0 - r_1)r_1}\right). \quad (4.6)$$

Using (4.5) in (4.3), the estimate (1.15) follows. To obtain the estimate (1.16), let  $\phi_2 \in [0, 1]$  be smooth such that  $\phi_2 \equiv 1$  in  $Q_{r_2}(Z_0, t_0)$  and  $\phi_2 \equiv 0$  outside  $Q_{r_0}(Z_0, t_0)$  satisfying (4.1). Choosing  $\phi_2$  as a test function in (4.2) and proceeding similarly as in the proof of energy estimate in Lemma 3.1, we get

$$\|\mu\|_{M^1(Q_{r_2}(Z_0, t_0))} \leq \|\phi_2 \mu\|_{M^1(\mathbb{R}^{2m} \times \mathbb{R}_-)} \leq r_0^{1+2m} c_1 \|u\|_{L^2(Q_{r_0}(Z_0, t_0))},$$

where  $c_1$  is given by (4.6). The last estimate, combined with (4.4) and (4.5), yields the estimate (1.16).  $\square$

#### 4.2. Proof of Theorem 1.2

As mentioned, we can, without loss of generality, assume that  $(Z_0, t_0) = (0, 0, 0)$ . For  $n \in \mathbb{N} \cup \{0\}$ , we define

$$r_n = r_\infty + (r_0 - r_\infty)2^{-n}, \quad T_n = -r_n^2, \quad k_n = \frac{1}{2}(1 - 2^{-n}), \quad u_n = (u - k_n)^+,$$

and

$$A_n := \sup_{t \in (T_n, 0)} \iint_{B(0, r_n) \times B(0, r_n^3)} u_n^2(\cdot, \cdot, t) \, dX \, dY.$$

By Lemma 3.2 we know that  $u_n$  is a weak sub-solution of (1.5). Thus applying Lemma 3.1 we obtain

$$A_n \leq c c_{n-1, n} \iiint_{Q_{r_{n-1}}} u_n^2 \, dX \, dY \, dt, \quad \forall n \geq 1, \quad (4.7)$$

where  $c = c(m, \Lambda) \geq 1$  and

$$c_{n-1, n} := \frac{1}{(r_{n-1} - r_n)^2} + \frac{r_{n-1}}{(r_{n-1} - r_n)r_n^2} + \frac{1}{(r_{n-1} - r_n)r_n} + 1 \leq \frac{2^{2n}}{r_\infty^2 (r_0 - r_\infty)^2}. \quad (4.8)$$



Now we will estimate the integral in the right hand side of (4.7). Let  $q = 2 + \frac{1}{2m}$ . By Hölder's inequality we have

$$\iiint_{Q_{r_{n-1}}} u_n^2 \, dX \, dY \, dt \leq \left( \iiint_{Q_{r_{n-1}}} u_n^q \, dX \, dY \, dt \right)^{\frac{2}{q}} \left| \{u_n > 0\} \cap Q_{r_{n-1}} \right|^{1-\frac{2}{q}}. \quad (4.9)$$

Since  $k_n > k_{n-1}$ , we get  $u_n \leq u_{n-1}$ . Using this fact, that  $0 < r_\infty < r_0 \leq 1$ , and Theorem 1.1, we get

$$\begin{aligned} \left( \iiint_{Q_{r_{n-1}}} u_n^q \, dX \, dY \, dt \right)^{\frac{2}{q}} &\leq \left( \iiint_{Q_{r_{n-1}}} u_{n-1}^q \, dX \, dY \, dt \right)^{\frac{2}{q}} \\ &\leq c^2 A_{n-2} \\ &\leq \left( \frac{c(m, \Lambda) 2^{3n}}{r_\infty^2 (r_0 - r_\infty)^3} \right)^2 A_{n-2}, \end{aligned} \quad (4.10)$$

for every  $n \geq 2$ , where we have used that

$$c = c(m, \Lambda) \left( 1 + \frac{1}{r_{n-2} - r_{n-1}} \right) c_{n-2, n-1} \leq \frac{c(m, \Lambda) 2^{3n}}{r_\infty^2 (r_0 - r_\infty)^3},$$

with  $c_{n-2, n-1}$  is as defined in (4.8).

Next, we observe that

$$\begin{aligned} \iiint_{Q_{r_{n-1}}} u_{n-1}^2 \, dX \, dY \, dt &\geq \iiint_{\{u_{n-1} \geq 2^{-n-1}\} \cap Q_{r_{n-1}}} u_{n-1}^2 \, dX \, dY \, dt \\ &\geq 2^{-2n-2} \left| \{u_{n-1} \geq 2^{-n-1}\} \cap Q_{r_{n-1}} \right|. \end{aligned}$$

Moreover,

$$\left| \{u_n > 0\} \cap Q_{r_{n-1}} \right| \leq \left| \{u_n \geq k_n - k_{n-1}\} \cap Q_{r_{n-1}} \right| = \left| \{u_n \geq 2^{-n-1}\} \cap Q_{r_{n-1}} \right|.$$

Combining the preceding two estimates and using  $0 < r_\infty < r_0 \leq 1$ , we get

$$\left| \{u_n > 0\} \cap Q_{r_{n-1}} \right| \leq 2^{2n+2} A_{n-1}. \quad (4.11)$$

Using the estimates (4.10) and (4.11) in (4.9), we get

$$\iiint_{Q_{r_{n-1}}} u_n^2 \, dX \, dY \, dt \leq c(m, \Lambda) \left( \frac{2^{4n}}{r_\infty^2 (r_0 - r_\infty)^3} \right)^2 A_{n-2}^{2-\frac{2}{q}}, \quad \forall n \geq 2, \quad (4.12)$$

where we have also used that  $A_{n-1} \leq A_{n-2}$ . Note that the latter is true since  $u_{n-1} \leq u_{n-2}$ ,  $r_{n-1} < r_{n-2}$  and  $T_{n-2} < T_{n-1}$  for every  $n \geq 2$ . Using (4.12) in (4.7), we obtain

$$A_n \leq c(m, \Lambda) \frac{2^{12n}}{r_\infty^6 (r_0 - r_\infty)^8} A_{n-2}^\alpha,$$

where  $\alpha = 2 - \frac{2}{q} > 1$ , since  $q > 2$ . Therefore, defining  $S_n := A_{2n}$ , we get

$$S_n \leq \beta^n S_{n-1}^\alpha \quad \forall n \geq 1,$$

where

$$\beta = c(m, \Lambda) \frac{2^{24}}{r_\infty^6 (r_0 - r_\infty)^8}.$$

Recursively we get

$$\begin{aligned} S_n &\leq \beta^{n+(n-1)\alpha+\dots+\alpha^{n-1}} S_1^{\alpha^{n-1}} \\ &\leq \left( \beta^{\frac{\alpha^2}{(\alpha-1)^2}} S_1 \right)^{\alpha^{n-1}} \\ &\leq \left( c(m, \Lambda) c_{0,1} \beta^{\frac{\alpha^2}{(\alpha-1)^2}} \|u\|_{L^2(Q_{r_0})}^2 \right)^{\alpha^{n-1}}, \end{aligned} \quad (4.13)$$

where we have used (4.7) and the estimate

$$n + \alpha(n-1) + \dots + \alpha^{n-1} \leq \frac{\alpha^{n+1}}{(\alpha-1)^2}.$$

Let

$$v := \frac{1}{\sqrt{2c(m, \Lambda) c_{0,1} \beta^{\frac{\alpha^2}{(\alpha-1)^2}} \|u\|_{L^2(Q_{r_0})}}} u.$$

We observe that

$$\gamma := c(m, \Lambda) c_{0,1} \beta^{\frac{\alpha^2}{(\alpha-1)^2}} \|v\|_{L^2(Q_{r_0})}^2 = \frac{1}{2} < 1.$$

Note that, by the property (ii) in Remark 2.1,  $v$  is again a weak sub-solution of (1.5). Thus the estimate (4.13) holds by replacing  $u$  with  $v$ . This fact combined with  $\gamma < 1$  gives  $v \leq \frac{1}{2}$  a.e. in  $Q_{r_\infty}$ . As a consequence we get

$$\sup_{Q_{r_\infty}} u \leq \sqrt{2c(m, \Lambda) c_{0,1} \beta^{\frac{\alpha^2}{(\alpha-1)^2}} \|u\|_{L^2(Q_{r_0})}} \leq c \left( \frac{1}{r_\infty^2 (r_0 - r_\infty)^3} \right)^{\frac{\theta}{2}} \|u\|_{L^2(Q_{r_0})},$$

for some  $c = c(m, \Lambda) \geq 1$  and  $\theta = \theta(m) > 1$ . Now, arguing similarly as in the proof of [23, Proposition 12, pages 7–8], the result follows.  $\square$

#### 4.3. Proof of Theorem 1.3

Using Remark 2.1, Theorem 1.2 along with Lemma 3.6, and following the lines of the proof of [23, Theorem 5, pages 13–14], the result follows.  $\square$

#### 4.4. Proof of Theorem 1.4

Using Remark 2.1, Lemma 3.6, and following the lines of the proof of [23, Theorem 7, pages 14–15], the result follows.  $\square$

### 5. Proof of Theorem 1.5

The purpose of the section is to prove Theorem 1.5. As  $g \in W(U_X \times V_{Y,t})$  we can in the following assume, without loss of generality, that  $g \equiv 0$ .

In domains of the form  $U_X \times U_Y \times I$  instead of  $U_X \times V_{Y,t}$ , one may attempt different approaches to prove Theorem 1.5, and perhaps the most natural first approach is to add the term  $\epsilon \Delta_Y$  to the operator and to instead consider the problem

$$\begin{cases} \nabla_X \cdot (A(\nabla_X u_\epsilon, X, Y, t)) + \epsilon \Delta_Y u_\epsilon - (\partial_t + X \cdot \nabla_Y) u_\epsilon = g^* & \text{in } U_X \times U_Y \times I, \\ u_\epsilon = 0 & \text{on } \partial_p(U_X \times U_Y \times I). \end{cases} \quad (5.1)$$

Here  $\partial_p(U_X \times U_Y \times I)$  is now the (standard) parabolic boundary of  $U_X \times U_Y \times I$ , i.e.,

$$\partial_p(U_X \times U_Y \times I) := (\partial(U_X \times U_Y) \times \bar{I}) \cup ((U_X \times U_Y) \times \{0\}).$$

The existence and uniqueness of weak solutions to (5.1) is classical and one easily deduces that

$$\begin{aligned} & \| \nabla_X u_\epsilon \|_{L^2(U_X \times U_Y \times I)}^2 + \epsilon \| \nabla_Y u_\epsilon \|_{L^2(U_X \times U_Y \times I)}^2 \\ & \leq c \| g^* \|_{L^2_{Y,t}(U_Y \times I, H^1_X(U_X))} \times \| u_\epsilon \| + \| \nabla_X u_\epsilon \|_{L^2(U_X \times U_Y \times I)}, \end{aligned} \quad (5.2)$$

for some positive constant  $c$ , independent of  $\epsilon$ . By the standard Poincaré inequality, applied on  $U_X$  to  $u_\epsilon(\cdot, Y, t)$  with  $(Y, t)$  fixed, we have

$$\| u_\epsilon \|_{L^2(U_X \times U_Y \times I)} \leq c \| \nabla_X u_\epsilon \|_{L^2(U_X \times U_Y \times I)}. \quad (5.3)$$

Hence, using Cauchy-Schwarz we can conclude that

$$\begin{aligned} & \| u_\epsilon \|_{L^2(U_X \times U_Y \times I)}^2 + \| \nabla_X u_\epsilon \|_{L^2(U_X \times U_Y \times I)}^2 + \epsilon \| \nabla_Y u_\epsilon \|_{L^2(U_X \times U_Y \times I)}^2 \\ & \leq c \| g^* \|_{L^2_{Y,t}(U_Y \times I, H^1_X(U_X))}^2, \end{aligned} \quad (5.4)$$

for a constant  $c$  which is independent of  $\epsilon$ . The idea is then to let  $\epsilon \rightarrow 0$  and in this way construct a solution to the problem in (2.7). To make this operational, already in the linear case,  $A(\xi, X, Y, t) = A(X, Y, t)\xi$ , one seems to need some uniform estimates up to the Kolmogorov boundary  $\partial_{\mathcal{K}}(U_X \times U_Y \times I)$  to get a solution in the limit. In addition, in the nonlinear case considered in this paper we also need to ensure that  $\nabla_X u_\epsilon \rightarrow \nabla_X u$  pointwise a.e. as  $\epsilon \rightarrow 0$  and how to achieve this is even less clear. One approach is to try to adapt the techniques of Boccardo and Murat [6] but it seems unclear how to make this approach operational in our case due to the presence of the term  $\epsilon \Delta_Y u_\epsilon$  in the approximating equation.

In this paper we will instead prove Theorem 1.5 by using a variational approach recently explored in Albritton-Armstrong-Mourrat-Novack [1] and Litsgård-Nyström [29]. We will prove that the solution to (1.21) can be obtained as the minimizer of a uniformly convex functional. The fact that a parabolic equation can be cast as the first variation of a uniformly convex integral functional was first discovered by Brezis-Ekeland [9, 10] and for a modern treatment of this approach, covering uniformly elliptic parabolic equations of second order in the more general context of uniformly monotone operators, we refer to [4] which in turn is closely related to [19], see also [18].

### 5.1. Variational representation of the symbol

To make the approach operational we will use a variational representation of the mapping  $\xi \mapsto A(\xi, X, Y, t)$ , for each  $(X, Y, t) \in \mathbb{R}^{N+1}$ , that we learned from [4] and [5] and we refer to these papers for

more background. Indeed, by [5, Theorem 2.9], there exists  $\tilde{A} \in L_{\text{loc}}^\infty(\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{N+1})$  satisfying the following properties, for  $\Gamma := 2\Lambda + 1$  and for each  $(X, Y, t) \in \mathbb{R}^{N+1}$ . First, the mapping

$$(\xi, \eta) \mapsto \tilde{A}(\xi, \eta, X, Y, t) - \frac{1}{2\Gamma}(|\xi|^2 + |\eta|^2) \quad \text{is convex.} \quad (5.5)$$

Second, the mapping

$$(\xi, \eta) \mapsto \tilde{A}(\xi, \eta, X, Y, t) - \frac{\Gamma}{2}(|\xi|^2 + |\eta|^2) \quad \text{is concave.} \quad (5.6)$$

Third, for every  $\xi, \eta \in \mathbb{R}^m$ , we have

$$\tilde{A}(\xi, \eta, X, Y, t) \geq \xi \cdot \eta, \quad (5.7)$$

and

$$\tilde{A}(\xi, \eta, X, Y, t) = \xi \cdot \eta \iff \eta = A(\xi, X, Y, t). \quad (5.8)$$

Note that the choice of  $\tilde{A}$  is in general not unique. Note also that (5.5) and (5.6) imply, in particular that

$$\begin{aligned} \frac{1}{2\Gamma}|\xi_1 - \xi_2|^2 &\leq \frac{1}{2}\tilde{A}(\xi_1, \eta, X, Y, t) + \frac{1}{2}\tilde{A}(\xi_2, \eta, X, Y, t) \\ &\quad - \tilde{A}\left(\frac{1}{2}\xi_1 + \frac{1}{2}\xi_2, \eta, X, Y, t\right) \leq \frac{\Gamma}{2}|\xi_1 - \xi_2|^2. \end{aligned} \quad (5.9)$$

## 5.2. Setting up the argument

To ease the notation we will in the following at instances use the notation

$$W := W(U_X \times V_{Y,t}), \quad W_0 := W_0(U_X \times V_{Y,t}),$$

and we let

$$\mathcal{L}u := \nabla_X \cdot (A(\nabla_X u, X, Y, t)) - (\partial_t + X \cdot \nabla_Y)u.$$

Given an arbitrary pair  $(f, \mathbf{j})$  such that

$$f \in L_{Y,t}^2(V_{Y,t}, H_X^1(U_X)) \quad \text{and} \quad \mathbf{j} \in L^2(V_{Y,t}, L^2(U_X))^m, \quad (5.10)$$

we introduce

$$\mathcal{J}[f, \mathbf{j}] := \iiint_{U_X \times V_{Y,t}} (\tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) - \nabla_X f \cdot \mathbf{j}) \, dX \, dY \, dt. \quad (5.11)$$

Using this notation, and given an arbitrary pair  $(f, f^*)$  such that

$$f \in L_{Y,t}^2(V_{Y,t}, H_X^1(U_X)) \quad \text{and} \quad f^*, f^* + (\partial_t + X \cdot \nabla_Y)f \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)), \quad (5.12)$$

we set

$$\mathcal{J}[f, f^*] := \inf \iiint_{U_X \times V_{Y,t}} \mathcal{J}[f, \mathbf{g}] \, dX \, dY \, dt, \quad (5.13)$$

where the infimum is taken with respect to the set

$$\{\mathbf{g} \in (L^2(V_{Y,t}, L^2(U_X)))^m \mid \nabla_X \cdot \mathbf{g} = f^* + (\partial_t + X \cdot \nabla_Y)f\}. \quad (5.14)$$

The condition

$$\nabla_X \cdot \mathbf{g} = f^* + (\partial_t + X \cdot \nabla_Y)f,$$

appearing in (5.14), should be interpreted as stating that

$$- \iiint_{U_X \times V_{Y,t}} \mathbf{g} \cdot \nabla_X \phi \, dX \, dY \, dt = \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), \phi \rangle \, dY \, dt, \quad (5.15)$$

for all  $\phi \in L^2(V_{Y,t}, H_{X,0}^1(U_X))$ . Finally, for  $g^* \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))$  fixed we introduce

$$\mathcal{A}(g^*) := \{(f, \mathbf{j}) \in W_0 \times (L^2(V_{Y,t}, L^2(U_X)))^m \mid \nabla_X \cdot \mathbf{j} = g^* + (\partial_t + X \cdot \nabla_Y)f\}. \quad (5.16)$$

5.3.  $\mathcal{J}$  is uniformly convex on  $\mathcal{A}(g^*)$

**Lemma 5.1.** *Let  $g^* \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))$  be fixed and let  $\mathcal{A}(g^*)$  be the set introduced in (5.16). Then  $\mathcal{A}(g^*)$  is non-empty.*

*Proof.* Take  $f \in W_0$  and consider the equation

$$\Delta_X v(X, Y, t) = (g^*(X, Y, t) + (\partial_t + X \cdot \nabla_Y)f(X, Y, t)) \in H_X^{-1}(U_X), \quad (5.17)$$

for  $dY \, dt$ -a.e.  $(Y, t) \in V_{Y,t}$ . By the Lax-Milgram theorem this equation has a (unique) solution  $v(\cdot) = v(\cdot, Y, t) \in H_{X,0}^1(U_X)$  and

$$\|\nabla_X v\|_{L_{Y,t}^2(V_{Y,t}, L^2(U_X))} \leq c \|g^* + (\partial_t + X \cdot \nabla_Y)f\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))} < \infty, \quad (5.18)$$

as  $f \in W_0$ . In particular,

$$(f, \nabla_X v) \in \mathcal{A}(g^*), \quad (5.19)$$

and hence  $\mathcal{A}(g^*)$  is non-empty.  $\square$

**Lemma 5.2.** *The functional  $\mathcal{J}$  introduced in (5.11) is uniformly convex on  $\mathcal{A}(g^*)$ .*

*Proof.* Note that if  $(f, \mathbf{j}) \in \mathcal{A}(g^*)$  and  $(\tilde{f}, \tilde{\mathbf{j}}) \in \mathcal{A}(0)$ , then  $(f + \tilde{f}, \mathbf{j} + \tilde{\mathbf{j}}) \in \mathcal{A}(g^*)$  and  $(f - \tilde{f}, \mathbf{j} - \tilde{\mathbf{j}}) \in \mathcal{A}(g^*)$ . Consider  $(f, \mathbf{j}) \in \mathcal{A}(g^*)$ . We first consider the term

$$- \iiint_{U_X \times V_{Y,t}} \nabla_X f \cdot \mathbf{j} \, dX \, dY \, dt.$$

We have

$$\begin{aligned} - \iiint_{U_X \times V_{Y,t}} \nabla_X f \cdot \mathbf{j} \, dX \, dY \, dt &= \iiint_{U_X \times V_{Y,t}} f \nabla_X \cdot \mathbf{j} \, dX \, dY \, dt \\ &= \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \\ &= \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \\ &\quad + \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt. \end{aligned} \quad (5.20)$$

Recall that  $W_0 = W_0(U_X \times V_{Y,t})$  is the closure in the norm of  $W(U_X \times V_{Y,t})$  of  $C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ . In particular, there exists  $\{f_j\}$ ,  $f_j \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$  such that

$$\|f - f_j\|_W \rightarrow 0 \text{ as } j \rightarrow \infty,$$

and consequently

$$\|(\partial_t + X \cdot \nabla_Y)(f - f_j)\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))} \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Using this we see that

$$\begin{aligned} & \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \\ & \geq \liminf_{j \rightarrow \infty} \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f_j(\cdot, Y, t), f_j(\cdot, Y, t) \rangle \, dY \, dt. \end{aligned} \quad (5.21)$$

However, using that  $f_j \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$  we see that

$$\begin{aligned} & \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f_j(\cdot, Y, t), f_j(\cdot, Y, t) \rangle \, dY \, dt \\ & = \iiint_{U_X \times V_{Y,t}} (\partial_t + X \cdot \nabla_Y)f_j f_j \, dX \, dY \, dt \\ & = \frac{1}{2} \iiint_{U_X \times V_{Y,t}} (\partial_t + X \cdot \nabla_Y)f_j^2 \, dX \, dY \, dt \\ & = \frac{1}{2} \int_{U_X} \iint_{\partial V_{Y,t}} f_j^2(X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \geq 0, \end{aligned} \quad (5.22)$$

by the divergence theorem and the definition of the Kolmogorov boundary. Hence,

$$\begin{aligned} - \iiint_{U_X \times V_{Y,t}} \nabla_X f \cdot \mathbf{j} \, dX \, dY \, dt & = \iiint_{U_X \times V_{Y,t}} f \nabla_X \cdot \mathbf{j} \, dX \, dY \, dt \\ & \geq \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt. \end{aligned} \quad (5.23)$$

Using this, and observing that,

$$\begin{aligned} & - \frac{1}{2} \iiint_{U_X \times V_{Y,t}} (\nabla_X(f + \tilde{f}) \cdot (\mathbf{j} + \tilde{\mathbf{j}}) + \nabla_X(f - \tilde{f}) \cdot (\mathbf{j} - \tilde{\mathbf{j}}) - 2\nabla_X f \cdot \mathbf{j}) \, dX \, dY \, dt \\ & = - \iiint_{U_X \times V_{Y,t}} \nabla_X \tilde{f} \cdot \tilde{\mathbf{j}} \, dX \, dY \, dt, \end{aligned} \quad (5.24)$$

we can conclude that

$$- \frac{1}{2} \iiint_{U_X \times V_{Y,t}} (\nabla_X(f + \tilde{f}) \cdot (\mathbf{j} + \tilde{\mathbf{j}}) + \nabla_X(f - \tilde{f}) \cdot (\mathbf{j} - \tilde{\mathbf{j}}) - 2\nabla_X f \cdot \mathbf{j}) \, dX \, dY \, dt \geq 0, \quad (5.25)$$

over the set  $\mathcal{A}(g^*)$ . Hence it suffices to prove that

$$\iiint_{U_X \times V_{Y,t}} \tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) \, dX \, dY \, dt$$

is uniformly convex over the set  $\mathcal{A}(g^*)$ . With  $(f, \mathbf{j}) \in \mathcal{A}(g^*)$  and  $(\tilde{f}, \tilde{\mathbf{j}}) \in \mathcal{A}(0)$  as above, (5.5) implies that

$$\frac{1}{2}\tilde{A}(\nabla_X(f + \tilde{f}), \mathbf{j} + \tilde{\mathbf{j}}, \cdot) + \frac{1}{2}\tilde{A}(\nabla_X(f - \tilde{f}), \mathbf{j} - \tilde{\mathbf{j}}, \cdot) - \tilde{A}(\nabla_X f, \mathbf{j}, \cdot) \geq \frac{1}{2\Gamma} (\|\nabla_X \tilde{f}\|^2 + \|\tilde{\mathbf{j}}\|^2).$$

We also have

$$\|(\partial_t + X \cdot \nabla_Y)\tilde{f}\|_{L^2_{Y,t}(V_{Y,t}, H^{-1}(U_X))} \leq \|\tilde{\mathbf{j}}\|_{L^2(U_X \times V_{Y,t})}.$$

Thus

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} \left( \frac{1}{2}\tilde{A}(\nabla_X(f + \tilde{f}), \mathbf{j} + \tilde{\mathbf{j}}, \cdot) + \frac{1}{2}\tilde{A}(\nabla_X(f - \tilde{f}), \mathbf{j} - \tilde{\mathbf{j}}, \cdot) - \tilde{A}(\nabla_X f, \mathbf{j}, \cdot) \right) dX dY dt \\ & \geq \frac{1}{4\Gamma} \left( \|\nabla_X \tilde{f}\|_{L^2(U_X \times V_{Y,t})}^2 + \|(\partial_t + X \cdot \nabla_Y)\tilde{f}\|_{L^2_{Y,t}(V_{Y,t}, H^{-1}(U_X))}^2 + \|\tilde{\mathbf{j}}\|_{L^2(U_X \times V_{Y,t})}^2 \right) \\ & \geq \frac{1}{4c\Gamma} \left( \|\tilde{f}\|_{W(U_X \times V_{Y,t})}^2 + \|\tilde{\mathbf{j}}\|_{L^2(U_X \times V_{Y,t})}^2 \right), \end{aligned}$$

by using the (standard) Poincaré inequality. Hence  $\mathcal{J}$  is uniformly convex on  $\mathcal{A}(g^*)$ .  $\square$

#### 5.4. Correspondence between weak solutions and minimizers

As the functional  $\mathcal{J}$  is uniformly convex over  $\mathcal{A}(g^*)$  there exists a unique minimizing pair  $(f_1, \mathbf{j}_1) \in \mathcal{A}(g^*)$  such that

$$\begin{aligned} (f_1, \mathbf{j}_1) & := \arg \min_{(f, \mathbf{j}) \in \mathcal{A}(g^*)} \mathcal{J}[f, \mathbf{j}] \\ & = \arg \min_{(f, \mathbf{j}) \in \mathcal{A}(g^*)} \iiint_{U_X \times V_{Y,t}} (\tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) - \nabla_X f \cdot \mathbf{j}) dX dY dt. \end{aligned}$$

Note that

$$\min_{(f, \mathbf{j}) \in \mathcal{A}(g^*)} \mathcal{J}[f, \mathbf{j}] = \min_{f \in W_0} J[f, g^*].$$

Moreover, by construction of  $\tilde{A}$ , see (5.7), we have

$$J[f_1, g^*] \geq 0. \tag{5.26}$$

**Lemma 5.3.** *There is a one-to-one correspondence between weak solutions in the sense of Eq (2.7) to  $\mathcal{L}u = g^*$  in  $U_X \times V_{Y,t}$ , such that  $u \in W_0$ , and null minimizers of  $J[\cdot, g^*]$ .*

*Proof.* To prove the lemma we need to prove that for every  $f \in W_0$ , we have

$$f \text{ solves } \mathcal{L}u = g^* \text{ in the weak sense in } U_X \times V_{Y,t} \iff J[f, g^*] = 0.$$

Indeed, the implication “ $\implies$ ” is clear since if  $f$  solves  $\mathcal{L}u = g^*$  in the weak sense, then

$$(f, A(\nabla_X f, X, Y, t)) \in \mathcal{A}(g^*) \quad \text{and} \quad \mathcal{J}[f, A(\nabla_X f, X, Y, t)] = 0 = J[f, g^*].$$

Conversely, if  $J[f, g^*] = 0$ , then  $f = f_1$  and

$$\mathcal{J}[f_1, \mathbf{j}_1] = \iiint_{U_X \times V_{Y,t}} (\tilde{A}(\nabla_X f_1, \mathbf{j}_1, X, Y, t) - \nabla_X f_1 \cdot \mathbf{j}_1) \, dX \, dY \, dt = 0. \quad (5.27)$$

Using (5.8), we see that the identity (5.27) implies that

$$\mathbf{j}_1 = A(\nabla f_1, \cdot, \cdot, \cdot) \quad \text{a.e. in } U_X \times V_{Y,t},$$

and by the definition of the set  $\mathcal{A}(g^*)$ ,

$$\nabla_X \cdot \mathbf{j}_1 = g^* + (\partial_t + X \cdot \nabla_Y) f_1.$$

Hence  $f_1$  indeed solves

$$\nabla_X \cdot A(\nabla f_1, \cdot, \cdot, \cdot) - (\partial_t + X \cdot \nabla_Y) f_1 = g^*$$

in the weak sense. I.e., we recover that  $f = f_1$  is indeed a weak solution of  $\mathcal{L}u = g^*$ . In particular, the fact that there is at most one solution to  $\mathcal{L}u = g^*$  is clear.  $\square$

### 5.5. An associated perturbed convex minimization problem

Using (5.26) and Lemma 5.3 we see that to complete the proof of Theorem 1.5 it remains to prove that

$$J[f_1, g^*] \leq 0. \quad (5.28)$$

In order to do so, we introduce the perturbed convex minimization problem defined, for every  $f^* \in L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))$ , by

$$G(f^*) := \inf_{f \in W_0} (J[f, f^* + g^*] - \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt).$$

As

$$G(0) = \inf_{f \in W_0} J[f, g^*],$$

we see that to prove (5.28) it suffices to prove that  $G(0) \leq 0$ .

**Lemma 5.4.**  *$G$  is a convex, locally bounded from above and lower semi-continuous functional on  $L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))$ .*

*Proof.* For every pair  $(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)$ , we have

$$\nabla_X \cdot \mathbf{j} = f^* + g^* + (\partial_t + X \cdot \nabla_Y) f,$$

and thus

$$\begin{aligned} \mathcal{J}[f, \mathbf{j}] &= \iiint_{U_X \times V_{Y,t}} (\tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) - \nabla_X f \cdot \mathbf{j}) \, dX \, dY \, dt \\ &= \iiint_{U_X \times V_{Y,t}} \tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) \, dX \, dY \, dt + \iint_{V_{Y,t}} \langle (f^* + g^*)(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \end{aligned}$$



$$+ \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt.$$

Hence

$$\begin{aligned} \mathcal{J}[f, \mathbf{j}] &= \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\ &= \iiint_{U_X \times V_{Y,t}} \tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) dX dY dt + \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\ &+ \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt. \end{aligned}$$

Taking the infimum over all  $(f, \mathbf{j})$  satisfying the affine constraint  $(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)$  we obtain the quantity  $G(f^*)$ , i.e.,  $G(f^*)$  can be expressed as

$$G(f^*) = \inf_{(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)} \left( \mathcal{J}[f, \mathbf{j}] - \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \right).$$

In particular,  $G(f^*)$  can be expressed as the infimum of

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} \tilde{A}(\nabla_X f, \mathbf{j}, X, Y, t) dX dY dt + \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\ & + \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \end{aligned} \quad (5.29)$$

with respect to  $(f, \mathbf{j})$  such that  $(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)$ . We now recall the argument in (5.21) and (5.22). In particular, given  $f \in W_0$  there exists  $\{f_j\}$ ,  $f_j \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$  such that

$$\|f - f_j\|_W \rightarrow 0 \text{ as } j \rightarrow \infty, \quad (5.30)$$

and consequently

$$\|(\partial_t + X \cdot \nabla_Y)(f - f_j)\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))} \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Using (5.21) and (5.22) we have

$$\begin{aligned} & \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\ & = \lim_{j \rightarrow \infty} \frac{1}{2} \int_{U_X} \iint_{\partial V_{Y,t}} f_j^2(X, 1) \cdot N_{Y,t} d\sigma_{Y,t} dX. \end{aligned}$$

Obviously we get the same limit in (5.31) independent of what sequence  $\{f_j\}$  chosen as long as (5.30) holds. Now consider  $f, g \in W_0$  and let  $\{f_j\}$ ,  $f_j \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ ,  $\{g_j\}$ ,  $g_j \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ , be such that

$$\|f - f_j\|_W + \|g - g_j\|_W \rightarrow 0 \text{ as } j \rightarrow \infty, \quad (5.31)$$

Then

$$\|(\tau f + (1 - \tau)g) - (\tau f_j + (1 - \tau)g_j)\|_W \rightarrow 0 \text{ as } j \rightarrow \infty, \quad (5.32)$$

for all  $\tau \in [0, 1]$ . Hence

$$\begin{aligned}
& \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)(\tau f + (1 - \tau)g)(\cdot, Y, t), (\tau f + (1 - \tau)g)(\cdot, Y, t) \rangle dY dt \\
&= \lim_{j \rightarrow \infty} \frac{1}{2} \int_{U_X} \iint_{\partial V_{Y,t}} (\tau f_j + (1 - \tau)g_j)^2 |(X, 1) \cdot N_{Y,t}| d\sigma_{Y,t} dX \\
&\leq \lim_{j \rightarrow \infty} \frac{1}{2} \int_{U_X} \iint_{\partial V_{Y,t}} \tau f_j^2 |(X, 1) \cdot N_{Y,t}| d\sigma_{Y,t} dX \\
&+ \lim_{j \rightarrow \infty} \frac{1}{2} \int_{U_X} \iint_{\partial V_{Y,t}} (1 - \tau)g_j^2 |(X, 1) \cdot N_{Y,t}| d\sigma_{Y,t} dX,
\end{aligned} \tag{5.33}$$

and we deduce that

$$\begin{aligned}
& \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)(\tau f + (1 - \tau)g)(\cdot, Y, t), (\tau f + (1 - \tau)g)(\cdot, Y, t) \rangle dY dt \\
&\leq \tau \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\
&+ (1 - \tau) \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)g(\cdot, Y, t), g(\cdot, Y, t) \rangle dY dt.
\end{aligned} \tag{5.34}$$

In particular, we can conclude that the mapping

$$f \rightarrow \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)(\tau f + (1 - \tau)g)(\cdot, Y, t), (\tau f + (1 - \tau)g)(\cdot, Y, t) \rangle dY dt$$

is convex on  $W_0$ . Using this, and (5.5), we see that the expression in (5.29) is convex as a function of  $(f, f^*, \mathbf{j})$  and this proves that  $G$  is convex. Furthermore, using (5.17)–(5.19) we can conclude that the infimum of the expression in (5.29) is finite, hence  $G(f^*) < \infty$ . In particular, the function  $G$  is locally bounded from above. These two properties imply that  $G$  is lower semi-continuous, see [17, Lemma 2.1 and Corollary 2.2].  $\square$

### 5.6. The convex dual of $G$

We denote by  $G^*$  the convex dual of  $G$ , defined for every

$$h \in (L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X)))^* = L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X)),$$

as

$$G^*(h) := \sup_{f^* \in L^2_{Y,t}(V_{Y,t}, H_X^{-1}(U_X))} \left( -G(f^*) + \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t), h(\cdot, Y, t) \rangle dY dt \right).$$

Let  $G^{**}$  be the bidual of  $G$ . Since  $G$  is lower semi-continuous, we have that  $G^{**} = G$  (see [17, Proposition 4.1]), and in particular,

$$G(0) = G^{**}(0) = \sup_{h \in L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X))} (-G^*(h)).$$

In order to prove that  $G(0) \leq 0$ , it therefore suffices to show that

$$G^*(h) \geq 0 \text{ for all } h \in L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X)). \quad (5.35)$$

To continue we note that we can rewrite  $G^*(h)$  as

$$G^*(h) = \sup_{(f, \mathbf{j}, f^*)} \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X f, \mathbf{j}, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j})) \, dX \, dY \, dt \right. \\ \left. + \iint_{V_{Y,t}} \langle f^*(\cdot, Y, t), (h(\cdot, Y, t) + f(\cdot, Y, t)) \rangle \, dY \, dt \right\}, \quad (5.36)$$

where the supremum is taken with respect to

$$(f, \mathbf{j}, f^*) \in W_0 \times (L^2_{Y,t}(V_{Y,t}, L^2_X(U_X)))^m \times L^2_{Y,t}(V_{Y,t}, H^{-1}_X(U_X)),$$

subject to the constraint

$$\nabla_X \cdot \mathbf{j} = f^* + g^* + (\partial_t + X \cdot \nabla_Y)f. \quad (5.37)$$

Furthermore, note that for every  $h \in L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X))$ , we have  $G^*(h) \in \mathbb{R} \cup \{+\infty\}$ .

**Lemma 5.5.** Consider  $h \in L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X))$ . Then

$$G^*(h) < +\infty \implies h \in W \cap L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X)). \quad (5.38)$$

*Proof.* To prove the lemma we need to prove that  $(\partial_t + X \cdot \nabla_Y)h \in L^2_{Y,t}(V_{Y,t}, H^{-1}_X(U_X))$ . Using that we take a supremum in the definition of  $G^*$  we can develop lower bounds on  $G^*$  by restricting the set with respect to which we take the supremum. Here, for  $f \in W_0$ , we choose to restrict the supremum to  $(f, \mathbf{j}, f^*)$  where  $\mathbf{j} = \mathbf{j}_0$  is a solution of  $\nabla_X \cdot \mathbf{j}_0 = g^*$  and  $f^* := -(\partial_t + X \cdot \nabla_Y)f$ . Recall from (5.19) that such a  $\mathbf{j}_0 \in (L^2_{Y,t}(V_{Y,t}, L^2_X(U_X)))^m$  exists. With these choices for  $\mathbf{j}$  and  $f^*$ , the constraint (5.37) is satisfied, and we obtain that

$$G^*(h) \geq \sup_{f \in W_0} \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X f, \mathbf{j}_0, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j}_0)) \, dX \, dY \, dt \right. \\ \left. - \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), (h(\cdot, Y, t) + f(\cdot, Y, t)) \rangle \, dY \, dt \right\}.$$

Consider  $f \in C^\infty_{\mathcal{K},0}(\overline{U_X \times V_{Y,t}}) \subset W_0$ . Then, again arguing as in (5.21), (5.22),

$$- \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \leq 0.$$

Furthermore, restricting to  $f \in C^\infty_0(U_X \times V_{Y,t}) \subset C^\infty_{\mathcal{K},0}(\overline{U_X \times V_{Y,t}})$  yields by the same argument that

$$- \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt = 0.$$

Hence we have the lower bound

$$G^*(h) \geq \sup \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X f, \mathbf{j}_0, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j}_0)) \, dX \, dY \, dt \right.$$

$$- \left. \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), h(\cdot, Y, t) \rangle dY dt \right\},$$

where the supremum now is taken with respect to  $f \in C_0^\infty(U_X \times V_{Y,t}) \subset C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ . Moreover, as  $G^*(h) < +\infty$ , we have that

$$\begin{aligned} & - \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), h(\cdot, Y, t) \rangle dY dt \\ & \leq \iiint_{U_X \times V_{Y,t}} (\tilde{A}(\nabla_X f, \mathbf{j}_0, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j}_0)) dX dY dt + G^*(h) < \infty, \end{aligned}$$

for every  $f \in C_0^\infty(U_X \times V_{Y,t})$  fixed. Note that by replacing  $f$  with  $-f$  in the above argument we also obtain a lower bound. In particular,

$$\sup \left| \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) h(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \right| < \infty,$$

where the supremum is taken over  $f \in C_0^\infty(U_X \times V_{Y,t})$  such that  $\|f\|_{L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))} \leq 1$ . Using that  $C_0^\infty(U_X \times V_{Y,t})$  is dense in  $L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))$  we can conclude that

$$(\partial_t + X \cdot \nabla_Y) h \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))$$

and this observation proves (5.38).  $\square$

### 5.7. Bounding $G^*$ from below

Lemma 5.5 gives at hand that in place of (5.35), we have reduced the matter to proving that

$$G^*(h) \geq 0 \text{ for all } h \in W \cap L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X)). \quad (5.39)$$

Furthermore, note that for  $\tilde{h} \in W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})$  we have

$$G^*(h) \geq G^*(\tilde{h}) - \|f^*\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))} \|h - \tilde{h}\|_{L_{Y,t}^2(V_{Y,t}, H_X^1(U_X))}. \quad (5.40)$$

As we are to establish a lower bound on  $G^*$ , we may restrict to taking the supremum over  $f^*$  such that

$$\|f^*\|_{L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X))} \leq 1. \quad (5.41)$$

In Lemma 5.6 below we prove that

$$G^*(h) \geq 0 \text{ for all } h \in W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}}).$$

By combining this with (5.40) and (5.41) we see that

$$G^*(h) \geq G^*(\tilde{h}) - \|h - \tilde{h}\|_{L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))} \geq -\|h - \tilde{h}\|_{L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))},$$

for all  $\tilde{h} \in W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})$ . Furthermore, by the definitions of  $W$ , and  $L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))$ , we can choose a sequence  $h_j \in W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})$  such that

$$\lim_{j \rightarrow \infty} \|h - h_j\|_{L_{Y,t}^2(V_{Y,t}, H_{X,0}^1(U_X))} = 0.$$

Hence the proof that  $G^*(h) \geq 0$ , and hence the final piece in the proof of existence in Theorem 1.5, is to prove the following lemma.

**Lemma 5.6.**

$$G^*(h) \geq 0 \text{ for all } h \in W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}}). \quad (5.42)$$

*Proof.* To start the proof of the lemma we first note that we have, as  $f \in W_0$ , that

$$(\partial_t + X \cdot \nabla_Y)f \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)),$$

and hence we can replace  $f^*$  by  $f^* - (\partial_t + X \cdot \nabla_Y)f$  in the variational formula (5.36) for  $G^*$  to get

$$G^*(h) \geq \sup_{(f, \mathbf{j}, f^*)} \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X f, \mathbf{j}, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j})) \, dX \, dY \, dt \right. \\ \left. + \iint_{V_{Y,t}} \langle (f^* - (\partial_t + X \cdot \nabla_Y)f)(\cdot, Y, t), (h(\cdot, Y, t) + f(\cdot, Y, t)) \rangle \, dY \, dt \right\},$$

where the supremum now is taken with respect to

$$(f, \mathbf{j}, f^*) \in (W \cap C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})) \times (L_{Y,t}^2(V_{Y,t}, L_X^2(U_X)))^m \times L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)), \quad (5.43)$$

subject to the constraint

$$\nabla_X \cdot \mathbf{j} = f^* + g^*. \quad (5.44)$$

Next using that  $f \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ ,  $h \in C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})$ , we have

$$\iint_{V_{Y,t}} -\langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), (h(\cdot, Y, t) + f(\cdot, Y, t)) \rangle \\ = \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)h(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \\ - \int_{U_X} \iint_{\partial V_{Y,t}} \left( \frac{1}{2}f^2 + fh \right)(X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX.$$

Using the identity in the last display we see that

$$G^*(h) \geq \sup_{(f, \mathbf{j}, f^*)} \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X f, \mathbf{j}, \cdot, \cdot, \cdot) - (\nabla_X f \cdot \mathbf{j})) \, dX \, dY \, dt \right. \\ \left. + \iint_{V_{Y,t}} \langle f^*, (h(\cdot, Y, t) + f(\cdot, Y, t)) \rangle + \langle (\partial_t + X \cdot \nabla_Y)h(\cdot, Y, t), f(\cdot, Y, t) \rangle \, dY \, dt \right. \\ \left. - \int_{U_X} \iint_{\partial V_{Y,t}} \left( \frac{1}{2}f^2 + fh \right)(X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \right\}, \quad (5.45)$$

where the supremum still is with respect to  $(f, \mathbf{j}, f^*)$  as in (5.43) subject to (5.44). Now, by arguing exactly as in the passage between displays (3.23) and (3.26) in [29], using the properties of  $\tilde{A}$ , we can conclude that it suffices to prove that  $\tilde{G}^*(h) \geq 0$  where

$$\tilde{G}^*(h) := \sup_{(\tilde{f}, \mathbf{j}, f^*, b)} \left\{ \iiint_{U_X \times V_{Y,t}} -(\tilde{A}(\nabla_X \tilde{f}, \mathbf{j}, \cdot, \cdot, \cdot) - (\nabla_X \tilde{f} \cdot \mathbf{j})) \, dX \, dY \, dt \right. \\ \left. + \iint_{V_{Y,t}} \langle f^*, (h(\cdot, Y, t) + \tilde{f}(\cdot, Y, t)) \rangle + \langle (\partial_t + X \cdot \nabla_Y)h(\cdot, Y, t), \tilde{f}(\cdot, Y, t) \rangle \, dY \, dt \right.$$

$$- \int_{U_X} \iint_{\partial V_{Y,t}} \left( \frac{1}{2} b^2 + bh \right) (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \Big\},$$

and where the supremum is taken with respect to all  $(\tilde{f}, \mathbf{j}, f^*, b)$  in the set

$$(W \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})) \times (L_{Y,t}^2(V_{Y,t}, L_X^2(U_X)))^m \times L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)) \times C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}}),$$

subject to the condition stated in [29], i.e., that

$$\Gamma(\tilde{f}, b) := \|\tilde{f}\|_{L_{Y,t}^2(V_{Y,t}, H_X^1(U_X))} + \|b\|_{L_{Y,t}^2(V_{Y,t}, H_X^1(U_X))} \leq \Gamma$$

for some large but fixed  $\Gamma \geq 1$ . However, this implies that  $\tilde{f} := -h$  is an admissible function. With this choice of  $\tilde{f}$ , we then let  $\mathbf{j} := A(-\nabla_X h, X, Y, t) \in (L_{Y,t}^2(V_{Y,t}, L_X^2(U_X)))^m$  and then

$$f^* = \nabla_X \cdot \mathbf{j} - g^* \in L_{Y,t}^2(V_{Y,t}, H_X^{-1}(U_X)).$$

Using this we deduce that

$$\tilde{G}^*(h) \geq \sup_b \left\{ - \int_{U_X} \iint_{\partial V_{Y,t}} \frac{1}{2} (b+h)^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \right\},$$

where supremum now is taken with respect to  $b \in C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ . Using Lemma 5.7 below it follows that

$$\sup_b \left\{ - \int_{U_X} \iint_{\partial V_{Y,t}} \frac{1}{2} (b+h)^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \right\} \geq 0.$$

The proof of the lemma is therefore complete.  $\square$

**Lemma 5.7.** Assume that  $h \in W(U_X \times V_{Y,t}) \cap C_{X,0}^\infty(\overline{U_X \times V_{Y,t}})$ . Then

$$\sup_{b \in W \cap C_{\mathcal{K},0}^\infty(U_X \times V_{Y,t})} - \iiint_{U_X \times \partial V_{Y,t}} (b+h)^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \geq 0. \quad (5.46)$$

Lemma 5.7 is Lemma 3.7 in [29] and in the next subsection we supply parts of the proof for completion.

### 5.8. Proof of Lemma 5.7

Let  $\psi(s) \in C^\infty(\mathbb{R})$  be such that  $0 \leq \psi \leq 1$ ,

$$\psi \equiv 1 \text{ on } [0, 1], \quad \psi \equiv 0 \text{ on } [2, \infty),$$

$|\psi'| \leq 2$  and such that  $\sqrt{1 - \psi^2} \in C^\infty(\mathbb{R})$ . Based on  $\psi$  we introduce for  $r, 0 \leq r < \infty$

$$\psi_r(X, Y, t) := \psi \left( r \frac{((X, 1) \cdot N_{Y,t})^+}{1 + |X|^2} \right), \quad (5.47)$$

where we use the notation  $s^+ := \max\{s, 0\}$  for  $s \in \mathbb{R}$ . As  $h$  is smooth, and  $U_X$  and  $V_{Y,t}$  are bounded domains, we have

$$\iiint_{U_X \times \partial V_{Y,t}} h^2 |(X, 1) \cdot N_{Y,t}| \, dX \, d\sigma_{Y,t} < \infty. \quad (5.48)$$

Let, for any  $r \geq 0$ ,

$$b_r := (\psi_r - 1)h. \quad (5.49)$$

As in the proof of Lemma 3.7 in [29] it follows that

$$b_r \in W(U_X \times V_{Y,t}). \quad (5.50)$$

By construction,  $b_r$  vanishes on  $\partial_{\mathcal{K}}(U_X \times V_{Y,t})$ . Together with (5.50), this yields that  $b_r \in W \cap C_{\mathcal{K},0}^\infty(\overline{U_X \times V_{Y,t}})$ . Furthermore,

$$- \iiint_{U_X \times \partial V_{Y,t}} (b_r + h)^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX = - \iiint_{U_X \times \partial V_{Y,t}} \psi_r^2 h^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX.$$

Letting  $r \rightarrow \infty$  we see that

$$\begin{aligned} & \lim_{r \rightarrow \infty} - \iiint_{U_X \times \partial V_{Y,t}} \psi_r^2 h^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \, dX \\ &= \iiint_{U_X \times \partial V_{Y,t}} h^2 ((X, 1) \cdot N_{Y,t})^+ \, d\sigma_{Y,t} \, dX \geq 0. \end{aligned}$$

### 5.9. The proof of Theorem 1.5

Retracing the argument we see by (5.26) and (5.28) that

$$J[f_1, g^*] = 0 \text{ for some } f_1 \in W_0. \quad (5.51)$$

Using Lemma 5.3 we can conclude that  $f_1$  is the unique weak solution  $f_1 \in W_0$  to  $\mathcal{L}u = g^*$  in  $U_X \times V_{Y,t}$  in the sense of Eq (2.7). This completes the proof of existence and uniqueness part of Theorem 1.5. The quantitative estimate follows in the standard way.

### 5.10. A comparison principle

Assume that  $u \in W(U_X \times V_{Y,t})$  is a weak sub-solution to the equation

$$\nabla_X \cdot (A(\nabla_X u, X, Y, t)) - (\partial_t + X \cdot \nabla_Y)u = g^* \text{ in } U_X \times V_{Y,t}. \quad (5.52)$$

By definition this means in particular that  $u \in W(U_X \times V_{Y,t})$ . Given  $u$  we now let  $v \in W(U_X \times V_{Y,t})$  be the unique weak solution to the problem

$$\begin{cases} \nabla_X \cdot (A(\nabla_X v, X, Y, t)) - (\partial_t + X \cdot \nabla_Y)v = g^* & \text{in } U_X \times V_{Y,t}, \\ v = u & \text{on } \partial_{\mathcal{K}}(U_X \times V_{Y,t}), \end{cases} \quad (5.53)$$

in the sense that

$$v \in W(U_X \times V_{Y,t}), \quad (v - u) \in W_0(U_X \times V_{Y,t}), \quad (5.54)$$

and in the sense that (2.7) holds for all  $\phi \in L^2_{Y,t}(V_{Y,t}, H^1_{X,0}(U_X))$ . By Theorem 1.5  $v$  exists and is unique. We want to prove that  $u \leq v$  a.e. in  $U_X \times V_{Y,t}$ . To achieve this we let  $\epsilon > 0$  be arbitrary and we use the test function  $\phi = (u - v - \epsilon)^+$ . Then  $\phi$  is a non-negative admissible test function and  $\phi = 0$  on  $\partial_{\mathcal{K}}(U_X \times V_{Y,t})$ . Hence,

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} A(\nabla_X u, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & + \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)u(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \leq 0, \end{aligned} \quad (5.55)$$

and

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} A(\nabla_X v, X, Y, t) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & + \iint_{V_{Y,t}} \langle g^*(\cdot, Y, t) + (\partial_t + X \cdot \nabla_Y)v(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt = 0. \end{aligned} \quad (5.56)$$

Subtracting these relations, we get

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} (A(\nabla_X v, X, Y, t) - A(\nabla_X u, X, Y, t)) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & + \iint_{V_{Y,t}} \langle \partial_t + X \cdot \nabla_Y (v - u)(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \geq 0. \end{aligned} \quad (5.57)$$

Using the property (1.8)-(ii), we now first note that

$$\begin{aligned} & \iiint_{U_X \times V_{Y,t}} (A(\nabla_X v, X, Y, t) - A(\nabla_X u, X, Y, t)) \cdot \nabla_X \phi \, dX \, dY \, dt \\ & \leq -\Lambda^{-1} \iiint_{U_X \times V_{Y,t}} |\nabla_X (u - v - \epsilon)^+|^2 \, dX \, dY \, dt. \end{aligned} \quad (5.58)$$

Second, again using the definition of  $W(U_X \times V_{Y,t})$  and that  $(v - u) \in W_0(U_X \times V_{Y,t})$ , we see that we see that there exists a sequence  $\{f_j\}$ ,  $f_j \in C^\infty_{\mathcal{K},0}(\overline{U_X \times V_{Y,t}})$  such

$$\begin{aligned} & \iint_{V_{Y,t}} (\partial_t + X \cdot \nabla_Y)(v - u)(\cdot, Y, t), \phi(\cdot, Y, t) \rangle \, dY \, dt \\ & = -\lim_{j \rightarrow \infty} \iiint_{U_X \times V_{Y,t}} (\partial_t + X \cdot \nabla_Y)(f_j - \epsilon)^+(f_j - \epsilon)^+ \, dX \, dY \, dt \\ & = -\frac{1}{2} \lim_{j \rightarrow \infty} \iint_{U_X \times \partial V_{Y,t}} ((f_j - \epsilon)^+)^2 (X, 1) \cdot N_{Y,t} \, d\sigma_{Y,t} \leq 0, \end{aligned} \quad (5.59)$$

as  $f_j = 0$  on  $\partial_{\mathcal{K}}(U_X \times V_{Y,t})$ . Hence, combining (5.57)–(5.59) we conclude that

$$\iiint_{U_X \times V_{Y,t}} |\nabla_X (u - v - \epsilon)^+|^2 \, dX \, dY \, dt \leq 0. \quad (5.60)$$



Finally, using, for a.e.  $(Y, t) \in V_{Y,t}$ , the Poincaré inequality on  $U_X$  we deduce from (5.60) that

$$\iiint_{U_X \times V_{Y,t}} |(u - v - \epsilon)^+|^2 \, dX \, dY \, dt \leq 0. \quad (5.61)$$

Hence  $(u - v - \epsilon)^+ = 0$  a.e. in  $U_X \times V_{Y,t}$  and hence  $u \leq v + \epsilon$  a.e. in  $U_X \times V_{Y,t}$ . We can conclude that we have proved the following theorem.

**Theorem 5.1.** *Let  $u \in W(U_X \times V_{Y,t})$  be a weak sub-solution to the equation in (5.52) in the sense of Definition 4. Given  $u$ , let  $v \in W(U_X \times V_{Y,t})$  be the unique weak solution to the problem in (5.53) in the sense of Definition 4. Then  $u \leq v$  a.e. in  $U_X \times V_{Y,t}$ . Similarly, if  $u \in W(U_X \times V_{Y,t})$  is a weak super-solution to the equation in (5.52) in the sense of Definition 4, then  $v \leq u$  a.e. in  $U_X \times V_{Y,t}$ .*

## 6. Future research and open problems

In this paper we have initiated the study of weak solutions, and their regularity, for what we call nonlinear Kolmogorov-Fokker-Planck type equations. We believe that there are many directions to pursue in this field and in the following we formulate a number of problems.

Let  $p$ ,  $1 < p < \infty$ , be given and let  $A = A(\xi, X, Y, t) : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$  be continuous with respect to  $\xi$ , and measurable with respect to  $X, Y$  and  $t$ . Assume that there exists a finite constant  $\Lambda \geq 1$  such that

$$\Lambda^{-1}|\xi|^p \leq A(\xi, X, Y, t) \cdot \xi \leq \Lambda|\xi|^p \quad (6.1)$$

for almost every  $(X, Y, t) \in \mathbb{R}^{N+1}$  and for all  $\xi \in \mathbb{R}^m$ . Given  $A$  and  $p$  we introduce the operator  $\mathcal{L}_{A,p}$  through

$$\mathcal{L}_{A,p}u := \nabla_X \cdot (A(\nabla_X u(X, Y, t), X, Y, t)) - (\partial_t + X \cdot \nabla_Y)u(X, Y, t). \quad (6.2)$$

This defines a class of strongly degenerate nonlinear parabolic PDEs modelled on the classical PDE of Kolmogorov and the  $p$ -Laplace operator, and to our knowledge there is currently no literature devoted to these operators. The results established in this paper concern  $\mathcal{L}_{A,2}$  assuming that  $A \in M(\Lambda)$  or  $A \in R(\Lambda)$ . We see a number of interesting research problems.

**Problem 1:** Establish existence and uniqueness of weak solutions to the Dirichlet problem

$$\begin{cases} \mathcal{L}_{A,p}u = g^*, & \text{in } U_X \times V_{Y,t}, \\ u = g, & \text{on } \partial_{\mathcal{K}}(U_X \times V_{Y,t}). \end{cases} \quad (6.3)$$

**Problem 2:** Prove higher integrability, local boundedness, Harnack inequalities and local Hölder continuity of weak solutions for the equation  $\mathcal{L}_{A,p}u = 0$  in the case  $p \neq 2$ . This is a challenging problem and the first step is probably to figure out how to replace the result of Bouchut [7], or the use of the fundamental solution constructed by Kolmogorov, in this case. The problem is already very interesting for the prototype

$$\nabla_X \cdot (|\nabla_X u(X, Y, t)|^{p-2} \nabla_X u(X, Y, t)) - (\partial_t + X \cdot \nabla_Y)u(X, Y, t) = 0. \quad (6.4)$$

**Problem 3:** Consider the equation in (6.4). Prove bounds for  $\nabla_X u$  and local Hölder continuity of  $\nabla_X u$ . Note that this must be a difficult problem in the nonlinear setting due to the lack of ellipticity in the variable  $Y$ . Again, the right place to start is probably to (simply) consider the equation

$$\nabla_X \cdot (A(\nabla_X u)) - (\partial_t + X \cdot \nabla_Y)u = 0,$$

where  $A(\xi)$  has linear growth, i.e., a nonlinear  $p = 2$  case.

Finally, we discuss the very formulation of the Dirichlet problem. Consider the geometry of  $U_X \times V_{Y,t}$  and let  $\Gamma := \partial U_X \times V_{Y,t}$  and

$$\begin{aligned} \Sigma^+ &:= \{(X, Y, t) \in \overline{U_X} \times \partial V_{Y,t} \mid (X, 1) \cdot N_{Y,t} > 0\}, \\ \Sigma_0 &:= \{(X, Y, t) \in \overline{U_X} \times \partial V_{Y,t} \mid (X, 1) \cdot N_{Y,t} = 0\}, \\ \Sigma^- &:= \{(X, Y, t) \in \overline{U_X} \times \partial V_{Y,t} \mid (X, 1) \cdot N_{Y,t} < 0\}. \end{aligned} \quad (6.5)$$

Using this notation  $\partial_{\mathcal{K}}(U_X \times V_{Y,t}) = \Gamma \cup \Sigma^-$ . Recall that  $W(U_X \times V_{Y,t})$  is defined as the closure of  $C^\infty(\overline{U_X \times V_{Y,t}})$  in the norm

$$\|u\|_{W(U_X \times V_{Y,t})} := \left( \|u\|_{L^2_{Y,t}(V_{Y,t}, H^1_X(U_X))}^2 + \|(\partial_t + X \cdot \nabla_Y)u\|_{L^2_{Y,t}(V_{Y,t}, H^{-1}_X(U_X))}^2 \right)^{1/2}. \quad (6.6)$$

Assuming  $u \in W(U_X \times V_{Y,t})$ , it is relevant to define and study the trace of  $u$  to  $\Gamma \cup \Sigma^+ \cup \Sigma_0 \cup \Sigma^-$ . We let  $= B^{2,2}_{1/2}(\partial U_X)$  denote the Besov space defined as the trace space of  $H^1_X(U_X)$  to  $\partial U_X$  (this space is often denoted  $H^{1/2}(\partial U_X)$  in the literature). It is well known, that if  $U_X$  is a bounded Lipschitz domain, then there exists a bounded continuous non-injective operator  $T : H^1_X(U_X) \rightarrow B^{2,2}_{1/2}(\partial U_X)$ , called the trace operator, and a bounded continuous operator  $E : B^{2,2}_{1/2}(\partial U_X) \rightarrow H^1_X(U_X)$  called the extension operator. The trace space of  $L^2_{Y,t}(V_{Y,t}, H^1_X(U_X))$  on  $\Gamma$  is therefore  $L^2_{Y,t}(V_{Y,t}, B^{2,2}_{1/2}(\partial U_X))$  and

$$\|u\|_{L^2_{Y,t}(V_{Y,t}, B^{2,2}_{1/2}(\partial U_X))} \leq c \|u\|_{W(U_X \times V_{Y,t})}.$$

The trace to  $\Sigma^+ \cup \Sigma_0 \cup \Sigma^-$  is less clear. Indeed, recall that the space  $W_{X,0}(U_X \times V_{Y,t})$  is defined as the closure in the norm of  $W(U_X \times V_{Y,t})$  of  $C^\infty_{X,0}(\overline{U_X \times V_{Y,t}})$ . In particular, given  $f \in W_{X,0}(U_X \times V_{Y,t})$  there exists  $\{f_j\}$ ,  $f_j \in C^\infty_{X,0}(\overline{U_X \times V_{Y,t}})$  such that

$$\|f - f_j\|_{W(U_X \times V_{Y,t})} \rightarrow 0 \text{ as } j \rightarrow \infty,$$

and consequently,

$$\|(\partial_t + X \cdot \nabla_Y)(f - f_j)\|_{L^2_{Y,t}(V_{Y,t}, H^{-1}_X(U_X))} \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Using this we see that

$$\begin{aligned} & \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \\ &= \lim_{j \rightarrow \infty} \iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y)f_j(\cdot, Y, t), f_j(\cdot, Y, t) \rangle dY dt \\ &= \frac{1}{2} \lim_{j \rightarrow \infty} \iint_{\Sigma^+ \cup \Sigma_0 \cup \Sigma^-} f_j^2(X, Y, t) (X, 1) \cdot N_{Y,t} d\sigma_{Y,t}. \end{aligned} \quad (6.7)$$

The first obstruction to a trace inequality is that  $(X, 1) \cdot N_{Y,t} d\sigma_{Y,t}$  is a signed measure on  $\Sigma^+ \cup \Sigma_0 \cup \Sigma^-$ . Assuming that  $f \in W_0(U_X \times V_{Y,t})$  we deduce that

$$\iint_{V_{Y,t}} \langle (\partial_t + X \cdot \nabla_Y) f(\cdot, Y, t), f(\cdot, Y, t) \rangle dY dt \quad (6.8)$$

$$= \frac{1}{2} \lim_{j \rightarrow \infty} \iint_{\Sigma^+ \cup \Sigma_0} f_j^2(X, Y, t) (X, 1) \cdot N_{Y,t} d\sigma_{Y,t}. \quad (6.9)$$

Hence, in this case

$$\lim_{j \rightarrow \infty} \iint_{\Sigma^+ \cup \Sigma_0} f_j^2(X, Y, t) (X, 1) \cdot N_{Y,t} d\sigma_{Y,t} \leq c \|f\|_{W(U_X \times V_{Y,t})}, \quad (6.10)$$

and we see that we can extract a subsequence of  $\{f_j\}$  converging in  $L^2(K, (X, 1) \cdot N_{Y,t} d\sigma_{Y,t})$  whenever  $K$  is a compact subset of  $\Sigma^+$ . At the expense of additional notation the roles of  $\Sigma^+$  and  $\Sigma^-$  can be interchanged in this argument. This observation highlights the difficulty concerning the possibility of a trace inequality and concerning the identification of the trace space for  $W(U_X \times V_{Y,t})$ . This explains why we in this paper, as in [29], have used the weaker formulation of the Dirichlet problem introduced.

**Problem 4:** What function space is the space of traces, to  $\Gamma \cup \Sigma^+ \cup \Sigma_0 \cup \Sigma^-$ , of  $W(U_X \times V_{Y,t})$ ?

### Conflict of interest

The authors declare no conflict of interest.

### References

1. D. Albritton, S. N. Armstrong, J. C. Mourrat, M. Novack, Variational methods for the kinetic Fokker-Planck equation, 2019, arXiv:1902.04037.
2. F. Anceschi, S. Polidoro, M. A. Ragusa, Moser's estimates for degenerate Kolmogorov equations with non-negative divergence lower order coefficients, *Nonlinear Anal.*, **189** (2019), 111568. <https://doi.org/10.1016/j.na.2019.07.001>
3. F. Anceschi, A. Rebusci, A note on the weak regularity theory for degenerate Kolmogorov equations, 2021, arXiv:2107.04441.
4. S. N. Armstrong, A. Bordas, J. C. Mourrat, Quantitative stochastic homogenization and regularity theory of parabolic equations, *Anal. PDE*, **11** (2018), 1945–2014. <https://doi.org/10.2140/apde.2018.11.1945>
5. S. N. Armstrong, J. C. Mourrat, Lipschitz regularity for elliptic equations with random coefficients, *Arch. Rational Mech. Anal.*, **219** (2016), 255–348. <https://doi.org/10.1007/s00205-015-0908-4>
6. L. Boccardo, F. Murat, Almost everywhere convergence of the gradients of solutions to elliptic and parabolic equations, *Nonlinear Anal. Theor.*, **19** (1992), 581–597. [https://doi.org/10.1016/0362-546X\(92\)90023-8](https://doi.org/10.1016/0362-546X(92)90023-8)
7. F. Bouchut, Hypocoelliptic regularity in kinetic equations, *J. Math. Pure. Appl.*, **81** (2002), 1135–1159. [https://doi.org/10.1016/S0021-7824\(02\)01264-3](https://doi.org/10.1016/S0021-7824(02)01264-3)

8. M. Bramanti, M. C. Cerutti, M. Manfredini,  $L^p$  estimates for some ultraparabolic operators with discontinuous coefficients, *J. Math. Anal. Appl.*, **200** (1996), 332–354. <https://doi.org/10.1006/jmaa.1996.0209>
9. H. Brézis, I. Ekeland, Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps, *C. R. Acad. Sci. Paris Sér. A-B*, **282** (1976), A1197–A1198.
10. H. Brézis, I. Ekeland, Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps, *C. R. Acad. Sci. Paris Sér. A-B*, **282** (1976), A971–A974.
11. C. Cercignani,  $H$ -theorem and trend to equilibrium in the kinetic theory of gases, *Arch. Mech.*, **34** (1982), 231–241.
12. C. Cinti, A. Pascucci, S. Polidoro, Pointwise estimates for a class of non-homogeneous Kolmogorov equations, *Math. Ann.*, **340** (2008), 237–264. <https://doi.org/10.1007/s00208-007-0147-6>
13. L. Desvillettes, Entropy dissipation estimates for the Landau equation in the Coulomb case and applications, *J. Funct. Anal.*, **269** (2015), 1359–1403. <https://doi.org/10.1016/j.jfa.2015.05.009>
14. L. Desvillettes, C. Mouhot, C. Villani, Celebrating Cercignani’s conjecture for the Boltzmann equation, *Kinet. Relat. Mod.*, **4** (2011), 277–294. <https://doi.org/10.3934/krm.2011.4.277>
15. L. Desvillettes, C. Villani, On the spatially homogeneous Landau equation for hard potentials II:  $H$ -theorem and applications, *Commun. Part. Diff. Eq.*, **25** (2000), 261–298. <https://doi.org/10.1080/03605300008821513>
16. E. DiBenedetto, U. Gianazza, V. Vespri, *Harnack’s inequality for degenerate and singular parabolic equations*, New York: Springer, 2012. <https://doi.org/10.1007/978-1-4614-1584-8>
17. I. Ekeland, R. Temam, *Convex analysis and variational problems*, Amsterdam-Oxford: North-Holland Publishing Co., 1976.
18. N. Ghoussoub, *Self-dual partial differential systems and their variational principles*, New York: Springer, 2009. <https://doi.org/10.1007/978-0-387-84897-6>
19. N. Ghoussoub, L. Tzou, A variational principle for gradient flows, *Math. Ann.*, **330** (2004), 519–549. <https://doi.org/10.1007/s00208-004-0558-6>
20. F. Golse, C. Imbert, C. Mouhot, A. F. Vasseur, Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, **19** (2019), 253–295. [https://doi.org/10.2422/2036-2145.201702\\_001](https://doi.org/10.2422/2036-2145.201702_001)
21. J. Guerand, Quantitative regularity for parabolic De Giorgi classes, 2019, arXiv:1903.07421.
22. J. Guerand, C. Imbert, Log-transform and the weak Harnack inequality for kinetic Fokker-Planck equations, 2021, arXiv:2102.04105.
23. J. Guerand, C. Mouhot, Quantitative de Giorgi methods in kinetic theory, 2021, arXiv:2103.09646.
24. L. Hörmander, Hypoelliptic second order differential equations, *Acta Math.*, **119** (1967), 147–171. <https://doi.org/10.1007/BF02392081>
25. A. N. Kolmogoroff, Zufällige Bewegungen (zur Theorie der Brownschen Bewegung), *Ann. Math. (2)*, **35** (1934), 116–117. <https://doi.org/10.2307/1968123>
26. E. Lanconelli, F. Lascialfari, A boundary value problem for a class of quasilinear operators of Fokker-Planck type, *Ann. Univ. Ferrara*, **41** (1996), 65–84. <https://doi.org/10.1007/BF02825256>

27. F. Lascialfari, D. Morbidelli, A boundary value problem for a class of quasilinear ultraparabolic equations, *Commun. Part. Diff. Eq.*, **23** (1998), 847–868. <https://doi.org/10.1080/03605309808821369>
28. P. L. Lions, On Boltzmann and Landau equations, *Philos. Trans. Roy. Soc. London Ser. A*, **346** (1994), 191–204. <https://doi.org/10.1098/rsta.1994.0018>
29. M. Litsgård, K. Nyström, The Dirichlet problem for Kolmogorov-Fokker-Planck type equations with rough coefficients, *J. Funct. Anal.*, **281** (2021), 109226. <https://doi.org/10.1016/j.jfa.2021.109226>
30. M. Litsgård, K. Nyström, Potential theory for a class of strongly degenerate parabolic operators of Kolmogorov type with rough coefficients, *J. Math. Pure. Appl.*, **157** (2022), 45–100. <https://doi.org/10.1016/j.matpur.2021.11.004>
31. M. Manfredini, The Dirichlet problem for a class of ultraparabolic equations, *Adv. Differential Equations*, **2** (1997), 831–866.
32. M. Manfredini, S. Polidoro, Interior regularity for weak solutions of ultraparabolic equations in divergence form with discontinuous coefficients, *Bollettino dell'Unione Matematica Italiana Serie 8*, **1-B** (1998), 651–675.
33. C. Mouhot, De Giorgi–Nash–Moser and Hörmander theories: new interplays, In: *Proceedings of the International Congress of Mathematicians (ICM 2018)*, Hackensack, NJ: World Sci. Publ., 2018, 2467–2493. [https://doi.org/10.1142/9789813272880\\_0146](https://doi.org/10.1142/9789813272880_0146)
34. A. Pascucci, S. Polidoro, The Moser's iterative method for a class of ultraparabolic equations, *Commun. Contemp. Math.*, **6** (2004), 395–417. <https://doi.org/10.1142/S0219199704001355>
35. S. Polidoro, M. A. Ragusa, Hölder regularity for solutions of ultraparabolic equations in divergence form, *Potential Anal.*, **14** (2001), 341–350. <https://doi.org/10.1023/A:1011261019736>
36. C. A. Truesdell, R. G. Muncaster, *Fundamentals of Maxwell's kinetic theory of a simple monatomic gas*, New York-London: Academic Press, Inc., 1980.
37. C. Villani, A review of mathematical topics in collisional kinetic theory, In: *Handbook of mathematical fluid dynamics, Vol I*, Amsterdam: North-Holland, 2002, 71–74. [https://doi.org/10.1016/S1874-5792\(02\)80004-0](https://doi.org/10.1016/S1874-5792(02)80004-0)
38. W. Wang, L. Zhang, The  $C^\alpha$  regularity of a class of non-homogeneous ultraparabolic equations, *Sci. China Ser. A-Math.*, **52** (2009), 1589–1606. <https://doi.org/10.1007/s11425-009-0158-8>
39. W. Wang, L. Zhang, The  $C^\alpha$  regularity of weak solutions of ultraparabolic equations, *Discrete Contin. Dyn. Syst.*, **29** (2011), 1261–1275. <https://doi.org/10.3934/dcds.2011.29.1261>
40. W. Wang, L. Zhang,  $C^\alpha$  regularity of weak solutions of non-homogenous ultraparabolic equations with drift terms, 2017, arXiv:1704.05323.
41. Y. Zhu, Velocity averaging and Hölder regularity for kinetic Fokker-Planck equations with general transport operators and rough coefficients, *SIAM J. Math. Anal.*, **53** (2021), 2746–2775. <https://doi.org/10.1137/20M1372147>

