# Ergodic Mean Field Games with Hörmander diffusions

Federica Dragoni * Ermal Feleqi †

June 18, 2018

**Abstract**

We prove existence of solutions for a class of systems of subelliptic PDEs arising from Mean Field Game systems with Hörmander diffusion. These results are motivated by the feedback synthesis Mean Field Game solutions and the Nash equilibria of a large class of $N$-player differential games.

**Keywords:** Mean Field Games, Hörmander condition, subelleptic PDEs, hypoelliptic.

## 1 Introduction

In this paper we consider a class of systems of degenerate elliptic PDEs of Hörmander type arising from certain ergodic differential games, more specifically, from the Mean Field Game (MFG) theory of J.M. Lasry and P.L. Lions [43, 44, 45]. These systems have been introduced to model differential games with a large number of players or agents with dynamics described by controlled diffusion processes, under simplifying features such as homogeneity of the agents and a coupling of Mean Field type. This allows to carry out a kind of limit procedure as the number of agents tends to infinity which leads to simpler effective models. Lasry and Lions have shown that for a large class of differential games (either deterministic or stochastic) the limiting model reduces to a Hamilton-Jacobi-Bellman equation for the optimal value function of the typical agent coupled with a continuity (or Fokker-Planck) equation for the density of the typical optimal dynamic, the so-called Mean Field Game equations. Solutions to these equations can be used to construct approximated Nash equilibria for games with a very large but still finite number of agents. The rigorous proof of the limit behaviour in this sense has been established by Lasry and Lions in [43, 45] for ergodic differential games and extended by one of the authors to several homogeneous populations of agents [31]. The time-dependent case with nonlocal coupling has been addressed in a general context by [20]. For a general overview on Mean Field Games, we refer the reader to the lecture notes of Guéant, Lasry, and Lions [37], Cardaliaguet [18], the lecture videos of P.-L. Lions at his webpage at Collège de France, the first papers of Lasry and Lions [43, 44, 45] and of M. Huang, P.E. Caines, R.P. Malhamé [39], [40], the survey paper [36], the book by Gomes and collaborators [34] and by Bensoussan, Frehse and Yan [13], the two special issues [8, 9] and the recent paper [20] on the master equation and its application to the convergence of games with a large population to a MFG. For applications to economics see e.g. [3], [22], [34], [37], [46], [42]. From the mathematical side, there are several important questions related to both the convergence and then the study of the limit MFG system itself, e.g. long time behaviour [19, 21], ergodic MFG systems [24, 11, 35], for homogenisation [23]. For further contributions see also [2, 4, 33]. The literature on Mean Field Games is very vast so the previous list is only partial and we refer to the references therein for a more extended bibliography.

The novelty of this paper consists in assuming that the dynamic of the average player is a diffusion of Hörmander type and hence the differential operators arising in the system are degenerate: the second order operator is not elliptic but only subelliptic. Roughly speaking this means that

*Cardiff School of Mathematics, Cardiff University, Cardiff, UK, e-mail: DragoniF@cardiff.ac.uk.
†Cardiff School of Mathematics, Cardiff University, Cardiff, UK, e-mail: FeleqiE@cardiff.ac.uk

the operators are elliptic only along certain directions of derivatives. Nevertheless the Hörmander condition ensures that the Laplacian induced by these selected derivatives is hypoelliptic. From the perspective of a single agent this means that the state cannot change in all directions, but the agent can move only along admissible directions: a subspace of the tangent space. This subspace depends on the state (position) of the agent. Similarly the growth conditions on the Hamiltonian are restricted to some selected directions of derivatives. This extension is not trivial and relies on recent deep achievements in the theory of Hörmander operators and subellipitc quasilinear equations. When the known regularity results will not be sufficient to proceed, we will use heat kernel estimates to overcome the problem. Moreover the techniques used here are different from the standard elliptic case and can also be used in other contexts to gain a-posteriori regularity.

Hamilton-Jacobi equations in the context of Hörmander regularity have been extensively studied, see e.g. [6, 17, 25, 28, 29], in particular because of the intriguing connection between the PDE theory and the underlying geometry induced by the admissible directions. This paper is to our knowledge the first one that connects these two recent and active areas.

We next state our main results:
1 - Under suitable assumptions (see Section 3) and assuming in particular that the Hamiltonian grows at most quadratically in the subgradient, we prove that there exists a solution $(u, m) \in C^2_{\mathcal{X}}(\mathbb{T}^d) \times C(\mathbb{T}^d)$ of the system

$$\begin{cases} \mathcal{L}u + \rho u + H(x, D_{\mathcal{X}}u) = V[m] \\ \mathcal{L}^* m - \mathrm{div}_{\mathcal{X}^*}(mg(x, D_{\mathcal{X}}u)) = 0 \\ \displaystyle\int_{\mathbb{T}^d} m\, dx = 1, \quad m > 0, \end{cases}$$

where $D_{\mathcal{X}}u$ is a subgradient associated to a family of Hörmander vector fields (e.g. $D_{\mathcal{X}}u = \left(u_x - \frac{y}{2}u_z, u_y + \frac{x}{2}u_z\right)^T$ on $\mathbb{R}^3$ in the Heisenberg case) and $\mathcal{L}$ is a hypoelliptic operator, $\mathcal{L}^*$ is the dual operator of $\mathcal{L}$ and $\mathrm{div}_{\mathcal{X}^*}$ is the corresponding divergence operator. Moreover by $C^2_{\mathcal{X}}(\mathbb{T}^d)$ we indicate the sets of functions whose first and second derivatives in the selected directions exist and are continuous (see Section 2 for more formal definitions).
2 - Under suitable assumptions (see Section 4)) and assuming in particular that the Hamiltonian grows at most linearly in the subgradient, we prove that there exists a solution $(\lambda, u, m) \in \mathbb{R} \times C^2_{\mathcal{X}}(\mathbb{T}^d) \times C(\mathbb{T}^d)$ of the system

$$\begin{cases} \mathcal{L}u + \lambda + H(x, D_{\mathcal{X}}u) = V[m] \\ \mathcal{L}^* m - \mathrm{div}_{\mathcal{X}^*}\big(mg(x, D_{\mathcal{X}}u)\big) = 0 \\ \displaystyle\int_{\mathbb{T}^d} u\, dx = 0, \quad \int_{\mathbb{T}^d} m\, dx = 1, \quad m > 0. \end{cases}$$

We also show uniqueness for both the systems under standard monotonicity assumptions.
Those results are applied to the feedback synthesis of MFG solutions and of Nash equilibria of a large class of $N$-player differential games.

The paper is organised as follows: in Section 2 we introduce the Hörmander condition and the corresponding first and second order operators and we state several regularity results and estimates which will be key in the proofs of our main results. In Section 3 we show existence for a stationary MFG system for at most quadratic Hamiltonians by a fixed-point argument in the presence of a regularisation. In Section 4 we remove this regularisation for Hamiltonians of at most linear growth and prove our main existence result. In the Appendix we show the convergence of Nash-equilibria as motivation for the MFG system studied. Since these results are very well-known in the non degenerate case and they do not lead to any substantial technical difference in the Hörmander case, we will omit the proofs, only reporting briefly the results.

**Acknowledgments:** The authors were supported by the EPSRC Grant "Random Perturbations of ultra-parabolic PDEs under rescaling". The authors would like to thank Nicolas Dirr and Pierre Cardaliaguet for the many interesting conversations and suggestions.

## 2 Preliminaries and notations

Let us consider $x \in \mathbb{T}^d$ the $d$-dimensional torus and $\mathcal{X} = \{X_1, \dots, X_m\}$ a family of smooth vector fields defined on $\mathbb{T}^d$ satisfying the Hörmander condition, i.e.

$$Span\Big(\mathcal{L}\big(X_1(x), \dots, X_m(x)\big)\Big) = T_x\mathbb{T}^d \equiv \mathbb{R}^d, \quad \forall\, x \in \mathbb{T}^d, \tag{2.1}$$

where $\mathcal{L}\big(X_1(x), \dots, X_m(x)\big)$ denotes the Lie algebra induced by the given vector fields and by $T_x\mathbb{T}^d$ we denote the tangent space at the point $x \in \mathbb{T}^d$. For more details on Hörmander vector fields we refer to [49]. Given a family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$ and $u : \mathbb{T}^d \to \mathbb{R}$, we define:

$$D_{\mathcal{X}}u = (X_1u, \dots, X_mu)^T \in \mathbb{R}^m, \tag{2.2}$$

$$\mathcal{L}u = -\frac{1}{2}\sum_{j=1}^{m} X_j^2 u \in \mathbb{R}. \tag{2.3}$$

For any vector-valued function $g : \mathbb{T}^d \to \mathbb{R}^m$, we will consider the divergence induced by the vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$, that is

$$\mathrm{div}_{\mathcal{X}} g = X_1\, g_1 + \dots + X_m\, g_m, \tag{2.4}$$

where $g_i$ indicates the $i$-component of $g$, for $i = 1, \dots, m$. In particular, later on, we will consider the divergence $\mathrm{div}_{\mathcal{X}^*} g$ induced by the dual vector fields $X_i^* = -X_i - \mathrm{div}X_i$ where $\mathrm{div}X_i$ indicate the standard (Euclidean) divergence of the vector fields $X_i : \mathbb{T}^d \to \mathbb{R}^d$, for $i = 1, \dots, m$. Given the family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$ we recall that any absolutely continuous curve $\gamma : [0,T] \to \mathbb{T}^d$ is called horizontal (or admissible) if there exists a measurable function $\alpha : [0,T] \to \mathbb{R}^m$ such that

$$\dot\gamma(t) = \sum_{i=1}^{m} \alpha_i(t) X_i(\gamma(t)), \quad \text{a.e. } t \in (0,T), \tag{2.5}$$

where $\alpha_i(t)$ is the $i$-component of $\alpha(t)$ for $i = 1, \dots, m$.

For all horizontal curves it is possible to define the length as:

$$l(\gamma) = \int_0^T \sqrt{\sum_{i=1}^{m} \alpha_i^2(t)}\; dt.$$

The *Carnot-Carathéodory distance* induced by the family $\mathcal{X} = \{X_1, \dots, X_m\}$ is denoted by $d_{CC}(\cdot,\cdot)$, and defined as

$$d_{CC}(x,y) = \inf\left\{ l(\gamma) \,|\, \gamma \text{ satisfying (2.5) with } \gamma(0) = x, \gamma(T) = y \right\}.$$

The Hörmander condition implies that the distance $d_{CC}(x,y)$ is finite and continuous w.r.t. the original Euclidean topology induced on $\mathbb{T}^d$ (see e.g. [49]). It is also known that there exists $C > 0$ such that

$$C^{-1}|x-y| \le d_{CC}(x,y) \le C|x-y|^{1/k} \tag{2.6}$$

for all $x, y \in \mathbb{T}^d$, where $k \in \mathbb{N}$ is the step, i.e. the maximum of the degrees of the iterated brackets occurring in the fulfillment of the Hörmander condition, see [50]. It was proved in [52, Lemma 5] and independently in [50] that there exists some $Q > 0$, called the *homogenous dimension*, such that, for all $\delta > 0$ sufficiently small and for some $C > 0$,

$$C^{-1}\delta^Q \le |B_{d_{CC}}(x,\delta)| \le C\delta^Q,$$

for all $x \in \mathbb{T}^d$, where $B_{d_{CC}}(x,\delta)$ is the ball of centre $x$ and radius $\delta$ w.r.t. the distance $d_{CC}$ and, for any $B \subset \mathbb{T}^d$, $|B|$ denotes the standard Lebesgue measure of $B$.

## 2.1 Hölder spaces and Hölder regularity estimates

Next we recall the definition of Hölder and Sobolev spaces associated to the family of vector fields $\mathcal{X}$ (we refer to [56] and [57] for more details on these spaces). For every multi-index $J = (j_1, \dots, j_m) \in \mathbb{Z}_+^m$ let $\mathcal{X}^J = X_{j_1} \cdots X_{j_m}$. The *length* of a multi-index $J$ is $|J| = j_1 + \cdots + j_m$, thus $\mathcal{X}^J$ is a linear differential operator of order $|J|$. For $r \in \mathbb{N}$ and $\alpha \in (0,1)$ we define the function spaces

$$C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d) = \left\{ u \in L^\infty(\mathbb{T}^d) \ : \ \sup_{\substack{x,y\in\mathbb{T}^d\\ x\neq y}} \frac{|u(x)-u(y)|}{d_{CC}(x,y)^\alpha} < \infty \right\},$$

$$C^{r,\alpha}_{\mathcal{X}}(\mathbb{T}^d) = \left\{ u \in L^\infty(\mathbb{T}^d) \ : \ \mathcal{X}^J u \in C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d) \ \ \forall |J| \leq r \right\}.$$

For any function $u \in C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ one can define a seminorm as

$$[u]_{C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} = \sup_{\substack{x,y\in\mathbb{T}^d\\ x\neq y}} \frac{|u(x)-u(y)|}{d_{CC}(x,y)^\alpha},$$

and, for every $u \in C^{r,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$, the norm is defined as

$$\|u\|_{C^{r,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} = \|u\|_{L^\infty(\mathbb{T}^d)} + \sum_{1\leq |J| \leq r} [\mathcal{X}^J u]_{C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)}.$$

Endowed with the above norm, $C^{r,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ are Banach spaces for any $r \in \mathbb{N}$ and $\alpha \in (0,1)$.

From estimates (2.6), it follows immediately

$$C^{-1} \|u\|_{C^{0,\frac{\alpha}{k}}(\mathbb{T}^d)} \leq \|u\|_{C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C \|u\|_{C^{0,\alpha}(\mathbb{T}^d)} \quad \Longrightarrow \quad C^{0,\alpha}(\mathbb{T}^d) \subset C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d) \subset C^{0,\frac{\alpha}{k}}(\mathbb{T}^d), \tag{2.7}$$

where $\|u\|_{C^{0,\alpha}(\mathbb{T}^d)}$ is the standard Hölder norm, $k$ is the step in the Hörmander condition and $C > 0$ is a global constant depending only on the dimension $d$ and the family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$. More in general, for all $r \in \mathbb{N}$, $C^{r,\alpha}(\mathbb{T}^d) \subset C^{r,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$.
Let $r$ be a non-negative integer and $1 \leq p \leq \infty$. We define the space

$$W^{r,p}_{\mathcal{X}}(\mathbb{T}^d) = \left\{ u \in L^p(\mathbb{T}^d) \ : \ \mathcal{X}^J u \in L^p(\mathbb{T}^d), \ \forall J \in \mathbb{Z}^m_+, \ |J| \leq r \right\}.$$

Endowed with the norm $\|u\|_{W^{r,p}_{\mathcal{X}}(\mathbb{T}^d)} = \left( \sum_{|J|\leq r} \int_{\mathbb{T}^d} |\mathcal{X}^J u|^p \, dx \right)^{1/p}$, $W^{r,p}_{\mathcal{X}}(\mathbb{T}^d)$ is a Banach space. For $p = 2$ we write $H^r_{\mathcal{X}}(\mathbb{T}^d)$ instead of $W^{r,p}_{\mathcal{X}}(\mathbb{T}^d)$ and in this case the space is Hilbert when endowed with the corresponding inner product. Moreover, for any $1 \leq p < \infty$, the embeddings

$$C^{kr,\alpha}_{\mathcal{X}}(\mathbb{T}^d) \hookrightarrow C^{r,\frac{\alpha}{k}}(\mathbb{T}^d),$$

$$W^{r,p}_{\mathcal{X}}(\mathbb{T}^d) \hookrightarrow W^{r/k,p}(\mathbb{T}^d),$$

hold true. The first is proved in [55] and the second in [54].

In proving one of our main results we will also need the following compact embedding.

**Lemma 2.1.** *$W^{1,p}_{\mathcal{X}}(\mathbb{T}^d)$ is compactly embedded into $L^p(\mathbb{T}^d)$.*

This follows from the previous embedding and the fact that the fractional Sobolev space $W^{k/m,p}(\mathbb{T}^d)$ is compactly embedded into $L^p(\mathbb{T}^d)$ (see e.g. [27]).

Next we want to recall some Hölder regularity results for linear and quasilinear subelleptic PDEs, key for the later existence results. Hölder and Schauder estimates for subelliptic linear and quasilinear equations have been proved by Xu [54, 56], Xu-Zuily [57] and [48]; see also the references therein. In particular we will consider the results proved in [57], but we will rewrite them in a stronger form, by combining them with some $L^p$-estimates proved by Sun-Liu-Li-Zheng [53].

The results in [57] are proved for subelliptic systems but we will apply them to the case of a single equation. We first consider linear equations of the form:

$$\mathrm{div}_{\mathcal{X}^*}\big(A(x)D_{\mathcal{X}}u\big) + g(x)\cdot D_{\mathcal{X}}u + c(x)\,u = f(x). \tag{2.8}$$

and assume that

$$A(x) \text{ is a } m\times m\text{-uniformly elliptic matrix.} \tag{2.9}$$

Note that in the case of the sub-Laplacian the previous assumption is trivially satisfied since $A(x)$ is equal to the identity $m\times m$-matrix.

**Theorem 2.2** ($C^{2,\alpha}_{\mathcal{X}}$-regularity for linear subelleptic PDEs, [57, 53].)**.** *Assuming* (2.9) *and that all coefficients of $A(x)$, $g(x)$, $c(x)$ and $f(x)$ are Hölder continuous, then any weak solution $u\in H^1_{\mathcal{X}}(\mathbb{T}^d)$ of* (2.8) *belongs to $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ for some $\alpha\in(0,1)$.*

*Moreover there exists a constant $C>0$ (depending only on the Hölder norms of the coefficients of the equation, on $d$ and on the vector fields $\mathcal{X}$) such that*

$$\|u\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \le C.$$

*Proof.* First we recall that, if the coefficients are $C^{0,\alpha}$ then they are also $C^{0,\alpha}_{\mathcal{X}}$ (see (2.7)). Then Theorem 3.4 and Theorem 3.5 in [57] ensure that, given any $u$ weak $H^1_{\mathcal{X}}$- solution, $u$ belongs to $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$, and the $C^{2,\alpha}_{\mathcal{X}}$-Hölder norm of $u$ is bounded by a constant depending on the Hölder norms of the coefficients, on the geometry of the problem (i.e. the step $r$, the dimension $d$ and the number of vector fields $m$), but also on a constant $M$ such that $\|u\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)}\le M$.
We can now use the uniform $L^p$ estimates proved in Theorem 1.4 in [53] to show that the constant $C$ is actually independent of $M$, i.e. independent of the $H^1_{\mathcal{X}}$-norm of $u$. Note that Hölder regularity on a compact domain implies all the necessary $L^p$-bounds to apply the result in [53]. □

Let us now consider a subelliptic quasilinear equation of the form:

$$\mathrm{div}_{\mathcal{X}^*}\big(A(x)D_{\mathcal{X}}u\big) = f(x,u,D_{\mathcal{X}}u). \tag{2.10}$$

and assume that $f(x,z,q)$ is a Hölder function with at most quadratic grow, i.e.

$$|f(x,z,q)| \le a|q|^2 + b, \tag{2.11}$$

for some non-negative constants $a$ and $b$.

**Theorem 2.3** ($C^{1,\alpha}_{\mathcal{X}}$-regularity for quasilinear subelleptic PDEs, [57, 53].)**.** *Assuming* (2.9), (2.11) *and that all the coefficients of the equation are Hölder continuous, then any weak solution $u\in H^1_{\mathcal{X}}(\mathbb{T}^d)\cap C(\mathbb{T}^d)$ belongs to $C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ for some $\alpha\in(0,1)$ and there exists a constant $C>0$ (depending only on the Hölder norms of the coefficients of $A(x)$ and of $f$, on $a$ and $b$ in* (2.11)*, on the step $r$, on $d$ and $m$) such that*

$$\|u\|_{C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \le C.$$

*Proof.* Combining once again the $L^p$-estimates in [53] with Theorem 4.1 in [57] one can immediately deduce the result. □

**Theorem 2.4** ($C^\infty$-regularity, Theorem 4.2, [57])**.** *Under the assumptions of Theorem 2.3, if in addition all coefficients in equation* (2.10) *are $C^\infty(\mathbb{T}^d)$ then $u\in C^\infty(\mathbb{T}^d)$.*

# 3 Discounted systems with at most quadratic Hamiltonians

In this section we consider a subelliptic MFG system with a first order nonlinear term that grows at most quadratic w.r.t. the horizontal gradient. We assume:

**(II-Q)** For $q = \sigma(x)p \in \mathbb{R}^m$ there exists a constant $C \geq 0$ such that

$$|H(x, q)| \leq C(|q|^2 + 1) \quad \forall x \in \mathbb{T}^d\,,\ q \in \mathbb{R}^m. \tag{3.1}$$

**(III)** The vector-valued function $g\colon \mathbb{T}^d \times \mathbb{R}^m \to \mathbb{R}^m$ is Hölder-continuous.

(Note that since $\mathbb{T}^d$ is compact and we will later prove global bounds for $D_{\mathcal{X}}u$, the continuity of $g$ implies also that $g$ is globally bounded).

**(IV)** Set $\mathcal{A} := \{m \in C(\mathbb{T}^d) \,:\, m > 0\,,\ \int_{\mathbb{T}^d} m(x)\,dx = 1\}$, then the map $V\colon \mathcal{A} \to L^\infty(\mathbb{T}^d)$ is assumed continuous and bounded. Moreover, we assume that $V$ is *regularising*, that is, $V[m] \in C^\alpha_{\mathcal{X}}(\mathbb{T}^d)$ for all $m \in \mathcal{A}$, and $\sup_{m\in\mathcal{A}} \|V[m]\|_{C^\alpha_{\mathcal{X}}(\mathbb{T}^d)} < \infty$.

**Theorem 3.1.** *Assume* (2.1)*,* **(II-Q)***,* **(III)***,* **(IV)** *and that $H(x, q)$ is locally Hölder, then given $\mathcal{L}$ defined in* (2.3) *with dual operator $\mathcal{L}^*$ and* $\mathrm{div}_{\mathcal{X}^*}$ *defined as in* (2.4) *w.r.t. the dual vector fields $X_i^* = -X_i - divX_i$, for every $\rho > 0$ the system*

$$\begin{cases} \mathcal{L}u + \rho u + H(x, D_{\mathcal{X}}u) = V[m] \\ \mathcal{L}^*m - div_{\mathcal{X}^*}(mg(x, D_{\mathcal{X}}u)) = 0 \\ \displaystyle\int_{\mathbb{T}^d} m\,dx = 1, \quad m > 0 \end{cases} \tag{3.2}$$

*has a solution $(u, m) \in C^2_{\mathcal{X}}(\mathbb{T}^d) \times C(\mathbb{T}^d)$. (Note that $u$ solves the system in the classical sense while $m$ is a weak solution in the distributional sense.)*

To prove the existence for the system (3.2) we need to look at both the equations involved, starting first from the associated linear PDE for $u$.

**Lemma 3.2.** *Assume* (2.1) *and that $\mathcal{L}$ is the corresponding sub-Laplacian defined in* (2.3)*, then for every $\rho > 0$ and $f \in C^{0,\alpha}(\mathbb{T}^d)$*

$$\mathcal{L}u + \rho u = f \ \text{ in } \mathbb{T}^d \tag{3.3}$$

*has a unique solution $u \in C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$. Moreover $\exists\, C \geq 0$ (independent of $u$ and $f$) such that*

$$\|u\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C\, \|f\|_{C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)}. \tag{3.4}$$

*Proof.* The solution is unique by the strong maximum principle of Bony [15] (see also Bardi and Da Lio [10]). We show the existence by vanishing viscosity methods, i.e. for all $\varepsilon > 0$ we consider the operator $\mathcal{L}_\varepsilon = -\varepsilon\Delta + \mathcal{L}$, with $\varepsilon > 0$ and the corresponding problem (3.3), replacing $\mathcal{L}$ by $\mathcal{L}_\varepsilon$. Note that $\mathcal{L}_\varepsilon u + \rho u = f$ is a linear uniformly elliptic equation. It is well-known that such a problem has a unique classical solution $u_\varepsilon$, which is of class $C^{2,\alpha}$ since $f \in C^{0,\alpha}$ (see e.g. [11, Lemma 2.7] and [32]). Moreover $\|u_\varepsilon\|_\infty \leq \frac{1}{\rho}\|f\|_\infty$. This implies that (up to a subsequence) $u_\varepsilon \to u$ in the weak*-topology of $L^\infty(\mathbb{T}^d)$. Therefore $u$ is a distributional solution of $\mathcal{L}u + \rho u = f$. Furthermore, if $f$ is smooth then, by Hörmader's hypoellipticity Theorem [38], $u$ is smooth. So let us assume for the moment that $f \in C^\infty(\mathbb{T}^d)$; then $u$ is in particular a classical solution satisfying the assumption of Zuily and Xu [57], thus Theorem 2.2 gives directly estimate (3.4).
If $f \in C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$, one can bypass this obstacle by mollifications and noticing that estimate (3.4) is stable w.r.t. the mollification parameter. More precisely, when $f$ is not smooth but only Hölder, we introduce $f_\zeta := f * \varphi_\zeta$, where $\varphi_\zeta(x) := \zeta^{-d}\varphi(x/\zeta)$ for $\zeta > 0$ and $x \in \mathbb{R}^d$, and $\varphi$ is a mollification kernel, that is, a nonnegative function of class $C^\infty$, with support in the unit ball of $\mathbb{R}^d$ and $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$. One can easily check that $f_\zeta \to f$ as $\zeta \to 0$ in $C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$. Let $\{\zeta_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers converging to zero. For every $n \in \mathbb{N}$ there exists

a unique solution $u_n \in C^{2,\alpha}(\mathbb{T}^d)$ to (3.3) for $f = f_n := f_{\zeta_n}$, and by estimate (3.4), we have $\|u_n - u_m\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C\,\|f_n - f_m\|_{C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)}$ for some constant $C > 0$ that does not depend on $n, m \in \mathbb{N}$. Thus $\{u_n\}_{n\in\mathbb{N}}$ is Cauchy in $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ (using $f_n \to f$ in $C^{\alpha}_{\mathcal{X}}(\mathbb{T}^d)$), hence it converges to some $u$ in $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$. Passing to the limit as $n \to \infty$ in the equation $\mathcal{L}u_n + \rho u_n = f_n$ and in the estimates $\|u_n\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C\|f_n\|_{C^{\alpha}_{\mathcal{X}}(\mathbb{T}^d)}$ we find that $u$ is a solution to (3.3) and that estimate (3.4) is satisfied. □

The existence and uniqueness for the subelliptic linear equation for $m$ is more technical. We first recall some heat kernel estimates and an ergodic result which will be key for the later results. Consider the Cauchy problem

$$\begin{cases} \dfrac{\partial z}{\partial t} - \mathcal{L}z - g \cdot D_{\mathcal{X}} z = 0 \\ z(0,x) = \phi(x) \end{cases} \tag{3.5}$$

where $\phi$ is Borel and bounded and $g$ is Hölder-continuous. Then we have the following representation for the unique solution of (3.5):

$$z(t,x) = \int_{\mathbb{T}^d} K(t,x,y)\phi(y)\,dy\,,$$

where the function $(t,x,y) \mapsto K(t,x,y)$, defined for $t > 0$, $x, y \in \mathbb{T}^d$, $x \neq y$, is the heat kernel associated to the ultraparabolic operator $\partial_t - \mathcal{L} - g \cdot D_{\mathcal{X}}$. We next recall some known Gaussian estimates satisfied by the heat kernel $K(t,x,y)$: there exist constants $C = C(T) > 0$ and $M > 0$ (depending only on the Hölder norm of $g$) such that

$$\frac{C^{-1}}{|B_{d_{CC}}(x,t^{1/2})|} e^{-M\,d_{CC}(x,y)^2/t} \leq K(t,x,y) \leq \frac{C}{|B_{d_{CC}}(x,t^{1/2})|} e^{-M\,d_{CC}(x,y)^2/t}, \tag{3.6}$$

for all $T > t > 0$ and $x \in \mathbb{T}^d$, where by $|B_{d_{CC}}(x,t^{1/2})|$ we indicate the Lebesgue measure of the Carnot-Carathéodory ball centred at $x$ and of radius $R = t^{1/2}$. This estimate has been firstly proved in the subelliptic case by [41] for "sums of squares" operators on compact manifolds and later generalised by many authors: in particular we refer to [16].

We now need to recall the following ergodic result.

**Theorem 3.3** ([12], Theorem II.4.1)**.** *Let $(S, \Sigma)$ be a compact metric space equipped with its Borel $\sigma$-algebra $\Sigma$. Let $P$ be a linear operator defined on the Banach algebra of Borel bounded functions on $S$. We assume that $\|P\| \leq 1$ and $P(1) = 1$, and there exists $\delta > 0$ such that*

$$P\mathbf{1}_E(x) - P\mathbf{1}_E(y) \leq 1 - \delta, \qquad \forall x, y \in S,\ E \in \Sigma\,, \tag{3.7}$$

*where by $\mathbf{1}_E(\cdot)$ we indicate the characteristic function of the Borel set $E$.*

*Under these assumptions there exists a unique probability measure $\pi$ on $S$ such that*

$$\left| P^n\phi(x) - \int_S \phi\,d\pi \right| \leq Ce^{-kn}\|\phi\|_\infty \quad \forall x \in S\,, \tag{3.8}$$

*where $C = 2/(1-\delta)$, $k = -\ln(1-\delta)$. Then the measure $\pi$ is the unique* invariant measure *of the operator $P$, that is the unique probability measure satisfying*

$$\int_S P\phi\,d\pi = \int_S \phi\,d\pi,$$

*for every bounded Borel function $\phi$ on $S$.*

The measure $\pi$ is called the *ergodic measure* of the operator $P$ (for more details on ergodic measure see e.g. [26]). Property (3.8) is a "strong" ergodic property: it implies the convergence

$$\lim_{n\to\infty} P^n\phi = \int_S \phi\,d\pi \qquad \text{uniformly}$$

but also provides an exponential decay estimate on the convergence rate.

*Remark* 3.4. As noted also in [12], when applying the ergodic theorem above usually one checks if the so-called Doeblin condition is satisfied. More precisely, we assume that $(S,\Sigma)$ is equipped with a probability measure $\mu$ and that $P$ has the form

$$P\phi(x) = \int_S k(x,y)\phi(y)\,d\mu(y),$$

for some Borel and bounded kernel $k\colon S\times S\to\mathbb{R}$, and that there exist a set $U$ with $\mu(U)>0$ and $\delta_0>0$ such that (*Doeblin condition*)

$$k(x,y)\geq\delta_0>0 \qquad \forall x\in S,\ y\in U\,. \tag{3.9}$$

It is easy to check that (3.9) implies (3.7) with $\delta=\mu(U)\delta_0$. In fact, using $S=\big(S\cap E\big)\cup\big(S\cap E^c\big)$:

$$P\mathbf{1}_E(x)-P\mathbf{1}_E(y)=1-\int_S k(y,z)\mathbf{1}_E(z)\,dz-\int_S k(x,z)\mathbf{1}_{E^c}(z)\,dz\leq 1-\delta_0\Big[\big|E^c\cap U\big|+\big|E\cap U\big|\Big]=1-\delta_0|U|.$$

Next we show existence and uniqueness for the weak solution of the subelliptic linear equation associated to $m$.

**Lemma 3.5.** *Assume* (2.1) *and that* $g\colon\mathbb{T}^d\to\mathbb{R}^m$ *is Hölder continuous. Then the problem*

$$\begin{cases}\mathcal{L}^*m-div_{\mathcal{X}^*}(mg)=0 \ \ in\ \mathbb{T}^d\,,\\ \displaystyle\int_{\mathbb{T}^d} m\,dx=1\,,\end{cases} \tag{3.10}$$

*has a unique weak solution $m$ in $H^1_{\mathcal{X}}(\mathbb{T}^d)$. Moreover $0<\delta_0\leq m\leq\delta_1$, for some $\delta_1,\delta_0$ depending only on the Hölder norm of $g$ and the coefficients of $\mathcal{L}$ (i.e. the coefficient of the vector fields $X_1,\dots,X_m$).*

A solution $m$ of the PDE in (3.10) is to be understood in the weak (or $H^1_{\mathcal{X}}$) sense, i.e. we define the bilinear form

$$\langle u,v\rangle := \int_{\mathbb{T}^d}\left(-\frac{1}{2}\sum_{i=1}^m X_iu\,X_i^*v-(g\cdot D_{\mathcal{X}}u)\,v\right)dx \tag{3.11}$$

and its dual $\langle u,v\rangle^*:=\langle v,u\rangle$ for all $u,v\in H^1_{\mathcal{X}}(\mathbb{T}^d)$. Then $m$ is a *solution* of the PDE in (3.10) if $\langle m,v\rangle^*=0$ for all $v\in H^1_{\mathcal{X}}(\mathbb{T}^d)$.

*Proof.* The proof follows the approach introduced in [14, Theorem 3.4] for uniformly elliptic operators and in [12, Theorem II.4.2 ]. We want first to show that, for $\eta>0$ large enough and for every $\varphi\in L^2(\mathbb{T}^d)$, the problem

$$\mathcal{L}u-g\cdot D_{\mathcal{X}}u+\eta\,u=\varphi \tag{3.12}$$

is well-posed in $H^1_{\mathcal{X}}(\mathbb{T}^d)$ in the standard weak sense, that is

$$\int_{\mathbb{T}^d}\left(-\frac{1}{2}\sum_{i=1}^m X_iu\,X_i^*v-(g\cdot D_{\mathcal{X}}u)\,v+\eta\,u\,v\right)dx=\int_{\mathbb{T}^d}\varphi\,v\,dx,\quad \forall\, v\in C_0^\infty\big(\mathbb{T}^d\big).$$

The previous well-posedness is proved by standard Hilbert space arguments. In fact, on the space $H^1_{\mathcal{X}}(\mathbb{T}^d)$, we consider the bilinear form

$$\begin{gathered}\langle\cdot,\cdot\rangle_\eta\colon H^1_{\mathcal{X}}(\mathbb{T}^d)\times H^1_{\mathcal{X}}(\mathbb{T}^d)\to\mathbb{R}\\ (u,v)\mapsto\langle u,v\rangle_\eta:=\langle u,v\rangle+\int_{\mathbb{T}^d}\eta\,u\,v\,dx\end{gathered}$$

for all $u, v \in H^1_{\mathcal{X}}(\mathbb{T}^d)$, where $\langle u, v\rangle$ is defined in (3.11). For $\eta > 0$ large enough and for some $c_1 > 0$, $c_2 \geq 0$ we claim that for all $u, v \in H^1_{\mathcal{X}}(\mathbb{T}^d)$

$$\langle u, u\rangle_\eta \geq c_1 \|u\|^2_{H^1_{\mathcal{X}}(\mathbb{T}^d)}, \tag{3.13}$$

$$|\langle u, v\rangle_\eta| \leq c_2 \|u\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)} \|v\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)}. \tag{3.14}$$

We first check estimate (3.13). Since $g$ and $\mathrm{div} X_i$ are by assumption continuous, hence bounded on $\mathbb{T}^d$, there exists $M \geq 0$ such that $\|g\|_\infty \leq M$ and $\|\mathrm{div} X_i\|_\infty \leq M$. Moreover

$$\begin{aligned}\langle u, u\rangle_\eta &= \int_{\mathbb{T}^d} \left( \frac{1}{2}\sum_{i=1}^m |X_i u|^2 + \frac{1}{2}\sum_{i=1}^m X_i u\, (\mathrm{div} X_i) u - \big(g \cdot D_{\mathcal{X}} u\big) u + \eta\, u^2 \right) dx \\ &\geq \int_{\mathbb{T}^d} \left( \frac{1}{2}\sum_{i=1}^m |X_i u|^2 - \frac{1}{2}\sum_{i=1}^m |X_i u| |\mathrm{div} X_i| |u| - |g| |D_{\mathcal{X}} u| |u| + \eta\, u^2 \right) dx\,.\end{aligned}$$

Using the inequality $ab \leq (1/4)a^2 + b^2$ and recalling $|D_{\mathcal{X}} u|^2 = \sum_{i=1}^m |X_i u|^2$, we find

$$\begin{aligned}\langle u, u\rangle_\eta &\geq \int_{\mathbb{T}^d} \left( \frac{1}{2}|D_{\mathcal{X}} u|^2 - \frac{1}{8}|D_{\mathcal{X}} u|^2 - |\mathrm{div} X_i|^2 |u|^2 - \frac{1}{4}|D_{\mathcal{X}} u|^2 - |g|^2|u|^2 + \eta\, u^2 \right) dx \\ &\geq \frac{1}{8}\int_{\mathbb{T}^d} |D_{\mathcal{X}} u|^2\, dx + (\eta - 2M^2) \int_{\mathbb{T}^d} u^2\, dx\,,\end{aligned}$$

from which, taking $\eta > 2M^2$, we obtain the first estimate (3.13) for a suitable $c_1 > 0$ (in particular $c_1 = \min\left\{1/8, \eta - 2M^2\right\} > 0$). For estimate (3.14) similarly

$$|\langle u, v\rangle_\eta| \leq \frac{1}{2}\int_{\mathbb{T}^d} |D_{\mathcal{X}} u| |D_{\mathcal{X}} v|\, dx + M \int_{\mathbb{T}^d} |D_{\mathcal{X}} u| |v|\, dx + \eta \int_{\mathbb{T}^d} |u|\, |v|\, dx\,,$$

and by the Cauchy-Schwarz inequality for integrals

$$\begin{aligned}|\langle u, v\rangle_\eta| &\leq \frac{1}{2}\|D_{\mathcal{X}} u\|_{L^2(\mathbb{T}^d)} \|D_{\mathcal{X}} v\|_{L^2(\mathbb{T}^d)} + (M + \eta) \|u\|_{L^2(\mathbb{T}^d)} \|v\|_{L^2(\mathbb{T}^d)} \\ &\leq \left( M + \eta + \frac{1}{2} \right) \|u\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)}\, \|v\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)}\,,\end{aligned}$$

where we have used simply $a\,b + c\,d \leq (a + c)\,(b + d)$ for every non-negative scalars $a, b, c$ and $d$. This gives (3.14) with $c_2 = \left(M + \eta + \frac{1}{2}\right) > 0$. Then the claim is proved.

Thus the bilinear form $\langle\cdot,\cdot\rangle_\eta$ is coercive and continuous. Clearly,

$$H^1_{\mathcal{X}}(\mathbb{T}^d) \ni u \mapsto \int_{\mathbb{T}^d} u\varphi\, dx\, \in \mathbb{R}$$

is a continuous linear functional on $H^1_{\mathcal{X}}(\mathbb{T}^d)$. Therefore by the Lax-Milgram Theorem there exists a unique $u \in H^1_{\mathcal{X}}(\mathbb{T}^d)$ such that, for all $v \in H^1_{\mathcal{X}}(\mathbb{T}^d)$,

$$\langle u, v\rangle_\eta = \int_{\mathbb{T}^d} \varphi v\, dx.$$

For every $\eta > 0$ large enough (i.e. $\eta > 2M^2$), we define the following linear operator $T_\eta \colon L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ by $T_\eta\varphi := u$, where $u$ is the unique solution to (3.12).

Note that $T_\eta\varphi = u \in H^1_{\mathcal{X}}(\mathbb{T}^d) \subset L^2(\mathbb{T}^d)$. Since the embedding of $H^1_{\mathcal{X}}(\mathbb{T}^d)$ into $L^2(\mathbb{T}^d)$ is compact (see Lemma 2.1), $T_\eta$ is a linear compact operator. Thus the equation

$$\mathcal{L}^* m - \mathrm{div}_{\mathcal{X}^*}(mg) = 0 \quad \text{in } \mathbb{T}^d$$

is equivalent to

$$(I - \eta T_\eta)^* m = 0\,, \tag{3.15}$$

where $I$ is the identity operator of $L^2(\mathbb{T}^d)$.
Since $T_\eta$ is compact, the Fredholm alternative applies. Indeed $I - \eta T_\eta$ is a Fredholm operator of index zero (see e.g. [1, Lemma 4.45]). This means that the kernels of $I - \eta T_\eta$ and $(I - \eta T_\eta)^*$ have the same dimension, in other words, the number of linearly independent solutions of the equation $(I - \eta T_\eta)^* = 0$ is equal to the number of linearly independent solutions of the equation $I - \eta T_\eta = 0$. Then we must find the number of linearly independent solutions of $(I - \eta T_\eta)u = 0$, that is

$$\mathcal{L}u + g(x) \cdot D_{\mathcal{X}} u = 0\,.$$

By [57, Theorem 3.3] the solution $u$ belongs to $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ for some $\alpha \in (0,1)$. Moreover the operator $\mathcal{L} + g \cdot D_{\mathcal{X}}$ satisfies the strong maximum principle, see [15] and [10]. Thus by the considerations above (Fredholm alternative) the equation (3.15), and hence (3.12), admits a unique solution $m \in H^1_{\mathcal{X}}(\mathbb{T}^d)$ up to a multiplicative constant.

The upper and lower bounds for $m$ (that imply in particular the positivity of $m$) are shown by its interpretation as the ergodic measure of the diffusion having generator $\mathcal{L} + g \cdot D_{\mathcal{X}}$. They rely on an ergodic theorem and on the Gaussian estimates (3.6). In fact, using that $d_{CC}(x, y)$ is continuous on $\mathbb{T}^d$ (compact), we can easily see from (3.6) (by simply taking the maximum and the minimum of $d^2_{CC}(x, y)$ on $\mathbb{T}^d$) that there exits $\delta_0, \delta_1 > 0$ such that

$$\delta_0 \le K(1, x, y) \le \delta_1 \qquad \forall x, y \in \mathbb{T}^d. \tag{3.16}$$

Therefore we can apply Theorem 3.3 and Remark 3.4 with $S = \mathbb{T}^d$, $\Sigma$ the Borel $\sigma$-algebra on $\mathbb{T}^d$, $\mu$ the Lebesgue measure on $\mathbb{T}^d$ and operator $P$ defined by

$$P\phi(x) = z(1, x) = \int_{\mathbb{T}^d} K(1, x, y)\phi(y)\, dy\,.$$

Note that $P^n\phi(x) = z(n, x)$. Then Theorem 3.3 implies the existence of a unique invariant probability measure $\pi$ such that

$$\left| z(n, x) - \int_{\mathbb{T}^d} \phi(y)\, d\pi(y) \right| \le Ce^{-kn}\|\phi\|\,. \tag{3.17}$$

Using $m$ defined as the unique solution of (3.10) and $z(t, x)$ defined as unique solution of (3.5), we want to show the following claim:

$$\int_{\mathbb{T}^d} z(t, x)m(x)\, dx = \int_{\mathbb{T}^d} \phi(x)m(x)\, dx,\ \forall\, t \ge 0. \tag{3.18}$$

To prove the previous claim, first note that for $t = 0$ (3.18) is trivially satisfied by the initial condition. We want to show that the right-hand side in (3.18) is constant in time, so we look at

$$\frac{d}{dt}\int_{\mathbb{T}^d} z(t, \cdot)m\, dx = \int_{\mathbb{T}^d} \partial_t z m\, dx = \int_{\mathbb{T}^d} \big(\mathcal{L}z + g \cdot D_{\mathcal{X}} z\big)\, m\, dx = \int_{\mathbb{T}^d} \big(\mathcal{L}^* m - \mathrm{div}_{\mathcal{X}^*}(m\, g)\big) z\, dx = 0.$$

Then $\int_{\mathbb{T}^d} z(t, x)m(x)\, dx = \int_{\mathbb{T}^d} z(0, x)m(x)\, dx$, for all $t \ge 0$, that proves the claim (3.18).

By using (3.17) and taking $t = n$ in (3.18) and passing to the limit as $n \to +\infty$, we can deduce:

$$\int_{\mathbb{T}^d} \phi(x)\, m(x)\, dx = \int_{\mathbb{T}^d} \bigg( \int_{\mathbb{T}^d} \phi(x)\, d\pi(x) \bigg) m(x) dx = \int_{\mathbb{T}^d} \phi(x)\, d\pi(x),$$

for any Borel bounded function $\phi$ on $\mathbb{T}^d$ (where we have used $\int_{\mathbb{T}^d} m = 1$). Thus $m$ is the density measure of the probability measure $\pi$ and therefore $m \ge 0$ a.e. on $\mathbb{T}^d$.

Using (3.16) together with (3.18) for $t = 1$, it follows that

$$\delta_1 \int_{\mathbb{T}^d} \phi(y)\, dy \ge \int_{\mathbb{T}^d} \phi(y)m(y)\, dy \ge \delta_0 \int_{\mathbb{T}^d} \phi(y)\, dy,$$

for any bounded and Borel function $\phi \ge 0$ on $\mathbb{T}^d$. Since $\phi \ge 0$ is arbitrary, one can deduce $\delta_0 \le m \le \delta_1$, thus Lemma 3.5 is proved. □

We can now prove our first existence result for a subelliptic MFG system.

*Proof of Theorem 3.1.* The proof is based on a corollary of Schauder's fixed point theorem. More precisely, we apply [32, Theorem 11.3] which states that, if $T\colon \mathcal{B}\to\mathcal{B}$ is a continuous and compact operator in the Banach space $\mathcal{B}$ such that the set $\{u\in\mathcal{B} \;:\; sTu=u,\ 0\le s\le 1\}$ is bounded, then $T$ has a *fixed point*, that is, there exists $u\in\mathcal{B}$ such that $Tu=u$. We define the Banach space $\mathcal{B}=C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ , where $0<\alpha<1$ is to be fixed later, and the operator $T\colon \mathcal{B}\to\mathcal{B}$ , according to the scheme $v\mapsto m\mapsto u$ . This means that, given $v\in\mathcal{B}$, we solve the second equation together with the corresponding conditions

$$\begin{cases} \mathcal{L}^* m - \mathrm{div}(m g(x, D_{\mathcal{X}}v)) = 0 \ \text{ in } \mathbb{T}^d \,, \\ \displaystyle\int_{\mathbb{T}^d} m\, dx = 1\,, \quad m>0 \ \text{ in } \mathbb{T}^d \end{cases}$$

and by Lemma 3.5 we find a unique solution $m\in H^1_{\mathcal{X}}(\mathbb{T}^d)\cap L^\infty(\mathbb{T}^d)$. Moreover $m$ is bounded. By assumption (**IV**), $V[\cdot]$ is regularizing, hence the function $f(x)=V[m](x)-H(x,D_{\mathcal{X}}v(x))$ belongs to $C^\alpha_{\mathcal{X}}(\mathbb{T}^d)$. Thus we apply Lemma 3.2 and deduce that

$$\mathcal{L}u+\rho u+H(x,D_{\mathcal{X}}v)=V[m] \tag{3.19}$$

admits a unique solution $u\in C^2_{\mathcal{X}}(\mathbb{T}^d)$. Set $Tv=u$, where $u$ is the unique solution of (3.19), it is easy to check that $T$ is continuous and compact, using that $C^2_{\mathcal{X}}(\mathbb{T}^d)$ is compactly embedded into $C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ for all $\alpha\in(0,1)$. Therefore, in order to apply [32, Theorem 11.3] we need to show that

$$\mathcal{A}=\{u\in\mathcal{B} \;:\; \exists\, 0\le s\le 1 \text{ such that } u=sTu\}$$

is bounded in $C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$. So note that: if $u$ is a fixed point of $sT$ (i.e. $sTu=u$), then it is also a solution of

$$\mathcal{L}u+\rho u+sH(x,D_{\mathcal{X}}u)=sV[m]. \tag{3.20}$$

Then looking at the minimum and maximum of $u$, we find

$$\|u\|_{\infty,\mathbb{T}^d}\le \frac{s}{\rho}\sup_{m\in H^1_{\mathcal{X}}(\mathbb{T}^d)}\|V[m]-H(\cdot,0)\|_{\infty,\mathbb{T}^d},$$

which is finite since $V[\cdot]$ is by assumption bounded.

The key step is now to apply $C^{1,\alpha}_{\mathcal{X}}$-regularity for semilinear equation (Theorem 2.3) that gives

$$\|u\|_{C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)}<C \tag{3.21}$$

for some constant $C>0$ and $\alpha\in(0,1)$ independ of $u$ and $s\in[0,1]$.

Note that, in order to apply the given theorem, we should write our equation in divergence form, which we can easily do by using the relation $X_i^*=-X_i-\mathrm{div}X_i$ (by adding the term $-\sum_{j=1}^n(\mathrm{div}X_j)X_j$ to the Hamiltonian). Observe that the new Hamiltonian has the same properties of the original Hamiltonian; in particular, it grows at most quadratically in $D_{\mathcal{X}}u$ (in fact the functions $\mathrm{div}X_j$ are bounded due to the $C^\infty$-regularity of the vector fields $X_j$). Using estimate (3.21) we can look at the semilinear PDE (3.20) as a linear PDE with an Hölder right-hand side $f(x)=sV[m]-sH(x,D_{\mathcal{X}}u)-\sum_{j=1}^m(\mathrm{div}X_j)X_ju$; hence we can apply the Schauder type result for linear equations proved in [57] (see Theorem 2.2), that implies $u\in C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ and $\|u\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)}<C$. To conclude we need only to remark that $\mathrm{div}X_j\in C^{0,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ (since the vector fields are smooth on a compact domain) in order to apply the previous $C^{2,\alpha}_{\mathcal{X}}$-estimates. □

# 4 Ergodic system with linear growth

We now want to study the ergodic problem that can be obtained by letting $\rho \to 0^+$ in (3.2). However, to study this, we need a more restrictive assumption on the Hamiltonian, i.e. we assume that $H$ grows at most linearly in $|D_{\mathcal{X}}u|$. More precisely:

**(II-L)** $H(x,q) = H(x,\sigma(x)p)$ grows at most linearly w.r.t. $q$, i.e. $\exists C \geq 0$ such that

$$|H(x,q)| \leq C(|q|+1) \quad \forall x \in \mathbb{T}^d,\ q \in \mathbb{R}^m. \tag{4.1}$$

We prove existence of solutions for the system of ergodic PDEs under condition **(II-L)**.

**Theorem 4.1** (Existence)**.** *Assume* (2.1)*,* **(II-L)***,* **(III)***,* **(IV)** *and that $H(x,q)$ is locally Hölder, then the system*

$$\begin{cases} \mathcal{L}u + \lambda + H(x, D_{\mathcal{X}}u) = V[m] \\ \mathcal{L}^* m - div_{\mathcal{X}^*}\big(mg(x, D_{\mathcal{X}}u)\big) = 0 \\ \int_{\mathbb{T}^d} u\, dx = 0, \quad \int_{\mathbb{T}^d} m\, dx = 1, \quad m > 0 \end{cases}. \tag{4.2}$$

*has a solution $(\lambda, u, m) \in \mathbb{R} \times C^2_{\mathcal{X}}(\mathbb{T}^d) \times C(\mathbb{T}^d)$.*

*Proof.* For $\rho > 0$ let $(u_\rho, m_\rho) \in C^2_{\mathcal{X}}(\mathbb{T}^d) \times \big(H^1_{\mathcal{X}}(\mathbb{T}^d) \cap L^\infty(\mathbb{T}^d)\big)$ be a solution of (3.2) by the existence result given in Theorem 3.1. Looking at the minima and maxima of $u_\rho$, we have

$$\|\rho u_\rho\|_\infty \leq \sup_{m \in H^1_{\mathcal{X}}(\mathbb{T}^d)} \|H(\cdot, 0) - V[m]\|_\infty\,. \tag{4.3}$$

Let $< u_\rho >:= \int_{\mathbb{T}^d} u_\rho\, dx$ be the average of $u_\rho$, the key estimate is given in the following claim: there exist $\rho_0 > 0$ and $C > 0$ (independent of $\rho$) such that

$$\|u_\rho - < u_\rho > \|_\infty \leq C, \quad \forall\, 0 < \rho < \rho_0. \tag{4.4}$$

To prove (4.4) we adapt some ideas from [5]. Assume by contradiction that there is a sequence $\rho_n \to 0$ such that $\|u_{\rho_n} - < u_{\rho_n} > \|_\infty \to +\infty$ or equivalently, such that the sequence

$$\varepsilon_n := \|u_{\rho_n} - < u_{\rho_n} > \|_\infty^{-1} \to 0.$$

Then the renormalised functions $\psi_n := \varepsilon_n(u_{\rho_n} - < u_{\rho_n} >)$ satisfy

$$\mathcal{L}\psi_n + \varepsilon_n H\left(x, \frac{D_{\mathcal{X}}\psi_n}{\varepsilon_n}\right) + \rho_n \psi_n = \varepsilon_n(V[m_{\rho_n}] - \rho_n < u_{\rho_n} >). \tag{4.5}$$

We now apply [54, Theorem 17] to deduce that the sequence $\{\psi_n\}$ is equi-Hölder continuous. In fact $\psi_n$ solve quasilinear equations of the same form as in [54] with $A_i(x,u,\xi) = \xi_i$ and $B(x,u,\xi) = \varepsilon_n H\left(x, \frac{\xi}{\varepsilon_n}\right) - \rho_n u - \varepsilon_n\big(V[m_{\rho_n}]\big) - \rho_n < u >$; then it is easy to check that all conditions on the equation are satisfied just taking $g = 0$, $f = 1$ and a $\Lambda$ depending only on the bound for $V[\cdot]$, the constant in **(II-L)** and the Lebesgue measure of $\mathbb{T}^d$. Thus [54, Theorem 17] tells us that, taking $\rho_n \leq 1$ and $\varepsilon_n \leq 1$, the Hölder norms of the solutions $\psi_{\rho_n}$ are equi-bounded independently on $n$, which implies that $\psi_{\rho_n}$ are equi-Hölder. Therefore (up to a subsequence) we get that $\psi_n$ converges uniformly to a function $\psi$. Note that the functions $\psi_n$ are all renormalised, then $\|\psi\|_\infty = 1$. Moreover, since $\int_{\mathbb{T}^d} \psi_n\, dx = 0$ by definition, then there exists a point $x_n \in \mathbb{T}^d$ such that $\psi_n(x_n) = 0$. Thus (up to a further subsequence) we get $\psi(\overline{x}) = 0$ for some $\overline{x} \in \mathbb{T}^d$. By using assumption **(II-L)** into equation (4.5), one finds out that $\psi_{\rho_n}$ are classical (and hence viscosity) subsolutions of

$$\mathcal{L}\psi_n - C\,|D_{\mathcal{X}}\psi_n| + \rho_n\psi_n - \varepsilon_n(V[m_{\rho_n}] - \rho_n < u_{\rho_n} > + C) = 0. \tag{4.6}$$

and classical (and hence viscosity) supersolutions of

$$\mathcal{L}\psi_n + C\,|D_{\mathcal{X}}\psi_n| + \rho_n\psi_n - \varepsilon_n(V[m_{\rho_n}] - \rho_n < u_{\rho_n} > -C) = 0. \tag{4.7}$$

Finally by taking $n \to \infty$ in (4.6) and (4.7) and by using the stability for viscosity subsolutions and viscosity supersolutions under uniform convergence (see e.g. [7]), $\psi$ is a viscosity subsolution of $\mathcal{L}\psi - C|D_{\mathcal{X}}\psi| = 0$ and a viscosity supersolution of $\mathcal{L}\psi + C|D_{\mathcal{X}}\psi| = 0$. Since $\mathcal{L}$ is the subelliptic Laplacian associated to smooth Hörmander vector fields and $\psi$ is periodic, we deduce from the strong maximum principle (see [15], [10]) that $\psi$ must be a constant, which contradicts $\|\psi\|_\infty = 1$ and $\psi(\overline{x}) = 0$, proving thus (4.4).

We complete the proof of the theorem by showing that there exists a sequence $\rho_n \to 0$ such that, for $w_\rho := u_\rho - < u_\rho >$,

$$(\rho_n < u_{\rho_n} >, w_{\rho_n}, m_{\rho_n}) \to (\lambda, u, m) \qquad \text{in } \mathbb{R} \times C^2_{\mathcal{X}}(\mathbb{T}^d) \times H^1_{\mathcal{X}}(\mathbb{T}^d), \tag{4.8}$$

where $(\lambda, u, m)$ is a solution of (4.2); the convergence $m_{\rho_n} \to m$ is in the weak topology of $H^1_{\mathcal{X}}(\mathbb{T}^d)$. Indeed, we note that $(w_\rho, m_\rho)$ solves

$$\begin{cases} \mathcal{L}w_\rho + \rho w_\rho + H(x, D_{\mathcal{X}}w_\rho) = V[m_\rho] - \rho < u_\rho > \quad \text{in } \mathbb{T}^d, \\ \mathcal{L}^* m_\rho - \operatorname{div}_{\mathcal{X}^*}(g(x, D_{\mathcal{X}}w_\rho)m_\rho) = 0, \\ \displaystyle\int_{\mathbb{T}^d} m_\rho(x)dx = 1, \quad m_\rho > 0\,. \end{cases} \tag{4.9}$$

By the a-priori Hölder estimates for quasilinear subelliptic equations recalled in Theorem 2.3 we know that

$$\|w_\rho\|_{C^{1,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C,$$

for some $\alpha \in (0,1)$ and $C > 0$ depending only on an upper bound of $\|w_\rho\|_\infty$ and on the data of the problem, in particular on the supremum norm of $V[m_\rho] - \rho < u_\rho >$, which is bounded uniformly in $\rho$ by **(IV)** and (4.3). In other words, $\alpha$ and $C$ can be chosen independent of $\rho$. Next by Schauder local estimates for subelliptic linear equations [57, Theorem 3.5], we have

$$\|w_\rho\|_{C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)} \leq C, \tag{4.10}$$

for some $C$ and $\alpha$ independent of $\rho$. On the other hand, by Lemma 3.5 and assumption **(III)**

$$\|m_\rho\|_{H^1_{\mathcal{X}}(\mathbb{T}^d)} \leq C, \tag{4.11}$$

for $C \geq 0$ independent of small enough $\rho$. Since $C^{2,\alpha}_{\mathcal{X}}(\mathbb{T}^d)$ is compactly embedded into $C^2_{\mathcal{X}}(\mathbb{T}^d)$, the previous estimates (4.10), (4.11) and the fact that the set $\{\rho < u_\rho > : \ \rho > 0\}$ is bounded (in $\mathbb{R}$) by (4.3), we can extract a sequence $\rho_n \to 0$ such that (4.8) holds. Furthermore, since $g$ is locally Hölder by assumption **(III)** and $D_{\mathcal{X}}w_{\rho_n} \to D_{\mathcal{X}}u$ in $C^1_{\mathcal{X}}(\mathbb{T}^d)$, then $g_n := g(\cdot, D_{\mathcal{X}}w_{\rho_n}(\cdot)) \to g(\cdot, D_{\mathcal{X}}u(\cdot))$. Let $\langle\cdot,\cdot\rangle_n$ denote the bilinear form associated with $g_n$ in the same fashion as $\langle\cdot,\cdot\rangle$ denotes the bilinear form associated with $g$ after the statement of Lemma 3.5. Since $m_{\rho_n}$ is the solution of the second equation in (4.9), $\langle m_{\rho_n}, \varphi\rangle_n^* = 0$ for all $\varphi \in H^1_{\mathcal{X}}(\mathbb{T}^d)$. From this and the fact that $g_n \to \bar{g}(\cdot, D_{\mathcal{X}}u(\cdot))$ in $L^2(\mathbb{T}^d)$, it is fairly easy to deduce that $\langle m, \varphi\rangle^* = 0$ for all $\varphi \in H^1_{\mathcal{X}}(\mathbb{T}^d)$. Thus $m$ is a solution of the second equation in (4.2). The normalising conditions in the third row of (4.2) are clearly preserved in the limit. Thus the triplet $(\lambda, u, m)$ is indeed a solution of (4.2). □

Exactly as in the elliptic case, both the previous MFG systems have unique solutions under suitable monotonicity assumptions.

Recall that an operator $V$, defined on some subset of $L^2(\mathbb{T}^d)$ with values in $L^2(\mathbb{T}^d)$, is *monotone* if $\int_{\mathbb{T}^d} (V[m_1] - V[m_2])\,(m_1 - m_2)\,dx \geq 0$, $\forall m_1, m_2$, and it is *strictly monotone* if the inequality is strict for all $m_1 \neq m_2$. Given a function $H\colon \mathbb{T}^d \to \mathbb{R}$ and a vector-valued map $g\colon \mathbb{T}^d \to \mathbb{R}^m$, we say that $H$ is *g-convex* if $H(q_2) - H(q_1) - g(q_1)\cdot(q_2 - q_1) \leq 0$, for all $q_1, q_2 \in \mathbb{T}^d$. If the inequality is strict for $q_1 \neq q_2$, $H$ is *strictly g-convex*.

**Theorem 4.2** (Uniqueness). *Assume that one of the two following assumptions holds:*

**(i)** *$V$ is monotone in $L^2$ and $H$ is strictly $(-g)$-convex, or*

**(ii)** *$V$ is strictly monotone in $L^2$ and $H$ is $(-g)$-convex.*

*Then the system* (4.2) *has a unique weak solution.*

The proof is standard so we omit it.

*Remark* 4.3.

1. Hamiltonians $H$ coming from optimal control are "$(-g)$-convex" and, under suitable assumptions, strictly $(-g)$-convex.
2. The strict $(-g)$-convexity can be relaxed requiring that $H(q_2) - H(q_1) - g(q_1) \cdot (q_2 - q_1) \leq 0$, implies $g(x, q_1) = g(x, q_2)$, instead of $q_1 = q_2$. In this way one can cover also the case $H(x, q) = |q|$ and $g(x, q) = -q/|q|$ for $q \neq 0$, $g(x, 0) = 0$.
3. Similarly one can state the uniqueness for the "discounted" system (3.2).

# 5 Appendix

Here we want to show briefly some applications to stochastic differential games, which motivate the study of our MFG system. These applications are standard and the Hörmander degenerate case is similar to the known uniformly elliptic case. We include them for completeness but omitting all details and proofs.

## 5.1 The optimal-control-fixed-point problem of MFG theory

The heuristics of Mean Field Games leads to a mathematical problem that consists of an optimal control problem followed by a fixed point problem. This heuristics is well explained in the literature, for example, in [43], [45], [18] and [4]. Moreover, the relation between $N$-player games and Mean Field Games has been considered in the literature since the very beginnings, [43, 44, 45], [39, 40]; see also [31] where results have been extended to several homogeneous populations of agents for ergodic problems. Recently the time-dependent case has been addressed by Cardaliaguet, Delarue, Lasry and Lions in [20]. For the first mathematical problem hinted above, we give a short self-contained description based mainly on the notes of Cardaliaguet [18] and some comments in the introduction of Araposthatis et al. [4].

As in the previous section, we consider a family of $m$ smooth vector fields $\mathcal{X} = \{X_1, \ldots, X_m\}$, $m \in \mathbb{N}$, satisfying the Hörmander condition (2.1) on the $d$-dimensional torus $\mathbb{T}^d$, and a map $V \colon P(\mathbb{T}^d) \to P(\mathbb{T}^d)$ that satisfies condition **(IV)**. Let $W_t = (W_t^1, \ldots, W_t^m)$ be an $\mathbb{R}^m$-valued Brownian motion in a complete filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F})_{t\in\mathbb{R}_+}, \mathbb{P}\big)$. The Brownian motion is assumed adapted with respect to the filter $(\mathcal{F})_{t\in\mathbb{R}_+}$ and the filter is required to satisfy the so-called *standard assumptions*, see e.g. [12], [51].

We consider the stochastic differential equation

$$\begin{cases} d\xi_t = \displaystyle\sum_{k=1}^{m} b^k(\xi_t, \alpha_t) X_k(\xi_t) dt + \sum_{k=1}^{m} X_k(\xi_t) \circ dW_t^k \,, \\ \xi_0 = x_0 \,, \end{cases} \tag{5.1}$$

where the notation "$\circ$" denotes Stratonovich integration and $x_0 \in \mathbb{T}^d$ is some fixed *initial condition*. The *controls* $\alpha = (\alpha^1, \ldots, \alpha^m) \colon [0, \infty) \times \Omega \to A$ are measurable $(F_t)_{t\geq 0}$-adapted maps taking values in some metric space $A$, while $\mathcal{A}$ denotes the set of admissible controls. We assume that the *drift* $b \colon \mathbb{T}^d \times A \to \mathbb{R}^m$ is Lipschitz continuous, locally in $a \in A$, then the *cost functional* is given by

$$J(\alpha, m) = \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \big(L(\xi_t, \alpha_t) + V[m](\xi_t)\big)\, dt\right], \tag{5.2}$$

for all $\alpha \in \mathcal{A}$, where $0 \le t \mapsto x(t) \in \mathbb{T}^d$ is the solution to (5.1) corresponding to the control $\alpha$. Under our assumptions this solution is uniquely determined by $\alpha$ and the initial condition $x_0 \in \mathbb{T}^d$. We will omit to write explicitly the dependence on $x_0$ in the functional $J$ since the *optimal value* is indeed independent of $x_0$. We assume that the Lagrangian $L\colon \mathbb{T}^d \times \mathbb{R}^m \to \mathbb{R}$ is measurable and locally bounded and that $V\colon P(\mathbb{T}^d) \to L^\infty(\mathbb{T}^d)$ is measurable. The standard MFG theory leads to the following mathematical problem:

(**P**): Find a pair $(m, \hat\alpha) \in P(\mathbb{T}^d) \times \mathcal{A}$ such that

1. $\hat\alpha = \hat\alpha(m)$ minimizes $J(\cdot, m)$ among $\alpha \in \mathcal{A}$,
2. $m$ is the ergodic measure of the *optimal dynamic* $\hat x(\cdot)$ corresponding to the *optimal control* $\hat\alpha$, i.e. the solution to (5.1) for $\alpha = \hat\alpha$ with initial state $x_0$.

The Hamiltonian has the standard structure

$$H(x,q) = \sup_{a\in A}\big( -b(x,a)\cdot q - L(x,a)\big) \qquad \forall x \in \mathbb{T}^d\,,\ q \in \mathbb{R}^m\,. \tag{5.3}$$

and we assume that there exists $\alpha\colon \mathbb{T}^d \times \mathbb{R}^m \to A$ Lipschitz continuous, locally in $q \in \mathbb{R}^m$ and such that $\forall x \in \mathbb{T}^d, q \in \mathbb{R}^m$ the function$A \ni a \mapsto -b(x,a)\cdot q - L(x,a) \in \mathbb{R}$ attains a maximum at $\bar\alpha(x,q)$. Finally the auxiliary map $g\colon \mathbb{T}^d \times \mathbb{R}^m \to \mathbb{R}^m$ is defined by $g(x,q) = b\big(x, \bar\alpha(x,q)\big)$, for all $x \in \mathbb{T}^d\,, q \in \mathbb{R}^m$. Note that by the above assumptions $g$ is Lipschitz, locally in $q \in \mathbb{R}^m$.

**Lemma 5.1** (Verification theorem)**.** *Under all previous assumptions, let* $(\lambda, u, m) \in \mathbb{R}\times C^2_{\mathcal{X}}(\mathbb{T}^d) \times P(\mathbb{T}^d)$ *be a solution to* (4.2) *and* $\hat\alpha$ *be the admissible control corresponding to the* feedback control $\bar\alpha\big(x, D_{\mathcal{X}}u(x)\big)$*, that is* $\hat\alpha_t = \bar\alpha\big(\hat\xi, D_{\mathcal{X}}u(\hat\xi_t)\big)$ *for all* $t \in [0,+\infty)$*, where* $\hat\xi$ *is the solution to* (5.1) *for* $\alpha_t = \hat\alpha_t$ *and for some* $x_0 \in \mathbb{T}^d$*, then the pair* $\big(\hat\alpha, m\big)$ *is a solution to problem* $(P)$*. Moreover* $\lambda = J(\hat\alpha, m)$*.*

*Proof.* The proof is trivial but we briefly sketch the main steps for sake of completeness. As consequence of Itô formula and the first equation in (4.2), the first property in problem (**P**) is satisfied. Then every Markov process with a compact state space has an invariant measure (e.g. see [30, Theorem 9.3]). In particular, the diffusion $\hat\xi$ has an invariant measure. This invariant (actually, ergodic) measure is a weak solution of the dual operator of the generator of $x \mapsto \hat\xi^x$, that means of the dual of $-\mathcal{L} + b\big(x, \bar\alpha(x, D_{\mathcal{X}}u(x))\big)\cdot D_{\mathcal{X}} = -\mathcal{L} - g(x, D_{\mathcal{X}}u)\cdot D_{\mathcal{X}}$. In other words, this invariant measure is a solution of equation (3.10). But by Lemma 3.5 that equation has a unique solution, so the invariant measure of $\hat\xi$ is precisely $m$. Then also condition 2 is satisfied. □

Combining together Theorem 4.1 and Lemma 5.1 one can derive the following result.

**Corollary 5.2** (Existence of solutions to the MFG problem)**.** *Under the assumptions of Lemma 5.1 and assuming in addition that* $L$ *is locally Hölder continuous, then for every* $x_0 \in \mathbb{T}^d$ *there exists a solution* $(\hat\alpha, m)$ *to problem* (**P**)*.*

## 5.2 Nash equilibria for a class of $N$-player games

The solvability theory for systems similar to (4.2) can be applied to build Nash equilibria in feedback form for a class of stochastic differential $N$-player games. This is a straightforward adaptation to the case of Hörmander diffusions of the results contained in [11]. We include the main ideas for completeness. Let $N \in \mathbb{N}$, $W^i_t$ be a $\mathbb{R}^{m_i}$-valued Brownian motion, for every $i = 1,\dots,N$ and for some $m_i \in \mathbb{N}$, adapted to the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\in\mathbb{R}_+}, \mathbb{P})$ and assume that $W^1_t,\dots,W^N_t$ are independent, then the dynamic of the game is described by the system of SDEs

$$\begin{cases} d\xi^i_t = \displaystyle\sum_{k=1}^{m_i} b^k_i(\xi^i_t, \alpha^i_t) X^i_k(\xi^i_t)dt + \sum_{k=1}^{m_i} X^i_k(\xi^i_t)\circ dW^i_t \quad \text{in } \mathbb{T}^{d_i} \\ \xi^i_0 = x^i_0 \end{cases}, \qquad i = 1,\dots,N, \tag{5.4}$$

where $\xi^i$ is the state of the $i$-th player, $x_0^i$ are given initial conditions, $A_i$ is a given metric space and the *set of control parameters of player* $i$ and each admissible control (namely also strategy) of player $i$, $\alpha^i\colon\ :\mathbb{R}_+\times\Omega\to A_i$, is a measurable and locally bounded map adapted to $W_t^i$. Let $\mathcal{A}_i$ denotes the set of all admissible controls for player $i$, assume that $\mathcal{X}_i=\{X_1^i,\dots,X_{m_i}^i\}$ is a set of smooth Hörmander vector fields on the flat torus $\mathbb{T}^{d_i}$ for some $d_i\in\mathbb{N}$, and the drift $b_i\colon\mathbb{T}^{d_i}\times A_i\to\mathbb{R}^{m_i}$ is a locally Lipschitz map. Under these assumptions, it is known that for any $N$-tuple of initial conditions $(x_0^1,\dots,x_0^N)$ and for every $N$-tuple of admissible controls $(\alpha^1,\dots,\alpha^N)\in\mathcal{A}_1\times\cdots\times\mathcal{A}_N$ there exists a unique solution $\xi=(\xi^1,\dots,\xi^N)$ to (5.4). Actually, the system (5.4) is "decoupled" in the sense that each SDE for $\xi^i$ is solved independently of all the other equations and the stochastic processes $\xi^1,\dots,\xi^N$ are independent of each other: each $\xi^i$ is adapted to its "own" Brownian motion $W_t^i$. The *cost* (or *performance criterion*) of player $i$ is given by

$$J_i(\alpha^1,\dots,\alpha^N)=\lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\int_0^T\big(L_i(\xi_t^i,\alpha_t^i)+F_i(\xi_t^1,\dots,\xi_t^N)\big)\,dt\right], \tag{5.5}$$

where $L_i\colon\mathbb{T}^{d_i}\times A_i\to\mathbb{R}$ and $F_i:\mathbb{T}^{d_1}\times\cdots\times\mathbb{T}^{d_N}\to\mathbb{R}$ are Hölder continuous. Each player seeks to optimise its performance criterion (minimising the cost) in presence of all the competitors. Clearly, the agents have conflicting goals: a *win-win* set of strategies that satisfies all players, i.e minimises all their costs simultaneously, in general will not exist. In these types of problems, a good notion of solution turns out to be the notion of Nash equilibrium: a set of admissible strategies $(\hat\alpha^1,\dots,\hat\alpha^N)\in\mathcal{A}_1\times\cdots\times\mathcal{A}_N$ is called a *Nash equilibrium* if, for every $i=1,\dots N$, $J_i(\hat\alpha^1,\dots,\hat\alpha^N)\le J_i(\hat\alpha^1,\dots,\hat\alpha^{i-1},\alpha^i,\hat\alpha^{i+1}\dots\hat\alpha^N)$. In other words, the player $i$ cannot "perform better" by moving away from $\hat\alpha^i$ if the opponents continue to stick to $(\hat\alpha^1,\dots,\hat\alpha^N)$. The problem of finding Nash equilibria reduces to finding solutions to a system of $2N$ PDEs, made by $N$ equations of HJB type coupled with $N$ equations of KFP type:

$$\begin{cases}\mathcal{L}_iu_i+\lambda_i+H_i(x,D_{\mathcal{X}_i}u_i)=V^i[m_1,\dots,m_N] & \text{in }\mathbb{T}^{d_i},\\ \mathcal{L}_i^*m_i+\operatorname{div}_{\mathcal{X}_i^*}\big(m_ig_i(x_i,D_{\mathcal{X}_i}u_i)\big)=0 & \text{in }\mathbb{T}^{d_i},\\ \displaystyle\int_{\mathbb{T}^{d_i}}u_i\,dx^i=0\,,\quad \int_{\mathbb{T}^{d_i}}m_i\,dx^i=0, & m_i>0,\end{cases}\qquad i=1,\dots,N\,, \tag{5.6}$$

where $\lambda_i\in\mathbb{R}$, $H_i(x,q)=\sup_{a\in A_i}\big(-b_i(x,a)\cdot q-L_i(x,q)\big)$, $\mathcal{L}_i=-\frac12\sum_{k=1}^{m_i}(X_k^i)^2$ (with dual operator $\mathcal{L}_i^*$), the auxiliary maps $g_i\colon\mathbb{T}^{d_i}\times\mathbb{R}^{m_i}\to\mathbb{R}^{m_i}$ are defined as $g_i(x,q)=b_i\big(x,\bar\alpha(x,q)\big)$ $\forall x\in\mathbb{T}^{d_i}$, $q\in\mathbb{R}^{m_i}$ and the operators $V_i\colon\prod_{\substack{1\le j\le N\\ j\neq i}}P(\mathbb{T}^{d_i})\to L^\infty(\mathbb{T}^{d_i})$ are defined by

$$V_i[m_1,\dots,m_{i-1},m_{i+1},\dots,m_N](x)=\int_{\prod_{\substack{j=1\\ j\neq i}}^N\mathbb{T}^{d_i}}F(x_1,\dots,x_N)\prod_{\substack{1\le j\le N\\ j\neq i}}m_j(dx^j)\,. \tag{5.7}$$

**Theorem 5.3** (PDEs and Nash equilibria). *Assume that there exist maps $\bar\alpha^i\colon\mathbb{T}^{d_i}\times\mathbb{R}^{m_i}\to A_i$, Lipschitz continuous, such that $\forall x\in\mathbb{T}^{d_i},q\in\mathbb{R}^{m_i}$ $\bar\alpha^i(x,q)$ is a maximum point of $A_i\ni a\mapsto -b_i(x,a)\cdot q-L_i(x,\alpha)\in\mathbb{R}$, and the Hamiltonians $H_i$ grow at most linearly in the gradient variable, uniformly w.r.t. $x$, then*

- *(i) There exists a solution $\lambda_i\in\mathbb{R}$, $u_i\in C^2_{\mathcal{X}_i}(\mathbb{T}^d)$, $m_i\in H^1_{\mathcal{X}_i}(\mathbb{T}^{d_i})\cap L^\infty(\mathbb{T}^{d_i})$, $i=1,\dots,N$ to system* (5.6).
- *(ii) Every solution of* (5.6) *determines a Nash equilibrium in* feedback form *by $\hat\alpha^i(x)=\bar\alpha^i\big(x,D_{\mathcal{X}_i}u_i(x)\big)$, for the game described above. Moreover $\lambda_i=J_i(\hat\alpha^1,\dots,\hat\alpha^N)$ and $m_i$ is the ergodic measure of $\hat\xi^i$, where $\hat\xi=(\hat\xi^1,\dots,\xi^N)$ is the optimal dynamic* (5.1) *corresponding to the optimal control $\hat\alpha=(\hat\alpha^1,\dots,\hat\alpha^N)$.*

Statement (i) is a corollary of Theorem 4.1. Statement (ii) is the analog of a so-called verification theorem in optimal control and differential games. As hinted in [45] the proof is standard and relies here also on an ergodic theorem for Hörmander diffusions. See [11] for more details in the case of uniformly elliptic diffusions.

## 5.3 Mean Field Games as limit of $N$-player games

Considering the same game as in the previous subsection, the goal is to let the number of the players $N \to \infty$ . In order to be able to pass to the limit as $N \to \infty$ we have to make two additional assumptions.

1. The players are *similar*. Mathematically this means that we are assuming that all $d_i$, $\mathcal{X}_i$, $H_i$, $L_i$, $F_i$ are the same, that is, independent of $i$. Being similar, the agents will reason "similarly", so we can assume also that $\bar{\alpha}^i = \bar{\alpha} \quad \forall i = 1, \dots, N$. As a consequence all $H_i$ and all $g_i$ will be the same, and we can call them $H$ and $g$.

2. Since players are "small" and their number is "large", each player can only have a "statistical visibility" of the game, each player cannot know all the individual states of the agents taking part in the game, but he knows, for example, the average or their states (some "macroeconomic parameter" say, that can somehow be measured or estimated). Mathematically this can be expressed assuming:

$$F(x^1, \dots, x^N) = W\left[\frac{1}{N}\sum_{j=1}^{N}\delta_{x^j}\right](x^i), \qquad \forall (x^1, \dots x^N) \in (\mathbb{T}^d)^N,$$

   for some map $W \colon P(\mathbb{T}^d) \to L^\infty(\mathbb{T}^d)$ , which we assume satisfying condition **(IV)**, (and $\delta_x$ denotes the usual Dirac delta measure). Thus we are assuming that each agent designs his cost as a function of the *empirical average* $\frac{1}{N}\sum_{j=1}^N \delta_{x^j}$ of the states of all the agents and of its own state $x^i$ (after all, it is reasonable to expect that each knows at least his own state $x^i$ and treats it separately from the states of the rest of the agents)

Under these assumptions system (5.6) reduces to

$$\begin{cases} \mathcal{L}u_i + \lambda_i + H(x, D_{\mathcal{X}}u_i) = V[m_1, \dots, m_N] \quad \text{in } \mathbb{T}^d, \\ \mathcal{L}^* m_i - \operatorname{div}_{\mathcal{X}^*}\big(m_i g(x_i, D_{\mathcal{X}}u_i)\big) = 0 \quad \text{in } \mathbb{T}^d \qquad i = 1, \dots, N\,, \\ \displaystyle\int_{\mathbb{T}^d} u_i\, dx = 0\,, \quad \int_{\mathbb{T}^d} m_i\, dx = 0, \quad m_i > 0, \end{cases} \tag{5.8}$$

with

$$V[m_1, \dots, m_N](x^i) = \int_{\mathbb{T}^{d(N-1)}} W\left[\frac{1}{N}\sum_{j=1}^{N}\delta_{x^j}\right](x^i) \prod_{\substack{1\le j\le N \\ j\neq i}} m_j(dx^j) \quad \forall x^i \in \mathbb{T}^d\,. \tag{5.9}$$

*Remark* 5.4 (Existence of symmetric solutions). Under the above assumptions it is easy to adapt the proof of Theorem 5.3 (i) or of Theorem 4.1 in order to show that the system of PDEs (5.8) has a symmetric solution $(\lambda, \dots, \lambda) \in \mathbb{R}^N$, $(u, \dots, u) \in C^2_{\mathcal{X}}(\mathbb{T}^d)$ and $(m, \dots, m) \in \big(H^1_{\mathcal{X}} \cap L^\infty(\mathbb{T}^d)\big)^N$.

System (5.8) does not have a unique solution in general. However, all its solutions symmetrise as $N \to \infty$, (as shown in the following theorem). Moreover, the limit points of these solutions satisfy the system of MFG equation. We recall that $P(\mathbb{T}^d) \subset C(\mathbb{T}^d)^*$ is a compact topological space for the topology of weak $^*$-convergence (Prokhorov's Theorem). Moreover, this topology is metrizable, for example, by the *Kantorovich-Rubinstein* distance $\mathbf{d}(m_1, m_2)$ defined for $m_1, m_2 \in P(\mathbb{T}^d)$ as $\mathbf{d}(m_1, m_2) = \sup\left\{\int_{\mathbb{T}^d} f(x)\, d(m_1 - m_2) \ : \ f \in C^{0,1},\ \operatorname{Lip}(f) \le 1\right\}$.

**Theorem 5.5** (Symmetrisation and MFG limit)**.** *Let* $(\lambda_1^N, \dots, \lambda_N^N) \in \mathbb{R}^N$, $(u_1^N, \dots, u_N^N) \in C^2_{\mathcal{X}}(\mathbb{T}^d)$, $(m_1^N, \dots, m_N^N) \in \big(H^1_{\mathcal{X}}(\mathbb{T}^d) \cap L^\infty(\mathbb{T}^d)\big)^N$ *be a solution to* (5.6) - (5.9)*. Then*

*(i)* $\{(\lambda_i^N, u_i^N, m_i^N)\}_{N\ge i}$ *is precompact in* $\mathbb{R} \times C^2_{\mathcal{X}}(\mathbb{T}^d) \times P(\mathbb{T}^d)$ *for every* $i \in \mathbb{N}$.

*(ii) (symmetrisation:)* $\lim_{N\to\infty}\Big(|\lambda_i^N - \lambda_j^N| + \|u_i^N - u_j^N\|_{C^2_{\mathcal{X}}(\mathbb{T}^d)} + d(m_i^N, m_j^N)\Big) = 0\,.$

*(iii) Let $(\lambda, u, m)$ be a limit point of $\{(\lambda_i^N, u_i^N, m_i^N)\}_{N\geq i}$ for some $i \in \mathbb{N}$. Then $(\lambda, u, m)$ is a solution of the MFG system*

$$\begin{cases} \mathcal{L}u + \lambda + H(x, D_{\mathcal{X}}u) = W[m] \ \textit{in}\ \mathbb{T}^d \\ \mathcal{L}^* u - div_{\mathcal{X}^*}\big(mg(x, D_{\mathcal{X}}u)\hat{\alpha})\big) = 0 \ \textit{in}\ \mathbb{T}^d \\ \displaystyle\int_{\mathbb{T}^d} u\, dx = 0, \quad \int_{\mathbb{T}^d} m\, dx = 1, \quad m > 0\,. \end{cases} \tag{5.10}$$

*Proof.* (i) is a consequence of the a priori estimates for solutions of system (5.8), which one can easily show that under the current assumptions hold true with constants independent of $N$. As hinted in [43], the proof of (ii) relies on the uniqueness and continuous dependence of solutions of HJB and KFP equations on the data, while for the proof of (iii) one needs also a law or large numbers. For additional details we refer to [31]. □

*Remark* 5.6. (i) If the MFG system has a unique solution $(\lambda, u, m)$ (see Theorem 4.2 for sufficient conditions), then clearly $(\lambda_i^N, u_i^N, m_i^n) \to (\lambda, u, m)$, as $N \to \infty$, for every $i \in \mathbb{N}$.

(ii) Solving the system of PDEs (5.10), for every $\varepsilon > 0$, we can build symmetric $\varepsilon$-Nash equilibria for the $N$-player game, provided $N$ is sufficiently large.

# References


[1] Y. A. Abramovich, C. D. Aliprantis. *An invitation to operator theory*, volume 50 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[2] Y. Achdou, M. Bardi, M. Cirant. Mean field games models of segregation. *Mathematical Models and Methods in Applied Sciences*, 27(01):75–113, 2017.

[3] Y. Achdou, P.-N. Giraud, J.-M. Lasry, P.-L. Lions. A Long-Term Mathematical Model for Mining Industries. *Applied Mathematics & Optimization*, 74(3), 579-618, 2016.

[4] M. Arapostathis, A. Biswas, J. Carroll. On solutions of mean field games with ergodic cost. *Comm. Partial Differential Equations*, 23(11-12):2187–2217, 2016.

[5] M. Arisawa, P.-L. Lions. On ergodic stochastic control. *Comm. Partial Differential Equations*, 23(11-12):2187–2217, 1998.

[6] Z.M. Balogh, A. Calogero, R. Pini. The Hopf-Lax formula in Carnot groups: a control theoretic approach, *Calc. Var. Partial Differential Equations*, 49 (2014), no. 3-4, 1379–1414.

[7] M. Bardi, I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.

[8] M. Bardi, P.E. Caines, I. Capuzzo Dolcetta (eds.). Special issue on mean field games, *Dyn. Games Appl.*, 3 (2013), no. 4.

[9] M. Bardi, P.E. Caines, I. Capuzzo Dolcetta (eds.), 2nd Special issue on mean field games, *Dyn. Games Appl.*, 4 (2014), no. 2.

[10] M. Bardi, F. Da Lio. Propagation of maxima and strong maximum principle for viscosity solutions of degenerate elliptic equations. I. Convex operators. *Nonlinear Anal.*, 44(8, Ser. A: Theory Methods):991–1006, 2001.

[11] M. Bardi, E. Feleqi. Nonlinear elliptic systems and mean-field games. *NoDEA Nonlinear Differential Equations Appl.*, 23(4):Art. 44, 32, 2016.

[12] A. Bensoussan. *Perturbation methods in optimal control.* Wiley/Gauthier-Villars Series in Modern Applied Mathematics. John Wiley & Sons, Ltd., Chichester; Gauthier-Villars, Montrouge, 1988. Translated from the French by C. Tomson.

[13] A. Bensoussan, J. Frehse, P. Yam. *Mean field games and mean field type control theory.* Springer, New York, 2013.

[14] A. Bensoussan, J.-L. Lions, G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam-New York, 1978.

[15] J.-M. Bony. Principe du maximum, inégalite de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés. *Ann. Inst. Fourier (Grenoble)*, 19(fasc. 1):277–304 xii, 1969.

[16] M. Bramanti, L. Brandolini, E. Lanconelli, F. Uguzzoni. Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities. *Mem. Amer. Math. Soc.*, 204(961):vi+123, 2010.

[17] I. Capuzzo Dolcetta, H. Ishii. The Hopf solution of Hamilton - Jacobi equations, *Elliptic And Parabolic Problems, Proceedings Of The 4th European Conference*, 2001.

[18] P. Cardaliaguet. Notes on mean field games. From P-L. Lions' lectures at Collège de France, 2010.

[19] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013.

[20] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions. The master equation and the convergence problem in mean field games. *arXiv preprint arXiv:1509.02505*, 2015.

[21] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta. Long time average of mean field games. *Netw. Heterog. Media*, 7(2):279–301, 2012.

[22] R. Carmona. *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications.* SIAM, 2016.

[23] A. Cesaroni, N. Dirr, C. Marchi. Homogenization of a mean field game system in the small noise limit. *SIAM Journal on Mathematical Analysis*, 48(4):2701–2729, 2016.

[24] M. Cirant. Stationary focusing mean-field games. *Comm. Partial Differential Equations*, 41(8):1324–1346, 2016.

[25] A. Cutrì, F. Da Lio. Comparison and existence results for evolutive non-coercive first-order Hamilton-Jacobi equations *ESAIM: Control, Optimisation and Calculus of Variations*, 13, (2007) n. 3, 484-502

[26] G. Da Prato. *An introduction to infinite-dimensional analysis.* Universitext. Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[27] E. Di Nezza, G. Palatucci, E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[28] N. Dirr, F. Dragoni, P. Mannucci, C. Marchi. Stochastic homogenization for functionals with anisotropic rescaling and non-coercive Hamilton-Jacobi equations. *arXiv preprint, arXiv:1707.00553* .

[29] F. Dragoni. Metric Hopf-Lax formula with semicontinuous data, *Discrete Contin. Dyn. Syst.*, 17 (2007), no. 4, 713–729.

[30] S. N. Ethier, T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[31] E. Feleqi. The derivation of ergodic mean field game equations for several populations of players. *Dyn. Games Appl.*, 3(4):523–536, 2013.

[32] D. Gilbarg, N. S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[33] D.A. Gomes, J. Mohr, R. Rigao Souza. Continuous time finite state mean field games. *Applied Mathematics & Optimization*, 68(1):99–143, 2013.

[34] D.A. Gomes, L. Nurbekyan, and E. Pimentel. Economic models and mean-field games theory. In *30o Co16quio Brasileiro de Matemática.[30th Brazilian Mathematics Colloquium]*, 2015.

[35] D.A. Gomes, S. Patrizi, V. Voskanyan. On the existence of classical solutions for stationary extended mean field games. *Nonlinear Analysis: Theory, Methods & Applications*, 99:49–79, 2014.

[36] D.A. Gomes, J. Saúde. Mean field games models – a brief survey. *Dyn. Games Appl.* 4 (2014), no. 2, 110–154.

[37] O. Guéant, J-M. Lasry, P.-L. Lions. *Mean field games and applications*, in "Paris-Princeton Lectures on Mathematical Finance 2010", R.A. Carmona et al. eds., vol. 2003 of Lecture Notes in Math., p. 205–266, Springer, Berlin, 2011.

[38] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[39] M. Huang, P.E. Caines, R.P. Malhamé. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6 (2006), 221–251.

[40] M. Huang, P.E. Caines, R.P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized $\epsilon$-Nash equilibria. *IEEE Trans. Automat. Control*, 52 (2007), 1560–1571.

[41] D.S. Jerison, A. Sánchez-Calle. Estimates for the heat kernel for a sum of squares of vector fields. *Indiana Univ. Math. J.*, 35(4):835–854, 1986.

[42] A. Lachapelle, J.-M. Lasry, C.-A. Lehalle and P.-L. Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *Mathematics and Financial Economics*, 10(3), 223-262, 2016.

[43] J.-M. Lasry, P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[44] J.-M. Lasry, P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

[45] J.-M. Lasry, P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[46] P.-L. Lions, J.-M.ZLasry. Instantaneous self-fulfilling of long-term prophecies on the probabilistic distribution of financial asset values. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, (Vol. 24, No. 3, pp. 361-368). Elsevier Masson.

[47] Guozhen Lu. Weighted Poincare and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications. *Rev. Mat. Iberoamericana*, 8(3):367–439, 1992.

[48] Guozhen Lu. Embedding theorems into lipschitz and bmo spaces and applications to quasilinear subelliptic differential equations. *Publicacions Matemátiques*, pages 301–329, 1996.

[49] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2002.

[50] A. Nagel, E.M. Stein, S. Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985.

[51] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[52] A. Sánchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.*, 78(1):143–160, 1984.

[53] B.-C. Sun, Z.-M. Liu, Q. Li, S.-Z. Zheng. $L^p$-estimates for quasilinear subelliptic equations with VMO coefficients under the controllable growth. *Bound. Value Probl.*, pages 2016:148, 18, 2016.

[54] C.-J. Xu. Subelliptic variational problems. *Bull. Soc. Math. France*, 118(2):147–169, 1990.

[55] C.-J. Xu. Regularity for quasilinear second-order subelliptic equations. *Comm. Pure Appl. Math.*, 45(1):77–96, 1992.

[56] C.-J. Xu. Semilinear subelliptic equations and the Sobolev inequality for vector fields satisfying Hörmander's condition. *Chinese Ann. Math. Ser. A*, 15(2):185–192, 1994.

[57] C.-J. Xu, C. Zuily. Higher interior regularity for quasilinear subelliptic systems. *Calc. Var. Partial Differential Equations*, 5(4):323–343, 1997.