

PENCIL-BEAM APPROXIMATION OF STATIONARY FOKKER-PLANCK

GUILLAUME BAL[†] AND BENJAMIN PALACIOS[‡]

ABSTRACT. Solutions of stationary Fokker-Planck equations in the narrow beam regime are commonly approximated by either ballistic linear transport or by a Fermi pencil-beam equation. We present a rigorous approximation analysis of these three models in a half-space geometry. Error estimates are obtained in a 1-Wasserstein sense, which is an adapted metric to quantify beam spreading. The required well-posedness and regularity results for the stationary Fokker-Planck equation with singular internal and boundary sources are also presented in detail.

1. INTRODUCTION

We consider the stationary (*forward*) Fokker-Planck equation in the half-space $\mathbb{R}_+^n = \{x = (x', x^n) \in \mathbb{R}^n; x^n > 0\}$ in dimension $n \geq 2$ given by:

$$(1.1) \quad -\epsilon^2 \sigma \Delta_\theta u + \theta \cdot \nabla_x u + \lambda u = f, \quad (x, \theta) \in Q := \mathbb{R}_+^n \times \mathbb{S}^{n-1},$$

with ∇_x the spatial gradient and Δ_θ the Laplace-Beltrami operator on the unit sphere \mathbb{S}^{n-1} . The coefficients $\sigma(x)$ and $\lambda(x)$ are spatially-dependent functions bounded above and below by positive constants and $\epsilon > 0$ is a scaling parameter. More concretely,

$$0 < \sigma_0 \leq \sigma(x) \quad \text{and} \quad 0 < \lambda_0 \leq \lambda,$$

for fixed constants σ_0, λ_0 .

The Fokker-Planck equation (1.1) may be used to model at a macroscopic level the propagation of high-frequency waves in heterogeneous media in the regime of highly peaked-forward scattering; see, e.g., [4]. It may also be formally derived from a linear Boltzmann equation, also in the highly peaked forward scattering regime [26, 2]. In (1.1), $f(x, \theta)$, $\epsilon^2 \sigma(x)$ and $\lambda(x)$, represent an internal source/sink of particles at a point x and direction θ , the level of diffusion, and the amount of absorption, respectively.

Our aim is to analyze the regime of narrow beam propagation. Heuristically, a beam propagating with a diffusion coefficient of order $O(1)$ changes directions over comparable distances or times with normalized speed. The narrow beam structure is therefore preserved only over short distances, which after appropriate rescaling takes the form (1.1) with a small diffusion coefficients of order $O(\epsilon^2)$.

The narrow beam will be initiated at the boundary of the domain, which for simplicity we assume to be a half space in most of the paper.

In a general smooth domain $\Omega \subset \mathbb{R}^n$, the boundary source of particles generates the following incoming boundary conditions for (1.1):

$$(1.2) \quad u = g \quad \text{on} \quad \Gamma_- := \{(x, \theta) \in \partial\Omega \times \mathbb{S}^{n-1} : \theta \cdot \nu(x) < 0\},$$

[†]Departments of Statistics and Mathematics, University of Chicago, Chicago, IL.
Email address: guillaumebal@uchicago.edu.

[‡]Department of Statistics, University of Chicago, Chicago, IL.
Email address: bpalacios@uchicago.edu.

where $\nu(x)$ is the outer unit normal vector at x on the boundary. Similarly, the outgoing boundary is defined as

$$\Gamma_+ := \{(x, \theta) \in \partial\Omega \times \mathbb{S}^{n-1} : \theta \cdot \nu(x) > 0\}.$$

Our subsequent analysis will repeatedly use solutions to the *backward Fokker-Planck equation*, which is obtained by changing the sign of the transport operator $T = \theta \cdot \nabla_x$ in (1.1) and imposing boundary conditions on the set Γ_+ instead.

We denote by P the Fokker-Planck operator and by P^t its adjoint. The parameter $\epsilon > 0$ measures the strength of the rescaled diffusion coefficient. We are interested in the asymptotic regime $\epsilon \ll 1$ and $\sigma = O(1)$, where we use the standard notation $O(\epsilon^m)$ whenever a term is bounded from above by a constant times the factor ϵ^m , $m \geq 0$, and write $O(\epsilon^\infty)$ when the latter holds for every $m > 0$. We use the symbol C as a generic positive constant depending on the dimension and coefficients of the problems but independent of ϵ , and which may vary from line to line.

There is a large literature on the analysis of the Fokker-Planck and similar kinetic equations. Existence and uniqueness questions are addressed in, e.g., [13, 10, 28] and [19, 20] from different points of view. The question of regularity of solutions can be traced back to the classical works of hypoelliptic operators [18] and subelliptic estimates [23]. A more specific treatment of regularity in Fokker-Planck, based on sub-elliptic estimates, is presented in [8]. A different approach to regularity can be found for instance in [15, 30], where iterative methods of the form of Moser or De Giorgi techniques are employed to prove Hölder continuity of solutions under sharper conditions on the coefficients. The results obtained from both strategies, subelliptic estimates (to obtain gains in regularity) and iterative methods (to gain continuity conditions), will be used in the derivations of this paper. There has been an increased interest in recent years in the study of the non-local Fokker-Planck models with fractional angular laplacians. Some references are [1, 16, 21]. Most of the analysis presented in this work should extend to the fractional case although we do not consider this issue in detail here. Let us also briefly mention recent work on the reconstruction of coefficients in the forward-peaked regime of Fokker-Planck in [11].

The regime of small diffusions $\epsilon \ll 1$ in (1.1) appears, for instance, in the modeling of laser light propagation in turbulent atmospheres [27, 17, 22, 12]. In the small diffusion scaling, one may be tempted to completely remove the diffusive part, which yields the following ballistic linear transport equation,

$$(1.3) \quad \theta \cdot \nabla_x v + \lambda v = f, \quad (x, \theta) \in \mathbb{R}_+^n \times \mathbb{S}^{n-1},$$

whose solutions can be explicitly written in terms of the source functions.

We will consider the accuracy of this simple model. A more interesting and more accurate asymptotic description is given by the Fermi pencil-beam equation,

$$(1.4) \quad -\tilde{\sigma} \Delta_V U + V \cdot \nabla_X U + \partial_{X^n} U + \tilde{\lambda} U = f, \quad (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1},$$

with Δ_V the Euclidean Laplacian. The above equation may be formally obtained from (1.1) (rescaled so that $\tilde{\sigma}$ is $O(1)$) by replacing the Laplace-Beltrami operator Δ_θ by its approximation on the tangent plane Δ_V and assuming that the beam direction along X^n is constant. Such solutions, unlike ballistic transport, capture beam spreading while completely neglecting backscattering (back to the boundary of \mathbb{R}_+^n). They admit reasonably explicit expressions [14]. For some derivations of the Fermi pencil-beam equation we refer to [6], while [7] formally investigates the accuracy of this model with respect to Fokker-Planck and linear transport.

To analyze the accuracy of the model, it is convenient to link the Fermi pencil-beam and Fokker-Planck equation through a precise choice of local coordinates on the unit sphere and a rescaling of variables at the level of the diffusion, thus depending on ϵ . This is done by means of *stereographic coordinates* (see [24, p.35]) from the south pole $S = (0, \dots, 0, -1)$, given by the map $\mathcal{S} : \mathbb{S}^{n-1} \setminus \{S\} \rightarrow \mathbb{R}^{n-1}$ such that

$$(1.5) \quad \mathcal{S}(\theta) := \left(\frac{\theta_1}{1 + \theta_n}, \dots, \frac{\theta_{n-1}}{1 + \theta_n} \right),$$

and with inverse $\mathcal{J} := \mathcal{S}^{-1} : \mathbb{R}^{n-1} \rightarrow \mathbb{S}^{n-1} \setminus \{S\}$,

$$(1.6) \quad \mathcal{J}(v) = \left(\frac{2v}{\langle v \rangle^2}, \frac{1 - |v|^2}{\langle v \rangle^2} \right), \quad \langle v \rangle := (1 + |v|^2)^{1/2}.$$

Under these coordinates the local representation of the metric on the sphere is

$$\dot{g} = c^2(v)dx^2, \quad \text{with} \quad c(v) := \frac{2}{\langle v \rangle^2},$$

while the volume form $d\theta$, the gradient ∇_θ and the Laplace-Beltrami operator are respectively given by

$$c^{n-1}(v)dv, \quad c^{-2}(v)\nabla_v, \quad \text{and} \quad c^{-(n-1)}(v)\nabla_v \cdot c^{n-3}(v)\nabla_v.$$

The coupling of stereographic coordinates with an ϵ -rescaling of the transversal and angular variables defines the *stretched coordinate system*. Namely, given macroscopic coordinates (x', x^n, θ) in $\mathbb{R}_+^n \times \mathbb{S}^{n-1}$ we denote by

$$(X', X^n, V) \in \mathbb{R}^{n-1} \times \mathbb{R}_+ \times \mathbb{R}^{n-1}$$

the stretched local coordinates defined by the relation

$$(x', x^n, \theta) = (2\epsilon X', X^n, \mathcal{J}(\epsilon V)).$$

This map is a diffeomorphism between $\mathbb{R}_+^n \times (\mathbb{S}^{n-1} \setminus \{S\})$ and $\mathbb{R}_+^n \times \mathbb{R}^{n-1} =: \Omega$.

Given the solution u to (1.1)-(1.2) with sources $f = 0$ and formally $g = \delta(x)\delta(\theta - \eta)$, for $|\eta - N| = O(\epsilon^2)$ (here $N := (0, \dots, 0, 1)$), its *pencil-beam approximation* takes the form

$$(1.7) \quad \mathbf{u}(x, \theta) := (2\epsilon)^{-2(n-1)}U((2\epsilon)^{-1}x', x^n, \epsilon^{-1}\mathcal{S}(\theta)),$$

for U solution to the Fermi pencil-beam equation (1.4) with null interior source, boundary condition $G(X', V) = \delta(X')\delta(V - \epsilon^{-1}\mathcal{S}(\eta))$, and coefficients

$$\tilde{\sigma} = \frac{1}{4}\sigma(X^n\eta) \quad \text{and} \quad \tilde{\lambda} = \lambda(X^n\eta).$$

A similar approximation can be constructed by superposing pencil-beams if instead g is a nonnegative and integrable function on Γ_- . This will be considered later in the paper.

Our analysis provides a rigorous comparison of solutions of Fokker-Planck, ballistic linear transport and the pencil-beam approximation. We establish error estimates in terms of powers of ϵ . Since the solutions model narrow beams and are therefore quite singular, we need a metric that accounts for mass transport at the ϵ -scale. We therefore use a natural setting for comparisons in such instances that is based on a version of the 1-Wasserstein distance.

For a given relative open and bounded set $\Omega \subset \bar{\mathbb{R}}_+^n$, we define

$$BL_{1,\kappa}(\Omega \times \mathbb{S}^{n-1}) := \{\psi \in C(\bar{\Omega} \times \mathbb{S}^{n-1}) : \|\psi\|_\infty \leq 1, \text{Lip}(\psi) \leq \kappa\}.$$

The 1-Wasserstein distance between two finite Borel measures in $\Omega \times \mathbb{S}^{n-1}$, $u(x, \theta)$ and $v(x, \theta)$, is defined as

$$(1.8) \quad \mathcal{W}_{\kappa, \Omega}^1(u, v) := \sup \left\{ \int \psi(x, \theta)(u - v) : \psi \in BL_{1, \kappa}(\Omega \times \mathbb{S}^{n-1}) \right\},$$

and analogously, if u and v are finite Borel measures defined in the whole space $\bar{\mathbb{R}}_+^n \times \mathbb{S}^{n-1}$, we define their 1-Wasserstein distance by replacing Ω with \mathbb{R}_+^n above and use the notation $\mathcal{W}_\kappa^1 = \mathcal{W}_{\kappa, \mathbb{R}_+^n}^1$. From the finiteness of the measures it is clear that

$$\lim_{\Omega \rightarrow \mathbb{R}_+^n} \mathcal{W}_{\kappa, \Omega}^1(u, v) = \mathcal{W}_\kappa^1(u, v).$$

The same distance was used in [3] in the context of inverse transport theory to model possible mis-alignments in the sources used in the probing as well as detector blurring of singular solutions. Consider two points $a \neq b$, two distributions $\delta_a(x) = \delta(x - a)$ and $\delta_b(x) = \delta(x - b)$, an approximation to the identity φ^ϵ , and regularizations in L^1 , $\varphi_a^\epsilon(x) = \varphi^\epsilon(x - a)$ and $\varphi_b^\epsilon(x) = \varphi^\epsilon(x - b)$. For small enough ϵ , $\|\varphi_a^\epsilon - \varphi_b^\epsilon\|_{L^1} = O(1)$ independent of the distance between a and b while $\mathcal{W}_\kappa^1(\delta_a, \varphi_a^\epsilon) = O(\epsilon\kappa)$ and $\mathcal{W}_\kappa^1(\delta_a, \delta_b) = \kappa|a - b|$ when a and b are close enough. Here, κ may be interpreted as a confidence level in our measurements to distinguish between points a and b , as well as the resolution we use to differentiate between measures. For the same reasons, this distance is a natural tool to quantify the spreading of narrowly focused beams.

The approximation by ballistic transport and Fermi pencil beams is interesting only in the setting of narrow beams. The ϵ -broadening of an already ‘‘broad’’ solution is hardly observable. We thus need to consider solutions with measure-valued boundary conditions. This poses no difficulty for the ballistic transport and Fermi pencil beam equations as those are explicitly solved by means of, e.g., Fourier transforms. We develop in this paper the appropriate theory for the Fokker-Planck equation to handle singular boundary conditions in the form of Radon measures. We obtain a unique solution u to (1.1)-(1.2) with internal and boundary sources given by finite compactly supported measures. In the particular cases of $f = 0$ and $g(x, \theta) = \delta_{x_0}(x)\delta_\eta(\theta)$, for $|\eta + \nu(x_0)| = O(\epsilon^2)$, or $g(x, \theta) \in L^1(\partial\mathbb{R}_+^n \times \mathbb{S}^{n-1})$ such that $g \geq 0$ and $\|g\|_{L^1} = 1$, compactly supported around $(x_0, -\nu(x_0))$ inside $\partial\mathbb{R}_+^n \times \{|\theta + \nu(x_0)| < C\epsilon^2\}$, we find a second order approximation $\mathbf{u}(x, \theta)$ by superposing delta-generated pencil-beams of the form (1.7). The corresponding ballistic transport solution of (1.3) is denoted by $v(x, \theta)$. In this context, the main result of this paper reads as follows.

Theorem 1.1. *For dimension $n \geq 2$, there exist constants $C, C' > 0$ (depending on n, σ and λ) such that for any $\kappa \gtrsim 1$,*

$$\mathcal{W}_\kappa^1(u, v) \leq C\kappa\epsilon \quad \text{and} \quad \mathcal{W}_\kappa^1(u, \mathbf{u}) \leq C'\epsilon^2\kappa.$$

In addition, if the resolution parameter is taken such that $\kappa \approx \epsilon^{-1}$, then

$$\frac{1}{C} \leq \mathcal{W}_\kappa^1(u, v) \leq C, \quad \text{and} \quad \mathcal{W}_\kappa^1(u, \mathbf{u}) \leq C'\epsilon.$$

The last statement reflects the reasonable fact that at the ϵ -scale of the beam, the ballistic transport solution v does not accurately model the beam spreading. The Fermi pencil beam approximation, however, is still quite accurate for small ϵ .

The rest of the paper is structured as follows. Section 2 studies the well-posedness of the forward and backward Fokker-Planck equation in the L^2 -setting based on a representation theorem that

goes back to the work [25]. Continuity of solution with respect to the data is given in the L^2 and L^∞ settings. We then consider Hölder and Sobolev regularity, which are obtained by means of a local normal form of (1.1). The former regularity is a direct consequence of recent developments in the theory of ultra-parabolic equations, while the latter follows from a well-known commutator technique introduced in the work on subelliptic estimates [23]. These results are collected in Section 3 to conclude, following a duality argument, on the existence and uniqueness result for (1.1) for rough sources and continuous coefficients. All necessary results on the the fundamental solution of the Fermi pencil-beam equation are presented in Section 4. Section 5 performs the approximation analysis using the 1-Wasserstein distance and states the main approximation result of this work, Theorem 1.1.

2. WELLPOSEDNESS AND REGULARITY OF TIME-INDEPENDENT FOKKER-PLANCK

We first establish existence and uniqueness of weak solutions based on a representation theorem for Hilbert spaces introduced in [25]. Continuity with respect to the boundary and interior sources is also proven in the L^2 -setting for general square integrable functions f and g , while for vanishing boundary sources we also prove continuity with respect to f in the L^∞ -norm. We then move to the question of regularity of solutions for which we introduce a (local) normal form for equation (1.1), and follow two different paths. As a direct application of recent developments on the study of Kolmogorov-type equations and more generally ultra-parabolic equations, we first deduce Hölder continuity of solutions under very weak assumptions on the coefficients. We then obtain an interior regularity result in Sobolev spaces by applying a well-known commutator technique that yields the necessary subelliptic estimates. Before going into the analysis a few remarks need to be taken into consideration regarding the results of this section.

Remark 2.1. (1) Everything we present in this section also holds for the backward Fokker-Planck equation and the proofs are identical but for minor differences such as interchanging the incoming boundary Γ_- with the outgoing one Γ_+ . (2) In all the results presented in this section, the domain $\mathbb{R}_+^n \times \mathbb{S}^{n-1}$ may be replaced by any set of the form $\Omega \times \mathbb{S}^{n-1}$, for Ω a bounded domain with sufficiently regular boundary. (3) The assumption of $\lambda_0 > 0$ is assumed to simplify the analysis and obtain global and integrable solutions on the whole half space. Such an assumption would not be necessary on a bounded domain as long as Poincaré-type estimates are provided, something we do not consider here.

We define $Q := \mathbb{R}_+^n \times \mathbb{S}^{n-1}$ and write $L^k(\Gamma_\pm)$, $k = 1, 2$, to denote the space of L^k functions at the boundary sets Γ_\pm endowed with the measure $|\theta \cdot \nu(x)| dS(x) d\theta$. Here, $d\theta$ stands for the volume form on \mathbb{S}^{n-1} while $dS(x) = dx'$ is the surface measure on $\partial\mathbb{R}_+^n$. We also denote by $L_{loc}^2(\bar{Q})$ the space of functions u so that $\chi u \in L^2(\bar{Q})$ for any $\chi \in C_c^\infty(\mathbb{R}^n)$.

2.1. L^2 -global theory and other properties. Consider the following Hilbert space

$$\mathcal{H} := \{u \in L^2(Q) : \nabla_\theta u \in L^2(Q)\}$$

with norm $\|u\|_{\mathcal{H}}^2 := \|u\|_{L^2}^2 + \epsilon^2 \|\nabla_\theta u\|_{L^2}^2$ and \mathcal{H}' its dual space. We denote by T the transport operator $Tu = \theta \cdot \nabla_x u$. We look for solutions in the sub- Hilbert space

$$\mathcal{Y} := \{u \in \mathcal{H} : Tu \in \mathcal{H}'\} \text{ with norm } \|u\|_{\mathcal{Y}}^2 := \|u\|_{\mathcal{H}}^2 + \|Tu\|_{\mathcal{H}'}^2.$$

Definition 2.2. A function $u \in \mathcal{Y}$ is a weak solution of (1.1)-(1.2) if for all $\varphi \in C^\infty(\bar{Q})$ such that $\varphi = 0$ on Γ_+ , it satisfies

$$(2.1) \quad \begin{aligned} \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} \varphi - u(\theta \cdot \nabla_x \varphi) + \lambda u \varphi) dx d\theta \\ = \int_Q f \varphi dx d\theta + \int_{\Gamma_-} g \varphi |\theta \cdot \nu(x)| dS(x) d\theta. \end{aligned}$$

Theorem 2.3 (Wellposedness). For any $f \in L^2(Q)$ and $g \in L^2(\Gamma_-)$, there exists a unique weak solution $u \in \mathcal{Y}$ to (1.1)-(1.2), which for some constant $C = C(\|\sigma\|_\infty, \sigma_0, \lambda_0) > 0$ satisfies

$$\|u\|_{\mathcal{Y}} \leq C(\|f\|_{L^2} + \|g\|_{L^2}).$$

Furthermore, the traces of u on Γ_- and Γ_+ are well defined, with $u|_{\Gamma_-} = g$ and $u|_{\Gamma_+} \in L^2(\Gamma_+)$.

The main ingredient in the proof of existence is given by the next theorem. The rest of the proof follows the approach of [13] (see also [10]).

Theorem 2.4 (Lions [25]). Let \mathcal{H} be a Hilbert space provided with the inner product $(\cdot, \cdot)_{\mathcal{H}}$ and the norm $\|\cdot\|_{\mathcal{H}}$. Let $\mathcal{F} \subset \mathcal{H}$ be a subspace provided with a prehilbertian norm $|\cdot|$, such that the injection of \mathcal{F} into \mathcal{H} is continuous. Let us consider a bilinear form $a : \mathcal{H} \times \mathcal{F} \rightarrow \mathbb{R}$, such that $a(\cdot, \varphi)$ is continuous on \mathcal{H} for any $\varphi \in \mathcal{F}$ and such that $a(\varphi, \varphi) \geq \alpha |\varphi|^2$ for any $\varphi \in \mathcal{F}$ with $\alpha > 0$. Then, given a linear form $L \in \mathcal{F}'$ continuous with the norm $|\cdot|$, there exists a solution $u \in \mathcal{H}$ of the problem: $a(u, \varphi) = L(\varphi)$ for any $\varphi \in \mathcal{F}$.

Proof of Theorem 2.3.

1) *Existence.* Denote by \mathcal{F} the set of test functions $\varphi \in C^\infty(\bar{Q})$ such that $\varphi = 0$ on Γ_+ , and consider the prehilbertian norm on \mathcal{F} ,

$$|\varphi|_{\mathcal{F}}^2 := \|\varphi\|_{\mathcal{H}}^2 + \frac{1}{2} \|\varphi\|_{L^2(\Gamma_-; |\theta \cdot \nu(x)| dS(x) d\theta)}^2.$$

It is clear that \mathcal{F} is a subspace of \mathcal{H} whose inclusion is continuous. In order to apply Theorem 2.4, we define the bilinear form $a : \mathcal{H} \times \mathcal{F} \rightarrow \mathbb{R}$ as the left-hand side of (2.1):

$$a(u, \varphi) := \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} \varphi - u(\theta \cdot \nabla_x \varphi) + \lambda u \varphi) dx d\theta,$$

and the bounded linear operator $L : \mathcal{F} \rightarrow \mathbb{R}$ as the right-hand side of (2.1):

$$L(\varphi) = \int_Q f \varphi dx d\theta + \int_{\Gamma_-} g \varphi |\theta \cdot \nu(x)| dS(x) d\theta.$$

Under these notation, a weak solution to the FP equation should satisfies the equality

$$(2.2) \quad a(u, \varphi) = L(\varphi), \quad \forall \varphi \in \mathcal{F}.$$

For an arbitrary $\varphi \in \mathcal{F}$, $a(\cdot, \varphi)$ is a bounded linear operator in \mathcal{H} and moreover $a(\cdot, \cdot)$ is coercive when restricted to $\mathcal{F} \times \mathcal{F}$. Indeed, for any $\varphi \in \mathcal{F}$ we have

$$\begin{aligned} a(\varphi, \varphi) &= \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} \varphi \cdot \nabla_{\theta} \varphi + \lambda(x) \varphi^2) dx d\theta + \frac{1}{2} \int_{\Gamma_-} \varphi^2 |\theta \cdot \nu(x)| dS(x) d\theta \\ &\geq \min\{1, \sigma_0, \lambda_0\} |\varphi|_{\mathcal{F}}^2. \end{aligned}$$

Here we used standard integration by parts in \mathbb{R}^n on the integral of $\varphi(\theta \cdot \nabla_x \varphi)$. It follows from Theorem 2.4 that there exists a weak solution $u \in \mathcal{H}$ satisfying (2.2). In fact, $u \in \mathcal{Y}$ because Tu is a distribution given by

$$(2.3) \quad \langle Tu, \varphi \rangle = - \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} \varphi + \lambda u \varphi - f \varphi) dx d\theta, \quad \forall \varphi \in C_c^\infty(Q),$$

thus $|\langle Tu, \varphi \rangle| \leq C \|\varphi\|_{\mathcal{H}}$. By density, the previous holds for all $\varphi \in \mathcal{H}$ therefore (2.3) defines $Tu \in \mathcal{H}'$ for solutions of (1.1). In addition we have:

$$(2.4) \quad \|Tu\|_{\mathcal{H}'} \leq \max\{\|\sigma\|_\infty, \|\lambda\|_\infty\} \|u\|_{\mathcal{H}} + \|f\|_{L^2}.$$

2) *Trace.* To show uniqueness, we first need to make sense of the trace of a solution u on Γ_- (and Γ_+), and we do so by a density argument. The subset $\tilde{\mathcal{Y}} = C_c^\infty(\bar{Q} \setminus \Gamma_0)$ of \mathcal{Y} is known to be dense [5, 10, 13]. Take an arbitrary $\varphi \in \tilde{\mathcal{Y}}$ that vanishes on Γ_+ . Then, Green's identity implies

$$\|\varphi\|_{L^2(\Gamma_-; |\theta \cdot \nu(x)| dS(x) d\theta)}^2 = -2 \int_Q \varphi(T\varphi) dx d\theta \leq 2 \|\varphi\|_{\mathcal{H}} \|T\varphi\|_{\mathcal{H}'} \leq \|\varphi\|_{\tilde{\mathcal{Y}}}^2.$$

Analogously, if $\varphi \in \tilde{\mathcal{Y}}$ vanishes on Γ_- instead, we get

$$\|\varphi\|_{L^2(\Gamma_+; |\theta \cdot \nu(x)| dS(x) d\theta)}^2 \leq 2 \|\varphi\|_{\mathcal{H}} \|T\varphi\|_{\mathcal{H}'} \leq \|\varphi\|_{\tilde{\mathcal{Y}}}^2.$$

Therefore, since any $\varphi \in \tilde{\mathcal{Y}}$ can be decomposed into $\varphi = \varphi_+ + \varphi_-$ with φ_\pm vanishing on Γ_\pm , we deduce

$$\|\varphi\|_{L^2(\Gamma_+ \cup \Gamma_-; |\theta \cdot \nu(x)| dS(x) d\theta)} \leq \|\varphi\|_{\tilde{\mathcal{Y}}}.$$

Consequently, the density of $\tilde{\mathcal{Y}}$ implies $u|_{\Gamma_\pm} \in L^2(\Gamma_\pm; |\theta \cdot \nu(x)| dS(x) d\theta)$ for all $u \in \mathcal{Y}$. In addition, for any pair $\varphi, \psi \in \tilde{\mathcal{Y}}$, we have

$$\int_Q \varphi(T\psi) dx d\theta + \int_Q \psi(T\varphi) dx d\theta = \int_{\Gamma_- \cup \Gamma_+} \varphi \psi (\theta \cdot \nu(x)) dS(x) d\theta.$$

By density we can extend the previous identity and write for any $u_1, u_2 \in \mathcal{Y}$,

$$(2.5) \quad \langle u_1, Tu_2 \rangle_{\mathcal{H}, \mathcal{H}'} + \langle u_2, Tu_1 \rangle_{\mathcal{H}, \mathcal{H}'} = \int_{\Gamma_- \cup \Gamma_+} u_1 u_2 (\theta \cdot \nu(x)) dS(x) d\theta.$$

In particular, if $u_1 = u$ is a weak solution of (1.1) and we take $u_2 = \varphi \in \tilde{\mathcal{Y}}$ such that it vanishes on Γ_+ , then we obtain

$$0 = \int_{\Gamma_-} (u - g) \varphi |\theta \cdot \nu(x)| dS(x) d\theta, \quad \forall \varphi \in \tilde{\mathcal{Y}}.$$

This follows by recalling that u satisfies (2.3) for all $\varphi \in \mathcal{H}$, thus in particular for all $\varphi \in \tilde{\mathcal{Y}}$ vanishing on Γ_+ , and on the other hand, by definition of weak solution, for the same functions φ

$$(2.6) \quad \langle u, T\varphi \rangle_{\mathcal{H}, \mathcal{H}'} = \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} \varphi + \lambda u \varphi - f \varphi) dx d\theta - \int_{\Gamma_-} g \varphi |\theta \cdot \nu(x)| dS(x) d\theta.$$

3) *Uniqueness.* Let $w_1, w_2 \in \mathcal{Y}$ be two weak solutions of (1.1)-(1.2). Therefore, $u = w_1 - w_2$ is a solution of the same equation replacing (f, g) by $(0, 0)$. We first use the definition of Tu in the distributional sense (2.3), which also holds if we take $\varphi = u \in \mathcal{Y} \subset \mathcal{H}$. Therefore,

$$\langle u, Tu \rangle_{\mathcal{H}, \mathcal{H}'} = - \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} u + \lambda u^2) dx d\theta.$$

On the other hand, by plugging $u_1 = u_2 = u$ into (2.5) we obtain

$$2\langle u, Tu \rangle_{\mathcal{H}, \mathcal{H}'} = \int_{\Gamma_+} u^2 |\theta \cdot \nu(x)| dS(x) d\theta.$$

From the previous two equalities, one deduces

$$-\frac{1}{2} \int_{\Gamma_+} u^2 |\theta \cdot \nu(x)| dS(x) d\theta = \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} u + \lambda u^2) dx d\theta \geq \min\{\sigma_0, \lambda_0\} \|u\|_{\mathcal{H}}^2,$$

and consequently, since the left-hand side is nonpositive, the above is true only for $u = 0$.

4) *Continuous dependence.* A weak solution u of (1.1)-(1.2) satisfies $Tu \in \mathcal{H}'$ given by (2.3). Choosing $\varphi = u$, we get

$$\langle u, Tu \rangle_{\mathcal{H}, \mathcal{H}'} = - \int_Q (\epsilon^2 \sigma(x) \nabla_{\theta} u \cdot \nabla_{\theta} u + \lambda u^2 - fu) dx d\theta.$$

This together with (2.5) applied to $u_1 = u_2 = u$ implies

$$(2.7) \quad \min\{\sigma_0, \lambda_0\} \|u\|_{\mathcal{H}}^2 \leq \int_Q f u dx d\theta + \frac{1}{2} \int_{\Gamma_-} g^2 |\theta \cdot \nu(x)| dS(x) d\theta,$$

where the first integral in the right-hand side is bounded as follows

$$\int_Q f u dx d\theta \leq \frac{1}{2} \delta^{-2} \|f\|_{L^2(\Omega \times \mathbb{S}^{n-1})}^2 + \frac{1}{2} \delta^2 \|u\|_{L^2(\Omega \times \mathbb{S}^{n-1})}^2.$$

We then choose $\delta > 0$ small enough in order to absorb the last term with the left-hand side of (2.7). Combining (2.4) with the above inequalities concludes the proof. \square

Lemma 2.5 (Non-negative solutions). *For f, g as in Theorem 2.3, if $f, g \geq 0$ a.e., the unique solution to (1.1)-(1.2) satisfies that $u \geq 0$ a.e. and the same holds for its trace on Γ_- .*

Proof. Let us denote $u^{\pm} = \max\{0, \pm u\}$ and set $A = Q \setminus \text{supp}(u^+)$. We denote the characteristic function of A as $\mathbb{1}_A$. Then, $u^- \in \mathcal{H}$ with $\nabla_{\theta} u^- = (\nabla_{\theta} u) \mathbb{1}_A$. Moreover, $Tu^- \in \mathcal{H}'$ is a distribution given by

$$\langle \phi, Tu^- \rangle = - \int_A (\epsilon^2 \sigma(x) \nabla_{\theta} u^- \cdot \nabla_{\theta} \phi + \lambda(x) u^- \phi + f \phi) dx d\theta, \quad \forall \phi \in C_c^{\infty}(A),$$

whose definition extends to all $\phi \in \mathcal{H}$ supported in A . On the other hand, identity (2.5) for u^- gives

$$2\langle u^-, Tu^- \rangle_{\mathcal{H}, \mathcal{H}'} = \int_{\Gamma_- \cup \Gamma_+} |u^-|^2 (\theta \cdot \nu(x)) dS(x) d\theta.$$

From the previous two equalities and using that $u|_{\Gamma_-} = g \geq 0$, thus $u^-|_{\Gamma_-} = 0$, we deduce that

$$\begin{aligned} & \int_A \lambda |u^-|^2 dx d\theta + \frac{1}{2} \int_{\Gamma_+} |u^-|^2 (\theta \cdot \nu(x)) dS(x) d\theta \\ &= -\epsilon^2 \int_A \sigma(x) |\nabla_{\theta} u^-|^2 dx d\theta - \int_A f u^- dx d\theta, \end{aligned}$$

where by virtue of $f \geq 0$ it leads to

$$\lambda_0 \int_A |u^-|^2 dx d\theta + \frac{1}{2} \int_{\Gamma_+} |u^-|^2 (\theta \cdot \nu(x)) dS(x) d\theta \leq 0$$

In consequence $u^- = 0$. \square

Following [19, Lemma 3.4], we obtain the following continuity in the L^∞ -norm for vanishing boundary source.

Lemma 2.6. *If $f \in L^\infty(Q)$ compactly supported and $g = 0$, then, the solution to (1.1) satisfies the estimate $\|u\|_\infty \leq \lambda_0^{-1} \|f\|_\infty$.*

Proof. Let $M = \|f\|_\infty$. Arguing by contradiction, let's assume there exists $\alpha > 0$ and a bounded set $A \subset Q$ with positive measure such that, with out lost of generality, $u(x, \theta) > M\lambda_0^{-1} + \alpha$ on A . Notice that $u \in L^2(Q)$ since f is also square integrable.

For any sufficiently small $\delta > 0$ we can find a ball $B \subset Q$ such that

$$(2.8) \quad \text{meas}(B \cap A) > (1 - \delta)\text{meas}(B),$$

(see for instance [29]). We take $h \in C_c^\infty(Q)$ so that $h \geq 0$, $\text{supp } h \subset \bar{B}$ and

$$\left| \int \left(\frac{\chi_B}{\text{meas}(B)} - h \right) dx d\theta \right| = \left| 1 - \int h dx d\theta \right| < \delta,$$

and consider φ , solution to the backward Fokker-Planck equation

$$(2.9) \quad -\epsilon^2 \sigma(x) \Delta_\theta \varphi - \theta \cdot \nabla_x \varphi + \lambda \varphi = h, \quad \varphi|_{\Gamma_+} = 0.$$

According to lemma 2.5 applied to (2.9) we deduce $\varphi \geq 0$. Then, we have

$$\int_Q f \varphi dx d\theta \leq M \int_Q \varphi dx d\theta \leq M \lambda_0^{-1} (1 + \delta).$$

The last inequality follows from the fact that φ is indeed a strong solution of (2.9) (as we will see in §2.2.3) and then integrating (2.9) over Q . Indeed,

$$\int_Q h dx d\theta = \int_Q (-\theta \cdot \nabla_x \varphi + \lambda \varphi) dx d\theta$$

where we used Green's identity on the unit sphere: $\int_{\mathbb{S}^{n-1}} \Delta_\theta \varphi d\theta = 0$. Stoke's theorem and the lower bound $\lambda_0 \leq \lambda$ then yield

$$\lambda_0 \int_Q \varphi dx d\theta \leq \int_Q h dx d\theta + \int_{\Gamma_-} \varphi(\theta \cdot \nu(x)) dS(x) d\theta \leq \|h\|_{L^1},$$

where by definition of h , $\|h\|_{L^1} \leq 1 + \delta$.

On the other hand,

$$\int_Q f \varphi dx d\theta = \int_Q u h dx d\theta = I_1 + I_2,$$

with

$$\begin{aligned} I_1 &= \int_{B \cap A} u h dx d\theta \\ &\geq (M\lambda_0^{-1} + \alpha) \left(\int_{B \cap A} \frac{\chi_B}{\text{meas}(B)} dx d\theta - \int_{B \cap A} \left(\frac{\chi_B}{\text{meas}(B)} - h \right) dx d\theta \right) \\ &\geq (M\lambda_0^{-1} + \alpha) \left(\frac{\text{meas}(B \cap A)}{\text{meas}(B)} - \delta \right) > (M\lambda_0^{-1} + \alpha)(1 - 2\delta), \end{aligned}$$

and

$$I_2 = \int_{B \setminus A} u h dx d\theta \leq \|u\|_{L^2} \|h\|_\infty \text{meas}(B \setminus A)^{1/2} \leq \|u\|_{L^2} \|h\|_\infty \text{meas}(B)^{1/2} \delta^{1/2}.$$

Bringing the above together, we obtain that for some $C > 0$ independent of δ ,

$$(M\lambda_0^{-1} + \alpha)(1 - 2\delta) - C\delta^{1/2} \leq M\lambda_0^{-1}(1 + \delta),$$

which then implies that

$$M\lambda_0^{-1} + \alpha \leq M\lambda_0^{-1} + C\delta^{1/2}.$$

It remains to take δ small enough to obtain a contradiction. \square

2.2. Regularity. The regularity of solution to Fokker-Planck is obtained by using an appropriate local representation of the equation. These *beam coordinates* were already used in [6] to derive the Fermi pencil-beam equation from an asymptotic expansions of Fokker-Planck and linear Boltzmann equations.

2.2.1. Beam coordinates. Consider the map $\mathcal{B} : \mathbb{S}_+^{n-1} \rightarrow \mathbb{R}^{n-1}$ transforming the upper hemisphere $\mathbb{S}_+^{n-1} := \{\theta \in \mathbb{R}^{n-1} : |\theta| = 1, \theta_n > 0\}$ and given by

$$\mathcal{B}(\theta) := \left(\frac{\theta_1}{\theta_n}, \dots, \frac{\theta_{n-1}}{\theta_n} \right),$$

and whose inverse is defined as the map $\mathcal{B}^{-1} : \mathbb{R}^{n-1} \rightarrow \mathbb{S}_+^{n-1}$ such that

$$\mathcal{B}^{-1}(v) = \left(\frac{v}{\langle v \rangle}, \frac{1}{\langle v \rangle} \right).$$

Under such coordinates, we obtain the local representations:

$$\begin{aligned} \dot{g} &= \frac{1}{\langle v \rangle^2} \left(Id - \left(\frac{v}{\langle v \rangle} \right) \left(\frac{v}{\langle v \rangle} \right)^T \right), & d\theta &= \langle v \rangle^{-n} dv, \\ \nabla_\theta &= \langle v \rangle^2 (Id + vv^T) \nabla_v, & \Delta_\theta &= \langle v \rangle^n \nabla_v \cdot \langle v \rangle^{2-n} (Id + vv^T) \nabla_v. \end{aligned}$$

2.2.2. Normal form and Hölder regularity. The regularity properties for Fokker-Planck solutions are obtained by assigning to (1.1) a normal form for which regularity issues are well understood. The same analysis applies to backward Fokker-Planck.

Let us pick an arbitrary $\theta_0 \in \mathbb{S}^{n-1}$. By embedding the unit sphere in \mathbb{R}^n , we consider spatial coordinates so that $\theta_0 = (0, \dots, 0, 1) =: N$. On the unit sphere, we choose beam-coordinates mapping the hemisphere $\{\theta \in \mathbb{S}^{n-1} : \theta \cdot \theta_0 > 0\}$ onto \mathbb{R}^{n-1} .

Let $U_x \subset \mathbb{R}_+^n$ be a neighborhood of an arbitrary point $x_0 \in \mathbb{R}_+^n$ and U_v a neighborhood of the origin in \mathbb{R}^{n-1} . For any test function $\tilde{\phi} = \langle v \rangle^{n+1} \phi \in C_c^\infty(U_x \times U_v)$, in local coordinates, equation (1.1) takes the form:

$$(2.10) \quad \int -u \partial_{x^n} \phi - uv \cdot \nabla_{x'} \phi + \left\langle \frac{\epsilon^2 \sigma(x)}{\langle v \rangle^{n-2}} (Id + vv^T) \nabla_v u, \nabla_v (\langle v \rangle^{n+1} \phi) \right\rangle + (\langle v \rangle \lambda) u \phi \, dx dv = \int (\langle v \rangle f) \phi \, dx dv.$$

We re-label coordinates by writing $(y, t) := (x', x^n)$, and define

$$(2.11) \quad A(t, y, v) = \epsilon^2 \sigma(y, t) \langle v \rangle^3 (Id + vv^T), \quad B(t, y, v) = \epsilon^2 (n+1) \sigma(y, t) \langle v \rangle^3 v;$$

and $c(y, v) = \langle v \rangle \lambda(y, t)$, $\hat{f}(t, y, v) = \langle v \rangle f(t, y, v)$. Then, assuming enough regularity on $u(y, t, v)$ one verifies that it satisfies the equation

$$(2.12) \quad \partial_t u + v \cdot \nabla_y u = \nabla_v \cdot (A \nabla_v u) - B \cdot \nabla_v u - cu + \hat{f}, \quad \text{in } U_x \times U_v,$$

in the weak sense, i.e. $u, \nabla_v u, \partial_t u + v \cdot \nabla_y u \in L^2_{loc}(U_x \times U_v)$. We call this equation the *normal form* of (1.1). We point out that the matrix-valued function A is positive semi-definite at every point while the vector-valued function B and the scalar function c are bounded in $U_x \times U_v$.

If a null boundary source is imposed, that is $g = 0$, we can also write equation (1.1) in normal form around (outgoing) boundary point $(x_0, \theta_0) \in \Gamma_+$ by extending the problem across Γ_+ . Indeed, let $(x_0, \theta_0) \in \Gamma_+$ and $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{S}^{n-1}$ a sufficiently small neighborhood of (x_0, θ_0) . Denote by u_{out} the unique solution to the exterior problem

$$(2.13) \quad -\epsilon^2 \sigma \Delta_\theta u_{out} + \theta \cdot \nabla_x u_{out} + \lambda u_{out} = 0 \quad \text{in } \mathbb{R}^n_- \times \mathbb{S}^{n-1},$$

with (exterior) incoming boundary conditions

$$u_{out} = u \quad \text{on } \Gamma_+.$$

Here, σ and λ are any positive extension of the coefficients that preserve the original regularity. It turns out that $u_{ext} = u|_{\mathbb{R}^n_+} + u_{out}|_{\mathbb{R}^n_-}$ satisfies that

$$\int \epsilon^2 \sigma \nabla_\theta u_{ext} \cdot \nabla_\theta \phi - u_{ext} (\theta \cdot \nabla_x \phi) + \lambda u_{ext} \phi \, dx d\theta = \int f_{ext} \phi \, dx d\theta,$$

for all $\phi \in C_c^\infty(\mathcal{U})$, with f_{ext} the zero extension of f . We can choose beam coordinates with respect to (x_0, θ_0) and deduce that u_{ext} satisfies (in a weak sense) an equation of the form (2.12), for all (y, t, v) in a bounded open set.

A first consequence of the normal form (2.12) is the Hölder continuity of solutions. This issue has been studied by several authors in the context of equations of the form (2.12) (see for instance the recent results [15, 30], and the survey [9]). The following is a direct consequence of Theorem 1.1 in [30] (or [Theorem 3 in [15]] in the case $c = 0$).

Theorem 2.7. *Let u be solution to (1.1) for $f, \sigma, \lambda \in L^\infty(\bar{Q})$. Then, there is $\alpha \in (0, 1)$ so that $u \in C^{0, \alpha}(Q)$. If in addition $g = 0$ in (1.1)-(1.2) then $u \in C^{0, \alpha}(Q \cup \Gamma_+)$.*

Remark 2.8. In the context of backward Fokker-Planck with null boundary condition $g = 0$ on Γ_+ , the analogous result gives $u \in C^{0, \alpha}(Q \cup \Gamma_-)$.

2.2.3. Sobolev regularity and strong solutions. Let us introduce the operators $D_x^s := (I - \Delta_x)^{s/2}$, with $s \in \mathbb{R}$, defined by

$$D_x^s u = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ik \cdot x} (1 + |k|^2)^{s/2} \hat{u}(k) dk, \quad \forall u \in C_c^\infty(\mathbb{R}^n),$$

where $\hat{\cdot}$ stands for the Fourier transform. Following the notation $x = (x', x^n)$, we analogously define $D_{x'}$ for an operator acting only on x' .

Theorem 2.9. *Let u be a weak solution to (1.1) with source $f \in L^2(Q)$. Then,*

$$\Delta_\theta u, \theta \cdot \nabla_x u, D_x^{2/3} u \in L^2_{loc}(Q),$$

and in particular u is a strong solution. Moreover, for every compact \mathcal{K} and open \mathcal{O} such that $\mathcal{K} \subset \mathcal{O} \subset Q$, there exists a constant $C > 0$ so that

$$(2.14) \quad \begin{aligned} & \epsilon^2 \|\Delta_\theta u\|_{L^2(\mathcal{K})} + \epsilon^{2/3} \|D_x^{2/3} u\|_{L^2(\mathcal{K})} + \|\theta \cdot \nabla_x u\|_{L^2(\mathcal{K})} \\ & \leq C (\|f\|_{L^2(\mathcal{O})} + \epsilon^2 \|\nabla_\theta u\|_{L^2(Q)} + \|u\|_{L^2(Q)}). \end{aligned}$$

Remark 2.10. Estimates of the form (2.14) are usually referred to as *subelliptic estimates*, in this case with a gain of 2 derivatives in the angular variable and $2/3$ derivatives in the spatial variable. Assuming smoothness in f , σ and λ , repeated differentiation of (1.1) combined with the subelliptic estimates lead to the *hypoellipticity* property.

Remark 2.11. According to §2.2.2, extending the solution near a point $(x, \theta) \in \Gamma_+$ across the boundary implies $\Delta_\theta u$, $\theta \cdot \nabla_x u$, $D_x^{2/3} u \in L^2(\mathcal{K})$ for any compact $\mathcal{K} \subset Q \cup \Gamma_+$.

We give a brief proof of the theorem and we refer the reader to Appendix A for a more detailed one.

Proof. It is enough to obtain the estimates in a neighborhood \mathcal{U} of an interior point $(x_0, \theta_0) \in Q$. Up to some rotation, we can always assume that θ_0 is contained in the span of $(1, 0, \dots, 0)$ and $(0, \dots, 0, 1) = N$. We first consider beam coordinates on the upper hemisphere $\mathbb{S}_+^{n-1} = \{\theta^n > 0\}$, thus $N = (0, \dots, 0, 1)$ is identified with $0 \in \mathbb{R}^{n-1}$.

After standard mollification and localization arguments we reduce the problem to obtaining the desired estimates for a smooth compactly supported function u , solution to the normal form equation

$$(2.15) \quad Tu - \nabla_v \cdot (A \nabla_v u) + B \cdot \nabla_v u + cu = f, \quad \forall (x, v) \in \mathcal{U},$$

with f (different from the original one) a smooth and compactly supported function with L^2 -norm bounded by the norm of the original source and the \mathcal{H} -norm of the Fokker-Planck solution. Here T stands for the transport operator

$$T = (\partial_{x^n} + v \cdot \nabla_{x'}),$$

which satisfies the commutator identity:

$$(2.16) \quad \partial_{x^j} T - T \partial_{v^j}, \quad j = 1, \dots, n-1.$$

The bulk of the proof is analogous to [8] where a central role is played by the (Hörmander-type) identity (2.16). The main difference is the more general form of the principal term, which is easily addressed due to the positive-definiteness of the anisotropic coefficient A . The technique consists in obtaining the estimates

$$(2.17) \quad \|\nabla_v \cdot A \nabla_v u\|_{L^2} \leq C \|f\|_{L^2} \quad \text{and} \quad \|D_{x'}^{2/3} u\|_{L^2} \leq C \epsilon^{-2/3} \|f\|_{L^2},$$

from which one deduces

$$\epsilon^2 \|\Delta_\theta u\|_{L^2} \leq C \|f\|_{L^2}.$$

On the other hand, by considering instead beam coordinates on the hemisphere $\{\theta \in \mathbb{S}^{n-1} : \theta^1 > 0\}$ we can repeat the computations leading to (2.17) but now in terms of the operator $D_{x''}^s = (1 - \Delta_{x''})^{s/2}$, where we decompose the spatial variables as $x = (x^1, x'')$. We then obtain

$$\|D_{x''}^{2/3} u\|_{L^2} \leq C \epsilon^{-2/3} \|f\|_{L^2},$$

which implies an analogous inequality for the full derivative $D_x^{2/3} u$.

The proof is complete by considering a convergent subsequence with respect to the mollification parameter, with respective derivatives converging weakly in L^2 , in a smaller neighborhood \mathcal{V} of (x_0, θ_0) , to $\Delta_\theta u$ and $D_x^{2/3} u$ for u now the Fokker-Planck solution. The estimates for the mollification yield the following subelliptic estimate for the limit,

$$\epsilon^2 \|\Delta_\theta u\|_{L^2(\mathcal{V})} + \|D_x^{2/3} u\|_{L^2(\mathcal{V})} \leq C (\|f\|_{L^2(\mathcal{W})} + \|u\|_{\mathcal{H}}),$$

where here f is the source term in (1.1) and $\mathcal{W} \subset Q$ a slightly larger open set containing \mathcal{U} . Finally, it follows directly from the equation that

$$\|\theta \cdot \nabla_x u\|_{L^2(\mathcal{V})} \leq C(\|f\|_{L^2(\mathcal{W})} + \|u\|_{\mathcal{H}}).$$

□

3. FOKKER-PLANCK WITH SINGULAR SOURCES

Our main motivation to study Fokker-Planck equation is the modeling of narrow beams with a source represented as a delta distribution. The above well-posedness theory does not allow for such objects, and hence it is necessary to extend the equation to delta sources or more general finite compactly supported measures. We achieve this following a duality argument which implies the need of continuity results for solution to the backward Fokker-Planck equation. These results follow from §2 and the paragraph at the beginning of that section.

All the measures considered here will be of Borel type. Let f and g be finite measures in Q and Γ_- respectively. Below we denote by P the forward Fokker-Planck operator while P^t stands for the adjoint (or backward) Fokker-Planck operator.

Definition 3.1. *A distribution $u \in \mathcal{D}'(\bar{Q})$ is said to be a solution of (1.1)-(1.2) whenever*

$$\langle Pu, \varphi \rangle := \langle u, P^t \varphi \rangle = \langle f, \varphi \rangle + \langle g|\theta \cdot \nu(x)|, \varphi \rangle_{\Gamma_-},$$

for all $\varphi \in C_0^\infty(\bar{Q})$ such that $\varphi = 0$ on Γ_+ .

Remark 3.2. (1) The factor $|\theta \cdot \nu(x)|$ is introduced to agree with the standard definition of weak solutions for Sobolev spaces in Definition (2.2). (2) Notice that if u is a finite measure, the above extends to all φ solutions to (3.1) for some $\psi \in C_c(\bar{Q})$, the set of all continuous functions with compact support in \bar{Q} . This is because those solutions belong to $C(\bar{Q}) \cap L^2(Q)$ and can be approximated (with respect to $\|\cdot\|_\infty$) by compactly supported smooth functions.

Theorem 3.3. *Let $\sigma, \lambda \in C(\bar{\mathbb{R}}_+^n)$ and f and g (positive) finite measures with compact support, defined respectively on Q and Γ_- . Then, there is a unique (positive) finite measure in Q satisfying (1.1)-(1.2).*

Proof. Due to the Hölder-regularity result of Theorem 2.7 and particularly the remark after it, and Lemma 2.6, the source-to-solution map associated to the backward equation

$$(3.1) \quad -\epsilon^2 \sigma(x) \Delta_\theta \varphi - \theta \cdot \nabla_x \varphi + \lambda \varphi = \psi, \quad \varphi|_{\Gamma_+} = 0,$$

maps $S_{adj} : C_c(\bar{Q}) \mapsto C(Q \cup \Gamma_-)$ continuously. It is a well defined map due to the inclusion $C_c(\bar{Q}) \subset L^2(Q)$.

We claim that u , defined by

$$(3.2) \quad \langle u, \psi \rangle := \langle f, S_{adj} \psi \rangle_Q + \langle g|\theta \cdot \nu(x)|, S_{adj} \psi \rangle_{\Gamma_-}, \quad \forall \psi \in C_c(\bar{Q}),$$

is a finite measure and satisfies (1.1)-(1.2) in the distributional sense. It is clearly a bounded linear map on $C_c(\bar{Q})$, thus by (Radon-Riesz representation) duality it defines a Radon measure on \bar{Q} , and it is finite since f and g are. Moreover, since any $\varphi \in C_c^\infty(\bar{Q})$ satisfying that $\varphi|_{\Gamma_+} = 0$ can be regarded as the unique solution to the backward system for $\psi = P^t \varphi \in C_c^\infty(\bar{Q})$, then

$$\langle f, \varphi \rangle_Q + \langle g|\theta \cdot \nu(x)|, \varphi \rangle_{\Gamma_-} = \langle u, P^t \varphi \rangle_Q,$$

which means u is a weak solution.

Let v be another finite measure solution to (1.1)-(1.2). For an arbitrary $\psi \in C_c(\bar{Q})$ and denoting $\varphi = S_{adj}\psi$ we have

$$\langle u - v, \psi \rangle = \langle f, \varphi \rangle + \langle g|\theta \cdot \nu(x)|, \varphi \rangle_{\Gamma_-} - \langle v, P^t \varphi \rangle = 0,$$

which implies $u = v$. The positivity is a direct consequence of the definition and the analogous result of Lemma 2.5 for the backward problem. \square

4. FERMI PENCIL-BEAM EQUATION

This section collects all the results we need on the Fermi pencil-beam solutions.

For $\tilde{\sigma}$ and $\tilde{\lambda}$ continuous functions depending only on $X^n \geq 0$ and bounded from below by positive constants, the Fermi pencil-beam equation is given by

$$(4.1) \quad \mathcal{P}(U) := -\tilde{\sigma}\Delta_V U + V \cdot \nabla_{X'} U + \partial_{X^n} U + \tilde{\lambda}U = F, \quad (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1},$$

and is endowed with boundary condition

$$(4.2) \quad U|_{X^n=0} = G(X', V), \quad (X', V) \in \mathbb{R}^{2(n-1)}.$$

The following solvability result goes back to work of Eyges [14] on transport theory of charged particles. Below we write $\mathfrak{F}_{X'}[\cdot]$ to denote the Fourier transformation with respect to the variable X' , that is

$$\mathfrak{F}_{X'}[F](\xi) := \frac{1}{(2\pi)^{(n-1)/2}} \int e^{-iX' \cdot \xi} F(X') dX'.$$

\mathfrak{F}_V is given analogously for a phase variable η . We also write $\tau_{Y'}$ to denote the translation map on the spatial transversal variables, i.e. $[\tau_{Y'} f](X') := f(X' - Y')$.

Proposition 4.1. *For any $F \in C(\mathbb{R}_+; \mathcal{S}'(\mathbb{R}^{2(n-1)}))$ and $G \in \mathcal{S}'(\mathbb{R}^{2(n-1)})$, there exists a unique solution $U \in C^1(\mathbb{R}_+; \mathcal{S}'(\mathbb{R}^{2(n-1)}))$ to (4.1)-(4.2) given explicitly by*

$$(4.3) \quad \begin{aligned} U(X, V) = & e^{-\int_0^{X^n} \tilde{\lambda}(s) ds} \tau_{X^n V}(\mathbf{H}_1 * G) \\ & + \int_0^{X^n} e^{-\int_t^{X^n} \tilde{\lambda}(s) ds} \tau_{X^n V}(\mathbf{H}_2(t) * \tau_{-tV} F) dt, \end{aligned}$$

for $\mathbf{H}_1(X, V)$ and $\mathbf{H}_2(X, V; t)$ Gaussian kernels, namely

$$(4.4) \quad \mathbf{H}_1(X, V) := \frac{e^{-\frac{1}{4(E_2 E_0 - E_1^2)}(E_0 |X'|^2 + 2E_1 X' \cdot V + E_2 |V|^2)}}{(4\pi \sqrt{E_2 E_0 - E_1^2})^{n-1}}$$

with $E_k(X^n) := \int_0^{X^n} s^k \tilde{\sigma}(s) ds$, and \mathbf{H}_2 defined analogously by replacing the terms E_k with $\tilde{E}_k(X^n; t) := \int_t^{X^n} s^k \tilde{\sigma}(s) ds$.

Proof. A solution U to (4.1) should satisfy (formally)

$$(\partial_{X^n} + iV \cdot \xi - \tilde{\sigma}(X^n)\Delta_V + \tilde{\lambda})\mathfrak{F}_{X'}[U] = \mathfrak{F}_{X'}[F].$$

Define

$$\begin{aligned} \mathbf{U}(\xi, X^n, V) &= e^{iX^n(V \cdot \xi) + \int_0^{X^n} \tilde{\lambda}(t) dt} \mathfrak{F}_{X'}[U] \\ \text{and } \mathbf{F}(\xi, X^n, V) &= e^{iX^n(V \cdot \xi) + \int_0^{X^n} \tilde{\lambda}(t) dt} \mathfrak{F}_{X'}[F], \end{aligned}$$

for which the following equality holds

$$\partial_{X^n} \mathbf{U} - \tilde{\sigma}(X^n) e^{iX^n(V \cdot \xi) + \int_0^{X^n} \tilde{\lambda}(t) dt} \Delta_V (e^{-iX^n(V \cdot \xi) - \int_0^{X^n} \tilde{\lambda}(t) dt} \mathbf{U}) = \mathbf{F}.$$

Denoting now $\hat{\mathbf{U}} = \mathfrak{F}_V[\mathbf{U}](\xi, X^n, \eta)$ and $\hat{\mathbf{F}} = \mathfrak{F}_V[\mathbf{F}](\xi, X^n, \eta)$, and recalling the identity

$$\mathfrak{F}_V[e^{iX^n(V \cdot \xi)} h(V)](\eta) = \mathfrak{F}_V[h(V)](\eta - X^n \xi),$$

by applying FT with respect to V in the previous equality, we obtain the following equation

$$\partial_{X^n} \hat{\mathbf{U}} + \tilde{\sigma}(X^n) |\eta - X^n \xi|^2 \hat{\mathbf{U}} = \hat{\mathbf{F}},$$

which we endow with the boundary condition $\hat{\mathbf{U}}(\xi, 0, \eta) = \mathfrak{F}_{X', V}[G](\xi, \eta) =: \mathbf{G}(\xi, \eta)$. $\hat{\mathbf{U}}$ is then given by

$$\hat{\mathbf{U}}(\xi, X^n, \eta) = \mathbf{G}(\xi, \eta) e^{-\int_0^{X^n} |\eta - t\xi|^2 \tilde{\sigma}(t) dt} + \int_0^{X^n} e^{-\int_t^{X^n} |\eta - s\xi|^2 \tilde{\sigma}(s) ds} \hat{\mathbf{F}}(\xi, t, \eta) dt,$$

with $\hat{\mathbf{F}}(\xi, t, \eta) = e^{\int_0^t \tilde{\lambda}(s) ds} \mathfrak{F}_{X', V}[F](\xi, t, \eta - t\xi)$. Consequently, the FT of U has the form

$$\mathfrak{F}_{X', V}[U](\xi, X^n, \eta) = e^{-\int_0^{X^n} \tilde{\lambda}(t) dt} \hat{\mathbf{U}}(\xi, X^n, \eta + X^n \xi).$$

The explicit expression for U follows directly from the previous two equalities. \square

Remark 4.2. A quick application of Plancherel's identity implies the following integration by part formula that will be used later: if U is a pencil-beam with null interior source and boundary condition $G = \delta(X' - Y')\delta(V - W)$, then for any $\Phi \in C(\mathbb{R}_+^n \times \mathbb{R}^{n-1}) \cap H^1(\mathbb{R}_+; L^2(\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}))$,

$$(4.5) \quad \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} U \partial_{X^n} \Phi dX' dX^n dV = - \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} \partial_{X^n} U \Phi dX' dX^n dV - \Phi(Y', 0, W).$$

The approximation analysis of the next section requires estimates on various derivatives of the solutions to Fermi pencil-beam equation. This is not an issue for large values of (X, V) due to the exponential decay at infinity but becomes more tedious when approaching $X^n = 0$, where the pencil-beam solutions are singular. The precise integrals that need to be controlled are summarized in the following lemma.

Lemma 4.3. *Let U be a solution to (4.1)-(4.2) for $F = 0$ and $G = \delta(X')\delta(V - \Theta)$, for some $\Theta \in \mathbb{R}^{n-1}$, and coefficients $\tilde{\sigma} \in C^3(\bar{\mathbb{R}}_+)$, $\tilde{\lambda} \in C(\bar{\mathbb{R}}_+)$. The integrals*

$$\| |X^i - X^n \Theta^i|^l |V^j|^m \partial_X^p \partial_V^q U \|_{L^1}, \quad i, j \in \{1, \dots, n-1\}$$

are finite:

- (i) in general for any integers $l, m \geq 0$ and multi-indices $p, q \geq 0$ so that $3l \geq 3|p| + |q|$;
- (ii) and for any $l, m \geq 0$ and $p, q \geq 0$ so that $3l + m \geq 3|p| + |q|$ if $\Theta = 0$.

Furthermore, the same conclusion holds for a solution \mathbf{U} to (4.1)-(4.2) with $G = 0$ and $F = ((X' - X^n \Theta) \cdot \xi(X^n)) \partial_V^q U$, $|q| \leq 3$ and ξ continuous and bounded in $\bar{\mathbb{R}}_+$ (this will be used in the proof of lemma 5.2).

Proof. Let us write $\eta = \eta(X^n) := e^{-\int_0^{X^n} \tilde{\lambda}(s) ds}$ so that for $F = 0$ and $G = \delta(X')\delta(V - \Theta)$ the Fermi pencil-beam equation can be written as

$$(4.6) \quad U(X, V) = \eta(X^n) \tau_{X^n V} \mathbf{H}_1(X, V - \Theta).$$

For any polynomial $p(X', V)$ and function $f(X', V)$ such that the following quantities are finite, we have

$$\int p(X', V)[\tau_{X^n V} f](X', V) dX' dV = \int p(X' + X^n V, V) f(X', V) dX' dV,$$

and thus, the previous identity along with a change of variables gives

$$(4.7) \quad \begin{aligned} & \| |X^i - X^n \Theta^i|^l |V^j|^m \partial_{X'}^p \partial_V^q U \|_{L^1} \\ & \leq \sum_{\tilde{q} \leq q} \binom{q}{\tilde{q}} \| \eta(X^n) |X^i + X^n V^i|^l |V^j + \Theta^j|^m (X^n)^{|\tilde{q}|} \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} \mathbf{H}_1(X, V) \|_{L^1}. \end{aligned}$$

We write the kernel $\mathbf{H}_1(X, V)$ as

$$(4.8) \quad \mathbf{H}_1(X, V) = \frac{1}{(4\pi\sqrt{\Delta})^{n-1}} e^{-(\alpha|X'|^2 + 2\beta X' \cdot V + \gamma|V|^2)},$$

where $\Delta = ac - b^2 > 0$,

$$(4.9) \quad a(X^n) = \int_0^{X^n} \tilde{\sigma}(s) ds, \quad b(X^n) = \int_0^{X^n} s \tilde{\sigma}(s) ds, \quad c(X^n) = \int_0^{X^n} s^2 \tilde{\sigma}(s) ds,$$

with $\tilde{\sigma} > 0$, and

$$\alpha = \frac{a}{4\Delta}, \quad \beta = \frac{b}{4\Delta} \quad \text{and} \quad \gamma = \frac{c}{4\Delta}.$$

Near $X^n = 0$ and for any extension of $\tilde{\sigma} \in C^3([0, \infty))$ to \mathbb{R} , we deduce from Taylor's theorem that there is $h_0 > 0$ and a function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$(4.10) \quad ac - b^2 = (X^n)^4 (h_0 + h(X^n)),$$

with $h(X^n) \rightarrow 0$ as $X^n \rightarrow 0$. In addition, in the limits $X^n \rightarrow 0$ and $X^n \rightarrow \infty$, the coefficients a , b and c behave respectively as X^n , $(X^n)^2$ and $(X^n)^3$, and this implies

$$\alpha \approx (X^n)^{-3}, \quad \beta \approx (X^n)^{-2}, \quad \text{and} \quad \gamma \approx (X^n)^{-1}.$$

To estimate the terms on the right hand side of (4.7) we need upper bounds for integrals of the form $I_k := \int_{-\infty}^{\infty} |t|^k e^{-\alpha t^2 + 2\beta t} dt$, for $k \geq 0$. We list a few of them:

$$I_k = \begin{cases} \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{\alpha}} & k = 0, \\ \frac{1}{\alpha} (1 + \beta \sqrt{\frac{\pi}{\alpha}} \operatorname{erf}(\frac{\beta}{\sqrt{\alpha}})) e^{\frac{\beta^2}{\alpha}} \leq \frac{1}{\alpha} (1 + |\beta| \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{\alpha}}) & k = 1, \\ \frac{\sqrt{\pi}(2\alpha + 4\beta^2)}{4\alpha^{5/2}} e^{\frac{\beta^2}{\alpha}} & k = 2, \\ \frac{1}{\alpha^2} (1 + \frac{\beta^2}{\alpha} + (\frac{2\beta^2}{\alpha} + 3) \frac{\beta}{2} \sqrt{\frac{\pi}{\alpha}} \operatorname{erf}(\frac{\beta}{\sqrt{\alpha}})) e^{\frac{\beta^2}{\alpha}} \leq \frac{1}{\alpha^2} (1 + \frac{\beta^2}{\alpha} + (\frac{2\beta^2}{\alpha} + 3) \frac{|\beta|}{2} \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{\alpha}}) & k = 3. \end{cases}$$

Performing the integrations over X' and V and using the values of the integrals above one can see that each factor $|X^i|$ scales in the estimates as $(X^n)^{3/2}$ since it bring a division by a factor $\sqrt{\alpha}$, while each $|V^j|$ scales as $(X^n)^{1/2}$ (division by $\sqrt{\gamma}$). Symmetrically, every derivative taken with respect to an X^i bring a factor $\alpha X^i + \beta V^i$ to the estimates which after integration translates into a division by $(X^n)^{3/2}$, and similarly every derivative with respect to V^j brings to the estimates a factor $\beta X^j + \gamma V^j$ which leads to a division by $(X^n)^{1/2}$. The condition imposed on exponent and multi-indices in (4.3) then implies that after integrating with respect to X' and V , integrals of the form

$$\| (X^n)^{|\tilde{q}|} |X^i|^l |V^j|^{\tilde{m}} \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} U \|_{L^1} \quad \text{and} \quad \| (X^n)^{l+|\tilde{q}|} |V^j|^{l+\tilde{m}} \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} U \|_{L^1},$$

with $\tilde{m} \in \{0, m\}$, are bounded by a constant factor times $\int_0^\infty \eta(X^n)(X^n)^s dX^n$, for some non-negative rational number s depending on l, m, p and q . The right hand side is finite since s is nonnegative and the exponential decay at infinity of η .

The two cases in the statement of the lemma derive from the previous estimation by noticing that for $\Theta = 0$ there is no integral with $\tilde{m} = 0$, and we just need $3l + m \geq 3|p| + |q|$ instead of the stronger requirement $3l \geq 3|p| + |q|$.

The second part of the proof consist in estimating the same integrals with U replaced by the more intricate function \mathbf{U} . We write

$$\mathbf{U}(X, V) = \int_0^{X^n} \eta_t(X^n) \tau_{X^n V}(\mathbf{H}_2(X^n; t) * \tau_{-tV} F(t)) dt,$$

for $\eta_t := e^{-\int_t^{X^n} \tilde{\lambda}(s) ds}$ and kernel

$$(4.11) \quad \mathbf{H}_2(X, V; t) = \frac{1}{(4\pi\sqrt{\Delta_t})^{n-1}} e^{-(\alpha_t |X'|^2 + 2\beta_t X' \cdot V + \gamma_t |V|^2)}, \quad t \in (0, X^n).$$

Here we use the notation $\Delta_t = a_t c_t - b_t^2$ for

$$(4.12) \quad a_t(X^n) = \int_t^{X^n} \tilde{\sigma}(s) ds, \quad b_t(X^n) = \int_t^{X^n} s \tilde{\sigma}(s) ds, \quad c_t(X^n) = \int_t^{X^n} s^2 \tilde{\sigma}(s) ds,$$

and $\alpha_t = \frac{a_t}{4\Delta_t}$, $\beta_t = \frac{b_t}{4\Delta_t}$, $\gamma_t = \frac{c_t}{4\Delta_t}$.

Let us analyze $F(t) = (X' - t\Theta) \cdot \xi(t) \partial_V^r U(t)$ (with $|r| \leq 3$) first. We see that

$$\begin{aligned} \tau_{-tV} F(t) &= \tau_{-tV} [(X' - t\Theta) \cdot \xi(t) \partial_V^r (\eta(t) \tau_{tV} \mathbf{H}_1(X', t, V - \Theta))] \\ &= \tau_{-tV} [(X' - t\Theta) \cdot \xi(t) \eta(t) \sum_{\tilde{r} \leq r} \binom{r}{\tilde{r}} t^{|\tilde{r}|} \tau_{tV} [\partial_{X'}^{\tilde{r}} \partial_V^{r-\tilde{r}} \mathbf{H}_1(X', t, V - \Theta)]] \\ &= \sum_{k=1}^{n-1} \sum_{\tilde{r} \leq r} \binom{r}{\tilde{r}} t^{|\tilde{r}|} \eta(t) \xi^k(t) (X^k + t(V^k - \Theta^k)) \partial_{X'}^{\tilde{r}} \partial_V^{r-\tilde{r}} \mathbf{H}_1(X', t, V - \Theta). \end{aligned}$$

Denoting by e_k the multi-index with a 1 in the k th-position and the rest all zeros, and also writing $r = (r_1, \dots, r_{n-1})$, then

$$\begin{aligned} &(X^k + t(V^k - \Theta^k)) \partial_{X'}^{\tilde{r}} \partial_V^{r-\tilde{r}} \mathbf{H}_1(X', t, V - \Theta) \\ &= \partial_{X'}^{\tilde{r}} \partial_V^{r-\tilde{r}} [(X^k + t(V^k - \Theta^k)) \mathbf{H}_1(X', t, V - \Theta)] \\ &\quad - (1 - \delta_{0\tilde{r}_k}) \partial_{X'}^{\tilde{r}-e_k} \partial_V^{r-\tilde{r}} \mathbf{H}_1(X', t, V - \Theta) \\ &\quad - (1 - \delta_{0(r_k - \tilde{r}_k)}) t \partial_{X'}^{\tilde{r}} \partial_V^{r-\tilde{r}-e_k} \mathbf{H}_1(X', t, V - \Theta), \end{aligned}$$

with δ_{ij} the Kronecker delta. Convoluting $\tau_{-tV} F(t)$ with \mathbf{H}_2 gives several terms of the form

$$\mathbf{H}_2 * t^{|\tilde{r}| + |s_2|} \eta(t) \xi^i(t) \partial_{X'}^{\tilde{r}-s_1} \partial_V^{r-\tilde{r}-s_2} [(X^k + t(V^k - \Theta^k))^{1-|(s_1, s_2)|} \mathbf{H}_1(X', t, V - \Theta)]$$

with $|(s_1, s_2)| \leq 1$. Therefore, the estimation of $\| |X^i - X^n \Theta^i|^l |V^j|^m \partial_{X'}^p \partial_V^q \mathbf{U} \|_{L^1}$ reduces to obtaining upper bounds for

$$(4.13) \quad \begin{aligned} &\| |X^i + X^n (V^i - \Theta^i)|^l |V^j|^m (X^n)^{|\tilde{q}|} \int_0^{X^n} t^{|\tilde{r}|} \eta(t) \xi^k(t) \eta_t(X^n) \\ &\quad \times \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} (\mathbf{H}_2 * \partial_{X'}^{\tilde{r}-s_1} \partial_V^{r-\tilde{r}-s_2} [(X^k + t(V^k - \Theta^k))^{1-|(s_1, s_2)|} \mathbf{H}_1(X', t, V - \Theta)]) dt \|_{L^1} \end{aligned}$$

with $i, j, k \in \{1, \dots, n-1\}$, and multi-indices $\tilde{p} \leq p$, $\tilde{q} \leq q$, $\tilde{r} \leq r$, $|(s_1, s_2)| \leq 1$.

Notice that

$$\begin{aligned} & \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} (\mathbf{H}_2 * \partial_{X'}^{\tilde{r}-s_1} \partial_V^{r-\tilde{r}-s_2} [(X^k + t(V^k - \Theta^k))^{1-|(s_1, s_2)|} \mathbf{H}_1(X', t, V - \Theta)]) \\ &= \partial_{X'}^{p+\tilde{q}+\tilde{r}-s_1} \partial_V^{q-\tilde{q}+r-\tilde{r}-s_2} (\mathbf{H}_2 * [(X^k + t(V^k - \Theta^k))^{1-|(s_1, s_2)|} \mathbf{H}_1(X', t, V - \Theta)]). \end{aligned}$$

On the other hand, denoting by $\Sigma = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$, then $\nabla_{X^k, V^k} \mathbf{H}_1 = -2\mathbf{H}_1 \Sigma \begin{pmatrix} X^k \\ V^k \end{pmatrix}$ and consequently

$$(X^k + tV^k) \mathbf{H}_1 = -\frac{1}{2} \left((\Sigma^{-1} \nabla_{X^k, V^k} \mathbf{H}_1)_1 + t(\Sigma^{-1} \nabla_{X^k, V^k} \mathbf{H}_1)_2 \right),$$

where $\Sigma^{-1} = 4 \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$. This yields

$$\begin{aligned} (X^k + t(V^k - \Theta^k)) \mathbf{H}_1(X', t, V - \Theta) &= 2(tb(t) - c(t)) \partial_{X^k} \mathbf{H}_1(X', t, V - \Theta) \\ &\quad - 2(ta(t) - b(t)) \partial_{V^k} \mathbf{H}_1(X', t, V - \Theta). \end{aligned}$$

Therefore,

$$\begin{aligned} & \partial_{X'}^{p+\tilde{q}} \partial_V^{q-\tilde{q}} (\mathbf{H}_2(X^n; t) * \partial_{X'}^{\tilde{r}-s_1} \partial_V^{r-\tilde{r}-s_2} [(X^k + t(V^k - \Theta^k))^{1-|(s_1, s_2)|} \mathbf{H}_1(X', t, V - \Theta)]) \\ &= \begin{cases} \partial_{X'}^{p+\tilde{q}+\tilde{r}-s_1} \partial_V^{q-\tilde{q}+r-\tilde{r}-s_2} (\mathbf{H}_2 * \mathbf{H}_1(X', t, V - \Theta)), & |(s_1, s_2)| = 1, \\ 2(tb(t) - c(t)) \partial_{X'}^{p+\tilde{q}+\tilde{r}+e_k} \partial_V^{q-\tilde{q}+r-\tilde{r}} (\mathbf{H}_2 * \mathbf{H}_1(X', t, V - \Theta)) \\ \quad - 2(ta(t) - b(t)) \partial_{X'}^{p+\tilde{q}+\tilde{r}} \partial_V^{q-\tilde{q}+r-\tilde{r}+e_k} (\mathbf{H}_2 * \mathbf{H}_1(X', t, V - \Theta)), & |(s_1, s_2)| = 0. \end{cases} \end{aligned}$$

The factors $(tb(t) - c(t))$ and $(ta(t) - b(t))$ behave respectively as t^3 and t^2 , in the limit $t \rightarrow 0$.

We now write $\mathbf{H} = \mathbf{H}_2 * \mathbf{H}_1$, therefore

$$(4.14) \quad \mathbf{H}(X, V; t) = \frac{1}{(4\pi\sqrt{\hat{\Delta}})^{n-1}} e^{-\frac{1}{4\hat{\Delta}}(\hat{a}|X'|^2 + 2\hat{b}X' \cdot V + \hat{c}|V|^2)}, \quad t \in (0, X^n),$$

for $\hat{\Delta} = \hat{a}\hat{c} - \hat{b}^2$ and

$$\hat{a}(X^n; t) = a(X^n) + a_t(X^n), \quad \hat{b}(X^n; t) = b(X^n) + b_t(X^n), \quad \hat{c}(X^n; t) = c(X^n) + c_t(X^n).$$

Noting that $\eta(X^n) = \eta_t(X^n)\eta(t)$ and recalling that $\xi(t)$ is bounded, we deduce that the estimation of (4.13) reduces to bounding from above several integrals of the form

$$\begin{aligned} & \int_0^\infty \eta(X^n) (X^n)^{|\tilde{q}|} \\ & \times \int_0^{X^n} t^{|\tilde{r}|} \| |X^i + X^n(V^i - \Theta^i)|^l |V^j|^m P_{X', V} \mathbf{H}(X', X^n, V - \Theta; t) \|_{L^1_{X', V}} dt dX^n, \end{aligned}$$

with $P_{X', V}$ any of the following differential operators:

$$\begin{aligned} & \partial_{X'}^{p+\tilde{q}+\tilde{r}-s_1} \partial_V^{q-\tilde{q}+r-\tilde{r}-s_2}, \quad |(s_1, s_2)| = 1 \\ & 2(tb(t) - c(t)) \partial_{X'}^{p+\tilde{q}+\tilde{r}+e_k} \partial_V^{q-\tilde{q}+r-\tilde{r}} \quad \text{and} \quad 2(ta(t) - b(t)) \partial_{X'}^{p+\tilde{q}+\tilde{r}} \partial_V^{q-\tilde{q}+r-\tilde{r}+e_k}. \end{aligned}$$

These last expressions are bounded by the same arguments as in the first part of the proof. \square

We conclude this section with a few words regarding the adjoint problem associated to (4.1)-(4.2). Following an analogous argument as in the existence result for equation (4.1), if we are given

a bounded and compactly supported function Ψ , the solution to the *backward Fermi pencil-beam* system

$$(4.15) \quad -\tilde{\sigma}\Delta_V W - V \cdot \nabla_{X'} W - \partial_{X^n} W + \tilde{\lambda}W = \Psi, \quad (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1},$$

augmented with a vanishing condition at infinity

$$(4.16) \quad \lim_{X^n \rightarrow \infty} W = 0,$$

can be written in the form

$$W(X, V) = \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \tau_{-X^n V}(\mathbf{H}_3(t) * \tau_{tV} \Psi(t)) dt,$$

for a given Gaussian kernel \mathbf{H}_3 defined similarly as \mathbf{H}_1 in (4.4) but with coefficients E_k replaced by $\int_{X^n}^t s^k \tilde{\sigma}(s) ds$. Formula (4.18) is derived similarly as in the proof of Proposition 4.1 and the X^n details of this can be found in Appendix B.

Using this explicit expression we directly obtain Lipschitz continuity of W . For a proof of this we refer the reader to Appendix C.

Lemma 4.4. *Let $\Psi \in C_c(\bar{Q})$ and Lipschitz continuous with respect to the variables $Z = (X', V)$. There exists $C > 0$ so that, for W solution to (4.15)-(4.16),*

$$|W(Z_1, X^n) - W(Z_2, X^n)| \leq C \text{Lip}(\Psi) |Z_1 - Z_2|,$$

for all $Z_1, Z_2 \in \mathbb{R}^{2(n-1)}$ and $X^n > 0$.

5. 1-WASSERSTEIN COMPARISON ANALYSIS

5.1. Approximation via pencil-beams. We now consider the approximation of narrow beam solutions to the Fokker-Planck equation by ballistic transport and Fermi pencil-beams, which correspond to the limit $\epsilon \ll 1$. The modeling of narrow beams is best modeled as beams propagating in a half space with singular boundary sources. Heuristically, each delta source term gives rise to Fokker-Planck, Fermi pencil-beam, and ballistic transport solutions. More general source terms may then be modeled as superpositions of such delta sources. We first state our comparison results for delta source terms and then consider some models of linear superposition. As we mentioned in the introduction, all comparisons are obtained in the well-adapted notion of 1-Wasserstein distance.

5.1.1. Delta boundary source. We consider first the case of a delta incoming boundary condition

$$(5.1) \quad g(x, \theta) = \delta(x) \delta_{\mathbb{S}^{n-1}}(\theta - \eta), \quad \text{for } \eta \in \mathbb{S}^{n-1} \text{ such that } |N - \eta| = O(\epsilon^2),$$

where $N = -\nu(0) = (0, \dots, 0, 1)$ and we let u be the solution to (1.1)-(1.2) from Theorem 3.3.

Associated to $\eta = (\eta', \eta^n)$, whose stereographic projection with respect to the south pole (see (1.5)) is given by

$$\mathcal{S}(\eta) = \frac{\eta'}{1 + \eta^n} \in \mathbb{R}^{n-1},$$

we set

$$G(X', V) = \delta(X') \delta(V - \epsilon^{-1} \mathcal{S}(\eta)), \quad \tilde{\sigma} = \frac{1}{4} \sigma(X^n \eta) \quad \text{and} \quad \tilde{\lambda} = \lambda(X^n \eta),$$

where the coefficients $\tilde{\sigma}$ and $\tilde{\lambda}$ have an indirect dependence on ϵ through the direction η . Letting U be the solution to (4.1) with the above choice of parameters, we define the *pencil-beam approximation* as

$$\mathbf{u}(x, \theta) := (2\epsilon)^{-2(n-1)} U((2\epsilon)^{-1} x', x^n, \epsilon^{-1} \mathcal{S}(\theta)).$$

Throughout this section, we assume $\sigma \in C^3(\bar{\mathbb{R}}_+^n)$ and $\lambda \in C^2(\bar{\mathbb{R}}_+^n)$. The extra regularity of σ is needed to control the integrals that appear when applying Lemma 4.3.

We then have the following approximation result:

Theorem 5.1. *There exists a positive constant $C = C(\|\sigma\|_{C^3}, \|\lambda\|_{C^2})$ such that, for $\kappa \gtrsim 1$,*

$$\mathcal{W}_\kappa^1(u, \mathbf{u}) \leq C\epsilon^2\kappa.$$

Proof. Consider an open and bounded set $\Omega \subset \bar{\mathbb{R}}_+^n$ and let $\psi \in BL_{1,\kappa}(\Omega \times \mathbb{S}^{n-1})$ which we extend to the entire half-space as a continuous and compactly supported function preserving its Lipschitz constant. Let φ the unique (strong) solution to the backward Fokker-Planck system,

$$(5.2) \quad -\epsilon^2\sigma\Delta_\theta\varphi - \theta \cdot \nabla_x\varphi + \lambda\varphi = \psi, \quad \varphi|_{\Gamma_+} = 0,$$

which is continuous in $Q \cup \Gamma_-$ due to Theorem 2.7 and the remark after it. The goal is to show that

$$\int_{\Omega \times \mathbb{S}^{n-1}} \psi(x, \theta)(u - \mathbf{u}) = O(\epsilon^2\kappa) + o(1),$$

where the error term goes to zero as $\Omega \rightarrow \bar{\mathbb{R}}_+$ since it is bounded by $\int_{\Omega^c \times \mathbb{S}^{n-1}} (u + \mathbf{u}) dx d\theta$. It appears due to the fact that we are extending ψ in space. From the definition of weak solutions, the left hand side of the above expression reads

$$\begin{aligned} \int_Q \psi(x, \theta) d(u - \mathbf{u}) &= \langle g, \varphi | \theta \cdot \nu(x) | \rangle_{\Gamma_-} - \int_Q \mathbf{u}(x, \theta) \psi(x, \theta) dx d\theta \\ &= \varphi(0, \eta) \eta^n - \int_Q \mathbf{u}(x, \theta) \psi(x, \theta) dx d\theta, \end{aligned}$$

and consequently our objective is to show that

$$(5.3) \quad \int_Q \mathbf{u} \psi dx d\theta = \varphi(0, \eta) \eta^n + O(\epsilon^2\kappa).$$

In order to achieve this we rescale the equation and pass to the stretched coordinates (at the level of the diffusion), which we recall is defined by the equality

$$(x', x^n, \theta) = (2\epsilon X', X^n, \mathcal{J}(\epsilon V)),$$

with $\mathcal{J} = \mathcal{S}^{-1}$ the inverse of the stereographic projection from the south pole (defined in (1.5)-(1.6)).

We then set Φ and Ψ as the following rescalings of φ and ψ , respectively, which preserve the L^∞ -norm:

$$(5.4) \quad \varphi(x, \theta) = \Phi((2\epsilon)^{-1}x', x^n, \epsilon^{-1}\mathcal{S}(\theta)) \quad \text{and} \quad \psi(x, \theta) = \Psi((2\epsilon)^{-1}x', x^n, \epsilon^{-1}\mathcal{S}(\theta)).$$

By noticing

$$\begin{aligned} \int_Q \mathbf{u}(x, \theta) \psi(x, \theta) dx d\theta &= \int_Q \mathbf{u}(-\epsilon^2\sigma\Delta_\theta\varphi - \theta \cdot \nabla_x\varphi + \lambda\varphi) dx d\theta \\ &= -\epsilon^2 \int_Q \sigma(\Delta_\theta\mathbf{u})\varphi dx d\theta - \int_Q \mathbf{u}(\theta \cdot \nabla_x\varphi) dx d\theta + \int_Q \lambda\mathbf{u}\varphi dx d\theta, \end{aligned}$$

we pass to stretched coordinates and analyze individually every term on the right hand side. The first one gives

$$\begin{aligned}
& -\epsilon^2 \int_Q \sigma(\Delta_\theta \mathbf{u}) \varphi dx d\theta \\
&= - \int_Q \tilde{\sigma}(X^n) (\Delta_V U) \Phi dX dV \\
&\quad + \int_Q \frac{1}{4} \sigma(2\epsilon X', X^n) \left(1 - 2^{3-n} c^{n-3}(\epsilon V)\right) (\Delta_V U) \Phi dX dV \\
&\quad - \epsilon(n-3) \int_Q c^{-3}(\epsilon V) \sigma(2\epsilon X', X^n) ((\nabla c)(\epsilon V) \cdot \nabla_V U) \Phi 2^{1-n} c^{n-1}(\epsilon V) dX dV \\
&\quad + \frac{1}{4} \int_Q (\sigma(X^n \eta) - \sigma(2\epsilon X', X^n)) (\Delta_V U) \Phi dX dV.
\end{aligned}$$

For the second term we have

$$\int_Q \mathbf{u}(\theta \cdot \nabla_x \varphi) dx d\theta = \int_Q \left[\left(\frac{V \cdot \nabla_{X'} \Phi}{1 + \epsilon^2 |V|^2} \right) U + \left(\frac{1 - \epsilon^2 |V|^2}{1 + \epsilon^2 |V|^2} \right) (\partial_{X^n} \Phi) U \right] \frac{c^{n-1}(\epsilon V)}{2^{n-1}} dX dV.$$

Thus, after integration by parts (where we use (4.5) for the X^n -variable) we obtain

$$\begin{aligned}
& - \int_Q \mathbf{u}(\theta \cdot \nabla_x \varphi) dx d\theta \\
&= \int_Q \left[\left(\frac{V \cdot \nabla_{X'} U}{1 + \epsilon^2 |V|^2} \right) \Phi + \left(\frac{1 - \epsilon^2 |V|^2}{1 + \epsilon^2 |V|^2} \right) (\partial_{X^n} U) \Phi \right] \frac{c^{n-1}(\epsilon V)}{2^{n-1}} dX dV \\
&\quad + \underbrace{\Phi(0, 0, \epsilon^{-1} \mathcal{S}(\eta)) \left(\frac{1 - |\mathcal{S}(\eta)|^2}{1 + |\mathcal{S}(\eta)|^2} \right) \langle \mathcal{S}(\eta) \rangle^{-2(n-1)}}_{= \eta^n (1 + O(\epsilon^2)) \quad (\text{due to (5.1)})} \\
&= \int_Q (V \cdot \nabla_{X'} U) \Phi dX dV + \Phi(0, 0, \epsilon^{-1} \mathcal{S}(\eta)) \eta^n + O(\epsilon^2) \\
&\quad + \int_Q \left(\frac{2^{1-n} c^{n-1}(\epsilon V)}{1 + \epsilon^2 |V|^2} - 1 \right) (V \cdot \nabla_{X'} U) \Phi dX dV \\
&\quad + \int_Q \left(\frac{1 - \epsilon^2 |V|^2}{1 + \epsilon^2 |V|^2} 2^{1-n} c^{n-1}(\epsilon V) - 1 \right) (\partial_{X^n} U) \Phi dX dV.
\end{aligned}$$

Finally, the third term gives

$$\begin{aligned}
\int_Q \lambda \mathbf{u} \varphi dx d\theta &= \int_Q \tilde{\lambda} U \Phi dX dV + \int_Q \lambda(2\epsilon X', X^n) (2^{1-n} c^{n-1}(\epsilon V) - 1) U \Phi dX dV \\
&\quad + \int_Q (\lambda(2\epsilon X', X^n) - \lambda(X^n \eta)) U \Phi dX dV.
\end{aligned}$$

Summarizing all the above and recalling that U is a solution to the Fermi-equation (i.e. $\mathcal{P}(U) = 0$) we have

$$\int_Q \mathbf{u}(x, \theta) \psi(x, \theta) dx d\theta = \varphi(0, \eta) \eta^n + J_1(\Phi) + J_2(\Phi) + O(\epsilon^2),$$

with

$$\begin{aligned}
(5.5) \quad J_1(\Phi) &= \int_{\Omega} \frac{1}{4} \sigma(2\epsilon X', X^n) \left(1 - 2^{3-n} c^{n-3}(\epsilon V)\right) (\Delta_V U) \Phi dX dV \\
&\quad - \epsilon(n-3) \int_{\Omega} c^{-3}(\epsilon V) \sigma(2\epsilon X', X^n) ((\nabla c)(\epsilon V) \cdot \nabla_V U) \Phi 2^{1-n} c^{n-1}(\epsilon V) dX dV \\
&\quad + \int_{\Omega} \left(\frac{2^{1-n} c^{n-1}(\epsilon V)}{1 + \epsilon^2 |V|^2} - 1 \right) (V \cdot \nabla_{X'} U) \Phi dX dV \\
&\quad + \int_{\Omega} \left(\frac{1 - \epsilon^2 |V|^2}{1 + \epsilon^2 |V|^2} 2^{1-n} c^{n-1}(\epsilon V) - 1 \right) (\partial_{X^n} U) \Phi dX dV \\
&\quad + \int_{\Omega} \lambda(2\epsilon X', X^n) (2^{1-n} c^{n-1}(\epsilon V) - 1) U \Phi dX dV \\
&= \sum_{i=1}^5 J_{1,i},
\end{aligned}$$

and

$$\begin{aligned}
(5.6) \quad J_2(\Phi) &= \frac{1}{4} \int_{\Omega} (\sigma(X^n \eta) - \sigma(2\epsilon X', X^n)) (\Delta_V U) \Phi dX dV \\
&\quad - \int_{\Omega} (\lambda(X^n \eta) - \lambda(2\epsilon X', X^n)) U \Phi dX dV.
\end{aligned}$$

To estimate $|J_1|$ we notice that for $m \geq n-3$ ($n \geq 2$),

$$(5.7) \quad 1 - 2^{-m} c^m(\epsilon V) = \begin{cases} \epsilon^2 \sum_{k=1}^m \frac{|V|^2}{(1 + \epsilon^2 |V|^2)^k}, & m > 0 \\ 0, & m = 0 \\ -\epsilon^2 |V|^2, & m = -1, \end{cases}$$

and thus for all $m \geq n-3$,

$$|1 - 2^{-m} c^m(\epsilon V)| \leq |m| \epsilon^2 |V|^2.$$

Furthermore, $(\nabla c)(\epsilon V) = -\frac{4\epsilon V}{(1 + \epsilon^2 |V|^2)^2}$ and

$$\frac{c^{n-4}(\epsilon V)}{2^{n-4} (1 + \epsilon^2 |V|^2)^2} = \frac{1}{(1 + \epsilon^2 |V|^2)^{n-2}} \leq 1.$$

We then have that

$$\begin{aligned}
&|J_{1,1}| + |J_{1,2}| + |J_{1,5}| \\
&\leq C \epsilon^2 \|\Phi\|_{\infty} (|n-3| \| |V|^2 \Delta_V U \|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})} \\
&\quad + \|V \cdot \nabla_V U\|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})} + \|U\|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})}).
\end{aligned}$$

The term $J_{1,3}$ is bounded since

$$\left| \frac{2^{1-n} c^{n-1}(\epsilon V)}{1 + \epsilon^2 |V|^2} - 1 \right| = \epsilon^2 |V|^2 \sum_{k=1}^n \frac{1}{1 + \epsilon^2 |V|^2} \leq n \epsilon^2 |V|^2,$$

[†]The equality follows from the difference of powers formula: $x^{p+1} - y^{p+1} = (x-y)(x^p + x^{p-1}y + \dots + xy^{p-1} + y^p)$.

while the estimate for $J_{1,4}$ follows from

$$\left| \frac{1 - \epsilon^2 |V|^2}{1 + \epsilon^2 |V|^2} 2^{1-n} c^{n-1} (\epsilon V) - 1 \right| \leq \left(1 - \frac{1}{(1 + \epsilon^2 |V|^2)^n} \right) + \frac{\epsilon^2 |V|^2}{(1 + \epsilon^2 |V|^2)^n} \leq (n+1) \epsilon^2 |V|^2.$$

We obtain

$$|J_{1,3}| + |J_{1,4}| \leq C \epsilon^2 \|\Phi\|_\infty (n \| |V|^2 V \cdot \nabla_{X'} U \|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})} + (n+1) \| |V|^2 \partial_{X^n} U \|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})}),$$

where we use the Fermi Pencil-beam equation to get

$$\| |V|^2 \partial_{X^n} U \|_{L^1} \leq \| |V|^2 \Delta_V U \|_{L^1} + \| |V|^2 V \cdot \nabla_{X'} U \|_{L^1} + \| |V|^2 U \|_{L^1}.$$

In summary, and recalling that $\|\Phi\|_\infty = \|\varphi\|_\infty \leq \lambda_0^{-1} \|\psi\|_\infty$ (from lemma (2.6)), we have obtained that

$$|J_1| \leq C \epsilon^2 (\| |V|^2 \Delta_V U \|_{L^1} + \| V \cdot \nabla_V U \|_{L^1} + \| |V|^2 V \cdot \nabla_{X'} U \|_{L^1} + \| |V|^2 U \|_{L^1} + \| U \|_{L^1}) \|\psi\|_\infty.$$

All the factors involving integrals of U are finite due to lemma 4.3.

We now estimate J_2 . When σ and λ are constant, then $J_2 = 0$. Otherwise, we notice first due to our hypothesis (5.1) we can write $\eta = N - 2\epsilon^2 \Theta$ for some $\Theta \in \mathbb{R}^n$, $|\Theta| = O(1)$, thus by Taylor expansion and the above we can write

$$J_2(\Phi) = -\frac{1}{2} \epsilon \int_{\Omega^\epsilon} (X' + \epsilon X^n \Theta', \epsilon X^n \Theta^n) \cdot \nabla \sigma(X^n \eta) (\Delta_V U) \Phi dX dV \\ + 2\epsilon \int_{\Omega^\epsilon} (X' + \epsilon X^n \Theta', \epsilon X^n \Theta^n) \cdot \nabla \lambda(X^n \eta) U \Phi dX dV + \|\Phi\|_\infty O(\epsilon^2)$$

where $\|\Phi\|_\infty \leq \lambda_0^{-1} \|\psi\|_\infty$ and with the remainder depending on integrals of the form

$$\|\sigma\|_{C^2} (\| |X' + \epsilon X^n \Theta'|^2 + \epsilon^2 |X^n \Theta^n|^2 \Delta_V U \|_{L^1} \\ \text{and } \|\lambda\|_{C^2} (\| |X' + \epsilon X^n \Theta'|^2 + \epsilon^2 |X^n \Theta^n|^2 U \|_{L^1})$$

which are uniformly bounded for Θ in a bounded region (guaranteed by (5.1)).

To gain the extra factor ϵ on the leading term, we notice that due to the symmetry of U with respect to the spatial direction $\mathcal{S}(\eta) = \frac{2}{1+\eta^n} \epsilon^2 \Theta'$, for any vector $\xi \in \mathbb{R}^{n-1}$,

$$\int_{\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}} ((X' + \epsilon X^n \Theta') \cdot \xi) U dX' dV = O(\epsilon^2)$$

(this can be obtain for instance by passing to Fourier domain) while

$$\int_{\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}} ((X' + \epsilon X^n \Theta') \cdot \xi) \Delta_V U dX' dV = 0.$$

Therefore we write

$$J_2 = -\frac{1}{4} \epsilon \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} (X' + \epsilon X^n \Theta') \cdot \nabla_{X'} \sigma(X^n \eta) \\ \times (\Delta_V U) (\Phi(X', X^n, V) - \Phi(-\epsilon X^n \Theta', X^n, 0)) dX dV \\ + 2\epsilon \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} (X' + \epsilon X^n \Theta') \cdot \nabla_{X'} \lambda(X^n \eta) \\ \times U (\Phi(X', X^n, V) - \Phi(-\epsilon X^n \Theta', X^n, 0)) dX dV + O(\epsilon^2).$$

Ideally, we would like to say that Φ is a Lipschitz continuous function with constant bounded by the $W^{1,\infty}$ -norm of Ψ , and this would give us the extra ϵ we need for the second order approximation. However, we only have shown Hölder continuity of Φ . A way of bypassing this lack of regularity is with the aid of the following lemma that allows us to substitute Φ with an $O(\epsilon)$ -approximation W that is indeed Lipschitz continuous. Its demonstration can be found at the end of this proof.

Lemma 5.2. *Let Φ , Ψ and U be as above, and consider*

$$f = (X' + \epsilon X^n \Theta') \cdot \xi(X^n)U \quad \text{or} \quad f = (X' + \epsilon X^n \Theta') \cdot \xi(X^n)\Delta_V U,$$

for some bounded function ξ . Then, for W solution to the backward Fermi pencil-beam system (4.15)-(4.16), there exists $C > 0$, depending on σ , λ and ξ , such that

$$\left| \int_{\Omega} f \Phi dX dV - \int_{\Omega} f W dX dV \right| \leq C \epsilon \|\Psi\|_{\infty}.$$

Using this lemma, we replace Φ with W in the last expression of J_2 to get

$$\begin{aligned} J_2 &= -\frac{1}{4}\epsilon \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} (X' + \epsilon X^n \Theta') \cdot \nabla_{X'} \sigma(X^n \eta) \\ &\quad \times (\Delta_V U)(W(X', X^n, V) - W(-\epsilon X^n \Theta', X^n, 0)) dX dV \\ &+ 2\epsilon \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} (X' + \epsilon X^n \Theta') \cdot \nabla_{X'} \lambda(X^n \eta) \\ &\quad \times U(W(X', X^n, V) - W(-\epsilon X^n \Theta', X^n, 0)) dX dV + O(\epsilon^2), \end{aligned}$$

which we rewrite more concisely as

$$J_2 = \epsilon \int_{\Omega} h(X, V)(W(X', X^n, V) - W(-\epsilon X^n \Theta', X^n, 0)) dX dV + O(\epsilon^2).$$

Lemma 4.4 implies

$$|J_2| \leq C \epsilon \| |(X' + \epsilon X^n \Theta', V)| h(X', V) \|_{L^1(\Omega)} \sup_{X^n > 0} \text{Lip}_{X', V}(\Psi(X^n)) + O(\epsilon^2),$$

where the finiteness of the integral involving h follows from lemma 4.3. Moreover, it follows directly from (5.4) that

$$\text{Lip}_{X', V}(\Psi(X^n)) \leq C \epsilon \text{Lip}_{x', v}(\psi(X^n)) \leq C \epsilon \kappa,$$

and thus we deduce the estimate

$$|J_2| \leq C \epsilon^2 (\|\psi\|_{\infty} + \kappa),$$

for a constant C independent of Ω , with the latter present in the estimate only through the support of ψ . The above estimate then depends only on the supremum norm and the Lipschitz constant of ψ , which are both uniformly bounded in $BL_{1,\kappa}$, thus we can take Ω arbitrary large. By taking supremum among all $\psi \in BL_{1,\kappa}(\Omega \times \mathbb{S}^{n-1})$ we deduce that

$$\mathcal{W}_{\kappa, \Omega}^1 \leq C \epsilon^2 \kappa + o(1),$$

and the proof is completed by letting $\Omega \rightarrow \bar{\mathbb{R}}_+^n$. □

Proof of lemma 5.2. Let \mathbf{U} be a solution to the (forward) Fermi pencil-beam problem

$$-\tilde{\sigma} \Delta_V \mathbf{U} + V \cdot \nabla_{X'} \mathbf{U} + \partial_{X^n} \mathbf{U} + \tilde{\lambda} \mathbf{U} = f, \quad (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1}, \quad \mathbf{U}|_{X^n=0} = 0,$$

given explicitly by

$$\mathbf{U}(X, V) = \int_0^{X^n} e^{-\int_t^{X^n} \tilde{\lambda}(s) ds} \tau_{X^n V}(\mathbf{H}_2(t)) * \tau_{-tV} f dt.$$

Denoting

$$\mathbf{u}(x, \theta) = \frac{1}{(2\epsilon)^{2(n-1)}} \mathbf{U}((2\epsilon)^{-1} x', x^n, \epsilon^{-1} \mathcal{S}(\theta)),$$

we have

$$\begin{aligned} \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} f W dX dV &= \int_{\Omega} \mathbf{U} \Psi dX dV \\ &= \underbrace{\int_{\Omega} \mathbf{u}(x, \theta) \psi(x, \theta) dx d\theta}_{I_1} + \underbrace{\int_{\Omega} \mathbf{U} \Psi (2^{1-n} \epsilon^{n-1} (\epsilon V) - 1) dX dV}_{I_2}. \end{aligned}$$

It is clear from (5.7) and lemma 4.3 that

$$|I_2| \leq \epsilon^2 \| |V|^2 \mathbf{U} \|_{L^1} \| \Psi \|_{\infty} \leq C \epsilon^2 \| \Psi \|_{\infty}.$$

On the other hand, we use that φ is a strong solution of (5.2) and we integrate by parts to get

$$I_1 = \int_{\Omega} f \Phi dX dV + J_0(\Phi) + J_1(\Phi) + J_2(\Phi),$$

where

$$(5.8) \quad J_0(\Phi) = \int_{\Gamma_-} \mathbf{u} \varphi(\theta \cdot \nu(x)) dS(x) d\theta = 0,$$

since $U = 0$ on $X^n = 0$, and J_1 and J_2 defined as in (5.5) and (5.6), respectively. Estimating J_1 as in the proof of the previous theorem and using one more time lemma 4.3, we obtain

$$\begin{aligned} |J_1| &\leq C \epsilon^2 (\| |V|^2 \Delta_V \mathbf{U} \|_{L^1} + \| V \cdot \nabla_V \mathbf{U} \|_{L^1} \\ &\quad + \| |V|^2 V \cdot \nabla_{X'} \mathbf{U} \|_{L^1} + \| (1 + |V|^2) \mathbf{U} \|_{L^1} + \| |V|^2 f \|_{L^1}) \| \Psi \|_{\infty} \leq C \epsilon^2 \| \Psi \|_{\infty}, \end{aligned}$$

while $|J_2|$ has a straightforward upper bound given by

$$|J_2| \leq C \epsilon \| \Phi \|_{\infty} (\| X' + \epsilon X^n \Theta' \|_{L^1} + \| X' + \epsilon X^n \Theta' \|_{L^1}) \leq C \epsilon \| \Psi \|_{\infty},$$

for a constant depending on $\| \sigma \|_{C^1}$ and $\| \lambda \|_{C^1}$. \square

5.1.2. L^1 -boundary sources. We now consider more general boundary sources such that for some $C > 0$ independent of ϵ ,

$$(5.9) \quad g(x, \theta) \in L^1(\Gamma_-), \quad \text{supp}(g) \Subset \partial \mathbb{R}_+^n \times \{ \theta \in \mathbb{S}^{n-1} : |\theta - N| < C \epsilon^2 \} \subset \Gamma_-.$$

Let $\{U_{y,\eta}(X, V)\}_{y,\eta}$ be a family of pencil-beams with respective boundary conditions $G = \delta(X') \delta(V - \epsilon^{-1} \mathcal{S}(\eta))$ on $\partial \mathbb{R}_+^n$ and null interior source $F = 0$. The subscript $(y, \eta) \in \text{supp}(g)$ indicates that each pencil beam is constructed using the coefficients

$$\tilde{\sigma}(X^n) = \frac{1}{4} \sigma(y + X^n \eta) \quad \text{and} \quad \tilde{\lambda}(X^n) = \lambda(y + X^n \eta),$$

and they are all given by the explicit formula (4.3).

We define $\mathbf{u}(x, \theta; y, \eta)$ as the following transformation of $U_{y,\eta}$:

$$(5.10) \quad \mathbf{u}(\cdot, \cdot; y, \eta) := (2\epsilon)^{-2(n-1)} U_{y,\eta} \circ T_y^\epsilon, \quad (y, \eta) \in \text{supp}(g),$$

where

$$T_y^\epsilon : \mathbb{R}^n \times \mathbb{S}^{n-1} \ni (x, \theta) \mapsto (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1}$$

is the transformation defined as

$$(X, V) = ((2\epsilon)^{-1}(x' - y'), x^n, \epsilon^{-1}\mathcal{S}(\theta)).$$

In other words, T_y^ϵ defines stretched coordinates after performing a spatial translation $x \mapsto x - y$. The rescaling is chosen such that it preserves (up to a $O(1)$ factor) the L^1 norm of the pencil-beam. Our superposition of pencil-beam model is the distribution in \bar{Q} given by

$$(5.11) \quad \mathbf{u}(x, \theta) := \int_{\Gamma_-} g(y, \eta) \mathbf{u}(x, \theta; y, \eta) dS(y) d\eta,$$

so that

$$\langle \mathbf{u}, \phi \rangle := \int_{\Gamma_-} \int_Q g(y, \eta) \mathbf{u}(x, \theta; y, \eta) \phi(x, \theta) dx d\theta dS(y) d\eta, \quad \forall \phi \in C_0^\infty(\bar{Q}).$$

We can extend the definition of \mathbf{u} to all continuous functions in \bar{Q} since for all $(z, \zeta) \in \text{supp}(g)$,

$$\|\mathbf{u}(\cdot, \cdot; z, \zeta)\|_{L^1} \leq \sup_{(y, \eta) \in \text{supp}(g)} \|U_{y, \eta}\|_{L^1(\mathbb{R}_+^n \times \mathbb{R}^{n-1})} < +\infty.$$

In this setting, we have the following approximation result:

Corollary 5.3. *Let u be a distributional solution to (1.1) with incoming boundary condition g as in (5.9). There exists a constant $C = C(\|\sigma\|_{C^3}, \|\lambda\|_{C^2}) > 0$, independent of ϵ , so that*

$$\mathcal{W}_\kappa^1(u, \mathbf{u}) \leq C \|g\|_{L^1} \epsilon^2 \kappa.$$

Proof. This proof is quite similar to that in the single pencil-beam case. We take an open and bounded $\Omega \subset \mathbb{R}_+^n$. For any $\psi \in BL_{1, \kappa}(\Omega \times \mathbb{S}^{n-1})$, which we extend as a continuous compactly supported function preserving its Lipschitz constant, let φ be the unique solution to the backward Fokker-Planck equation,

$$-\epsilon^2 \sigma \Delta_\theta \varphi - \theta \cdot \nabla_x \varphi + \lambda \varphi = \psi, \quad \varphi|_{\Gamma_+} = 0.$$

We have

$$\begin{aligned} \int_Q \psi(x, \theta) (u - \mathbf{u}) dx d\theta &= \langle g | \theta \cdot \nu(x) |, \varphi \rangle_{\Gamma_-} \\ &\quad - \int_{\Gamma_-} \int_Q g(y, \eta) \mathbf{u}(x, \theta; y, \eta) \psi(x, \theta) dx d\theta dS(y) d\eta \\ &= \int_{\Gamma_-} g(y, \eta) \varphi(y, \eta) |\eta \cdot \nu(y)| dS(y) d\eta \\ &\quad - \int_{\Gamma_-} g(y, \eta) \int_Q \mathbf{u}(x, \theta; y, \eta) \psi(x, \theta) dx d\theta dS(y) d\eta. \end{aligned}$$

The proof reduces to showing that

$$(5.12) \quad \left| \int_Q \mathbf{u}(x, \theta; y, \eta) \psi(x, \theta) dx d\theta - \varphi(y, \eta) \eta^n \right| \leq C \epsilon^2 \kappa,$$

uniformly for $(y, \eta) \in \text{supp}(g)$. Notice that applying the translation $z = x - y$ we get

$$\int_Q \mathbf{u}(x, \theta; y, \eta) \psi(x, \theta) dx d\theta = (\sqrt{2}\epsilon)^{-2(n-1)} \int_Q U_{y, \eta}(\epsilon^{-1}z', z^n, \epsilon^{-1}\mathcal{S}(\theta)) \tilde{\psi}(z, \theta) dz d\theta,$$

with $\tilde{\psi}(z, \theta) = \psi(z + y, \theta)$. Thus, (5.12) is precisely what we obtained in the proof of Theorem 5.1 (see (5.3)). Going over those computations, one realizes that the constant in the estimate are uniform in (y, η) and only depend on $\|\sigma\|_{C^3}$, $\|\lambda\|_{C^2}$ and the dimension n . We then have

$$\int_{\Omega \times \mathbb{S}^{n-1}} \psi(x, \theta)(u - \mathbf{u}) dx d\theta \leq C\epsilon^2 \kappa \|g\|_{L^1} + o(1),$$

for all $\psi \in BL_{1,\kappa}(\Omega \times \mathbb{S}^{n-1})$, for a constant $C > 0$ independent of the set Ω , thus we conclude by taking supremum among all those functions and then letting $\Omega \rightarrow \bar{\mathbb{R}}_+^n$. \square

5.2. Comparison with ballistic linear transport. We now compare the Fokker-Planck solution to the ballistic transport equation in the narrow-beam regime. Narrow beams are by essence well approximated by a distribution supported on a half line, at least when one considers approximations in the 1-Wasserstein distance. We consider such an approximation and show that the Fermi pencil-beam solution is significantly more accurate than the ballistic transport solution. At the appropriate beam scaling, i.e., at distances from the beam center scaled in ϵ , we show (using $\kappa \approx \epsilon^{-1}$) that the Fermi pencil-beam is still accurate while the ballistic transport solution (obviously) fails to account for dispersion.

Since diffusion is weak in a narrow beam regime, simply neglecting its effects leads to the ballistic transport model:

$$(5.13) \quad \theta \cdot \nabla_x v + \lambda v = 0, \quad (x, \theta) \in \mathbb{R}_+^n \times \mathbb{S}^{n-1}, \quad \text{with } v|_{\Gamma_-} = g.$$

Rather than comparing v with u , we compare v with \mathbf{u} instead since we already know how close \mathbf{u} is to u . We obtain the following result.

Lemma 5.4. *For $\sigma \in C^3(\bar{\mathbb{R}}_+^n)$ and $\lambda \in C(\bar{\mathbb{R}}_+^n)$, let g and \mathbf{u} be as in §5.1.2, and let v be the solution to (5.13). There exists a constant $C(\|\sigma\|_{C^3}, \lambda_0) > 0$ independent of ϵ such that*

$$\mathcal{W}_\kappa^1(v, \mathbf{u}) \leq C\kappa\epsilon \|g\|_{L^1}.$$

Proof. The transport solution v can be written as

$$v(x, \theta) = e^{-\int_0^{\tau_-(x, \theta)} \lambda(x - s\theta) ds} g(x - \tau_-(x, \theta)\theta, \theta).$$

Thus, applying the change of variables

$$(x, \theta) = \mathcal{J}(y', t, \eta) := ((y' + t\eta', t\eta^n), \eta), \quad |\nabla \mathcal{J}(y', t, \eta)| = |\eta^n|$$

we get that for any $\psi \in BL_{1,\kappa}(\bar{\mathbb{R}}_+^n \times \mathbb{S}^{n-1})$,

$$\begin{aligned} \int v \psi dx d\theta &= \int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}^{n-1}} \int_0^\infty g((y', 0), \eta) e^{-\int_0^t \lambda((y', 0) + s\eta) ds} \psi((y', 0) + t\eta, \eta) \eta^n dt dy' d\eta \\ &= \int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}^{n-1}} \int_0^\infty g((y', 0), \eta) e^{-\int_0^t \lambda((y', 0) + s\eta) ds} \psi((y', 0) + t\eta, \eta) dt dy' d\eta \\ &\quad + O(\epsilon^2 \|g\|_{L^1} \|\psi\|_\infty), \end{aligned}$$

where the last equality is obtained from the fact that for all $x \in \partial\mathbb{R}_+^n$, $g(x, \cdot)$ is supported in a ϵ^2 -neighborhood of N . On the other hand, recalling the definition of \mathbf{u} in (5.10)-(5.11) and denoting $\Psi(X, V) = \psi(2\epsilon X', X^n, \mathcal{J}(\epsilon V))$ with \mathcal{J} the inverse of the stereographic projection with respect to

the south pole, we have that

$$\begin{aligned}
& \int_Q \mathbf{u} \psi dx d\theta \\
&= \int_{\Gamma_- \times \mathbb{R}^{n-1} \times \mathbb{R}_+^{n-1}} g((y', 0), \eta) U_{y, \eta}(X' - \frac{1}{2\epsilon} y, X^n, V) \Psi(X, V) \frac{c^{n-1}(\epsilon V)}{2^{n-1}} dX dV dy' d\eta \\
&= \int_{\Gamma_- \times \mathbb{R}^{n-1} \times \mathbb{R}_+^{n-1}} g((y', 0), \eta) U_{y, \eta}(X' - \frac{1}{2\epsilon} y', X^n, V) \Psi(X, V) dX dV dy' d\eta + R(\epsilon^2),
\end{aligned}$$

with remainder $|R(\epsilon^2)| \leq C\epsilon^2 \|\psi\|_\infty \|g\|_{L^1}$. Combining both integrals above and performing the dilation $Y' = \frac{1}{2\epsilon} y'$ we obtain

$$\begin{aligned}
& \int_Q \psi(x, \theta) (v - \mathbf{u}) dx d\theta \\
&= \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) e^{-\int_0^{X^n} \lambda((2\epsilon Y', 0) + s\mathcal{J}(\epsilon W)) ds} \\
&\quad \times \Psi(Y', X^n, W) dX^n dY' dW \\
(5.14) \quad & - \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) \left(\int_{\mathbb{R}^{2(n-1)}} U_{(2\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) \right. \\
&\quad \left. \times \Psi(X' + Y, X^n, V) dX' dV \right) dX^n dY' dW \\
& + R(\epsilon^2),
\end{aligned}$$

where $\tilde{g}(Y', W) := (2\epsilon)^{2(n-1)} g((2\epsilon Y', 0), \mathcal{J}(\epsilon W))$, thus its L^1 -norm is of the order of $\|g\|_{L^1}$. We recall that $U_{y, \eta}$ is the solution to the Fermi pencil-beam equation for $\tilde{\sigma}(X^n) = \frac{1}{4}\sigma(y + X^n\eta)$ and $\tilde{\lambda}(X^n) = \lambda(y + X^n\eta)$, and moreover

$$\int_{\mathbb{R}^{2(n-1)}} U_{y, \eta}(X', X^n, V) dX' dV = e^{-\int_0^{X^n} \lambda(y + s\eta) ds}.$$

Consequently,

$$\begin{aligned}
& \int_Q \psi(x, \theta) (v - \mathbf{u}) dx d\theta \\
&\leq \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) \left| \int_{\mathbb{R}^{2(n-1)}} U_{(2\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) \Psi(Y', X^n, W) dX' dV \right. \\
&\quad \left. - \int_{\mathbb{R}^{2(n-1)}} U_{(2\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) \Psi(X' + Y', X^n, V) dX' dV \right| dX^n dY' dW \\
&\leq \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) \int_{\mathbb{R}^{2(n-1)}} U_{(2\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) \\
&\quad \times |\Psi(Y', X^n, W) - \Psi(X' + Y', X^n, V)| dX' dV dX^n dY' dW \\
&\leq C\epsilon \text{Lip}(\psi) \left(\sup_{(y, \eta) \in \text{supp}(g)} \|U_{y, \eta}(X, V)|(X', V)\|_{L^1(\Omega)} \right) \|\tilde{g}\|_{L^1}.
\end{aligned}$$

We conclude by noticing that $\|\tilde{g}\|_{L^1} = \|g\|_{L^1}$ and

$$\sup_{(y, \eta) \in \text{supp}(g)} \|U_{y, \eta}(X, V)|(X', V)\|_{L^1(\Omega)} < \infty,$$

and taking supremum over all $\psi \in BL_{1,\kappa}(\bar{\mathbb{R}}_+^n \times \mathbb{S}^{n-1})$. \square

We now obtain a lower bound on the mismatch between ballistic transport and the pencil-beam model. This shows that spreading is indeed not accounted for by the ballistic transport solution, and this effect becomes visible once we set a high measuring resolution of order ϵ^{-1} or higher.

Lemma 5.5. *For a resolution parameter $\kappa \gtrsim \epsilon^{-1}$ and $\sigma, \lambda \in C(\bar{\mathbb{R}}_+^n)$, there exists a constant $C(n, \sigma_0, \|\lambda\|_\infty) > 0$ such that*

$$\frac{1}{C} \|g\|_{L^1} \leq \mathcal{W}_\kappa^1(v, \mathbf{u}).$$

Proof. For any $\psi \in BL_{1,\kappa}(\bar{\mathbb{R}}_+^n \times \mathbb{S}^{n-1})$, we have equality (5.14) for the rescaling $\Psi(X, V) = \psi(2\epsilon X', X^n, \mathcal{J}(\epsilon V))$. We take an specific $\psi(x, \theta) = \psi_1(x^n)\psi_2(\theta)$ satisfying that:

1. $0 \leq \psi_1, \psi_2 \leq 1$ and $\|\psi\|_\infty = 1$;
2. $\psi_2(\theta) = 1$ in the support of g (i.e. for all $|N - \theta| \leq C\epsilon^2$);
3. denoting $\Psi_2(V) = \psi_2(\mathcal{J}(\epsilon V))$, $\text{supp } \Psi_2 \subset B(\eta)$ with $\eta > 0$ to be chosen;
4. $\text{supp } \psi_1 \subset (a, b) \Subset (0, \infty)$.

For this choice of test function (and recalling $\tilde{g}(Y', W) = (2\epsilon)^{2(n-1)}g(2\epsilon Y', 0, \mathcal{J}(\epsilon W))$) we have

$$\begin{aligned} & \int \psi(x, \theta)(v - \mathbf{u}) dx d\theta \\ & \geq \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) e^{-\int_0^{X^n} \lambda((\epsilon Y', 0) + s(\epsilon W, 1)) ds} \\ & \quad \times \Psi_1(X^n) \Psi_2(W) dX^n dY' dW \\ & \quad - \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) \Psi_1(X^n) \\ & \quad \times \left(\int_{\mathbb{R}^{2(n-1)}} U_{(\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) \Psi_2(V) dX' dV \right) dX^n dY' dW \\ & \quad - |R(\epsilon^2)|. \end{aligned}$$

We rewrite this inequality as

$$(5.15) \quad \int \psi(x, \theta)(v - \mathbf{u}) dx d\theta \geq \int_{\mathbb{R}^{2(n-1)}} \int_0^\infty \tilde{g}(Y', W) e^{-\int_0^{X^n} \lambda((\epsilon Y', 0) + s(\epsilon W, 1)) ds} \\ \times I_\eta(X^n; Y', W) \Psi_1(X^n) dX^n dY' dW - |R(\epsilon^2)|,$$

where $I_\eta(X^n; Y', W)$ is defined as

$$\begin{aligned} I_\eta(X^n; Y', W) \\ = 1 - \int_{\mathbb{R}^{n-1} \times B(\eta)} e^{\int_0^{X^n} \lambda((\epsilon Y', 0) + s(\epsilon W, 1)) ds} U_{(\epsilon Y', 0), \mathcal{J}(\epsilon W)}(X', X^n, V) dX' dV. \end{aligned}$$

We see that

$$\begin{aligned} I_\eta(X^n; Y', W) &= 1 - \int_{B(\eta)} \frac{e^{-\frac{|V|^2}{4E_0(X^n; Y', W)}}}{(4\pi E_0(X^n; Y', W))^{\frac{n-1}{2}}} dV \\ &= \int_{|V| > \eta} \frac{e^{-\frac{|V|^2}{4E_0(X^n; Y', W)}}}{(4\pi E_0(X^n; Y', W))^{\frac{n-1}{2}}} dV, \end{aligned}$$

with $E_0(X^n; Y', W) = 4 \int_0^{X^n} \sigma((\epsilon Y', 0) + s(\epsilon W, 1)) dt$. By considering spherical coordinates and noticing that

$$\frac{\eta}{\sqrt{4E_0(X^n; Y', W)}} \leq \frac{\eta}{4\sqrt{a\sigma_0}} =: \eta_0, \quad \forall (Y', W) \in \text{supp}(\tilde{g}), \quad X^n \in (a, b),$$

we have

$$I_\eta(X^n; Y', W) \geq c_n \int_{\eta_0}^{\infty} e^{-r^2} r^{n-2} dr,$$

for $c_n = \frac{|\mathbb{S}^{n-2}|}{\pi^{(n-1)/2}} = 2\Gamma((n-1)/2)^{-1}$. For $n = 2$, we can choose $R > \eta^0$ so that

$$(5.16) \quad \int_{\eta_0}^{\infty} e^{-r^2} dr \geq \frac{1}{R} \int_{\eta_0}^R e^{-r^2} r dr = \frac{1}{2R} (e^{-\eta_0^2} - e^{-R^2}).$$

If now $n \geq 3$, it is not hard to verify that

$$(5.17) \quad \int_{\eta_0}^{\infty} e^{-r^2} r^{n-2} dr \geq e^{-\eta_0^2} \sum_{k=0}^{n-3} \frac{k!}{2^k} \eta_0^{n-3-k}.$$

The parameter η imposes a restriction on the size of the support of the (angular) test function $\Psi_2(V)$, which says that $\text{supp} \Psi_2 \subset B(\eta)$. In order to fulfill this support condition we want our test function not to decay too slowly, which means we need to have a function Ψ as above and such that

$$\|\Psi_2\|_{\infty} = 1 \quad \text{and} \quad \text{Lip}_V(\Psi_2) \geq \eta^{-1}.$$

Going back to the original angular variable θ this translates into

$$C \text{Lip}_{\theta}(\psi_2) \geq (\epsilon \eta)^{-1},$$

for some constant $C > 0$ (independent of the parameters), hence, this condition is fulfilled by choosing η such that

$$\kappa \gtrsim \epsilon^{-1} \eta^{-1}.$$

In dimension 2, we choose $\eta = O(\epsilon^{-1} \kappa^{-1})$ so that $\eta_0 = \epsilon^{-1} \kappa^{-1}$, and $R = \sqrt{2} \eta_0$, thus we obtain

$$I_\eta(X^n; Y', W) \geq c \epsilon \kappa e^{-(\epsilon \kappa)^{-2}} (1 - e^{-(\epsilon \kappa)^{-2}}),$$

for some constant c independent of ϵ and uniform with respect to X^n, Y' and W . According to our assumption $\kappa \gtrsim \epsilon^{-1}$ we deduce that for some $c > 0$,

$$I_\eta(X^n; Y', W) \geq c \epsilon \kappa.$$

Similarly in dimension $n \geq 3$, for η as above we get the next uniform lower bound

$$I_\eta(X^n; Y', W) \geq e^{-(\epsilon \kappa)^{-2}} \sum_{k=0}^{n-3} \frac{k!}{2^k} (\epsilon \kappa)^{-(n-3-k)} \geq c,$$

for another constant $c > 0$. Introducing this lower bounds into (5.15) yields that for $\kappa \gtrsim \epsilon^{-1}$

$$\|\lambda\|_{\infty}^{-1} \|g\|_{L^1} (1 - \epsilon^2) \leq C \int \psi(x, \theta) (v - \mathbf{u}) dx d\theta \leq C \mathcal{W}_{\kappa}^1(v, \mathbf{u}).$$

□

We summarize the previous results of this section in the form of our main theorem which we rephrase here.

THEOREM 1.1. *Let u be the solution to Fokker-Planck with boundary source g as in (5.9) and $\|g\|_{L^1} = 1$. Let \mathbf{u} the superposition of pencil beam in (5.11) and v the ballistic transport solution of (5.13). For any dimension $n \geq 2$ there exists a constant $C(n, \|\sigma\|_{C^3}, \|\lambda\|_{C^2}) > 0$ such that for $\kappa \gtrsim 1$,*

$$\mathcal{W}_\kappa^1(u, v) \leq C\kappa\epsilon \quad \text{and} \quad \mathcal{W}_\kappa^1(u, \mathbf{u}) \leq C\epsilon^2\kappa.$$

Moreover, if the resolution parameter is such that $\kappa \approx \epsilon^{-1}$, then there is $C > 0$ so that

$$C^{-1} \leq \mathcal{W}_\kappa^1(u, v) \leq C \quad \text{and} \quad \mathcal{W}_\kappa^1(u, \mathbf{u}) = O(\epsilon).$$

Proof. It follows directly from corollary 5.3 and lemmas 5.4 and 5.5. □

ACKNOWLEDGEMENT

The authors would like to thank Luis Silvestre for multiple comments on the manuscript and for bringing [21] to their attention. This research was partially supported by the Office of Naval Research, Grant N00014-17-1-2096 and by the National Science Foundation, Grant DMS-1908736.

APPENDIX A. COMPLETE PROOF OF THEOREM 2.9

We follow the method presented in [8] which is based on a Hörmander-type identity for the commutator of certain operators.

1) *Localization and Mollification.* Let η^δ be a mollifier and χ a cutoff function supported in a neighborhood $\mathcal{U} \Subset Q$ of a point (x_0, θ_0) . Let \mathcal{V} and \mathcal{W} be two open sets containing (x_0, θ_0) and such that $\chi = 1$ in \mathcal{V} and $\mathcal{V} \Subset \mathcal{U} \Subset \mathcal{W}$. We choose $\delta > 0$ small enough so that $\mathcal{U} + \text{supp}(\eta^\delta) \subset \mathcal{W}$. Since we can always rotate the coordinate system in \mathbb{R}^n , we lose no generality in assuming that θ_0 is contained in the span of $(1, 0, \dots, 0)$ and $(0, \dots, 0, 1) = N$. We first consider beams coordinates on \mathbb{S}^{n-1} with respect to the north pole N . Then N is identified with $0 \in \mathbb{R}^{n-1}$ and we write $\eta^\delta(x, v)$ and $\chi(x, v)$ for the respective representative function of η^δ and χ in the local coordinates. We abuse the notation and write \mathcal{U} interchangeably for the set considered above and its image under local coordinates.

Let's plug

$$\phi(y, \xi) = \eta^\delta(x - y, \theta - \xi)\chi(x, \theta)$$

into equation (2.10). From the transport part we get

$$\begin{aligned} & -\chi(x, v) \int u(y, w)(\partial_{y^n} + w \cdot \nabla_{y'})[\eta^\delta(x - y, v - w)]dydw \\ &= \chi(x, v) \int u(y, w)(\partial_{x^n}\eta^\delta(x - y, v - w) + v \cdot \nabla_{x'}\eta^\delta(x - y, v - w))dydw \\ &\quad - \chi(x, v) \int u(y, w)((v - w) \cdot \nabla_{x'}\eta^\delta(x - y, v - w))dydw \\ &= (\partial_{x^n} + v \cdot \nabla_{x'})[\chi(u * \eta^\delta)] - \underbrace{(u * \eta^\delta)(\partial_{x^n}\chi + v \cdot \nabla_{x'}\chi) + \chi(u * (v \cdot \nabla_{x'}\eta^\delta))}_{=: g_1}. \end{aligned}$$

Using the notation in (2.11) the diffusion term gives

$$\begin{aligned}
& \chi(x, v) \int \left\langle \frac{\epsilon^2 \sigma(y)}{\langle w \rangle^{n-2}} (Id + ww^T) \nabla_w u, \nabla_w (\langle w \rangle^{n+1} \eta^\delta(x-y, v-w)) \right\rangle dy dw \\
&= -\chi(x, v) \int \nabla_w u \cdot (A(y, w) \nabla_v \eta^\delta(x-y, v-w)) + u(B(x, w) \cdot \nabla_v \eta^\delta(x-y, v-w)) dy dw \\
&= -\nabla_v \cdot (A \nabla_v (\chi(u * \eta^\delta))) + (u * \eta^\delta) (\nabla_v \cdot A \nabla_v \chi) \\
&\quad + B \cdot \nabla_v (\chi(u * \eta^\delta)) - (u * \eta^\delta) (B \cdot \nabla_v \chi) \\
&\quad - \chi(x, v) \int \nabla_w u \cdot H_A(y, w) [(x-y, v-w), \nabla_v \eta^\delta(x-y, v-w)] dy dw \\
&\quad - \chi(x, v) \int u H_B(y, w) [(x-y, v-w), \nabla_v \eta^\delta(x-y, v-w)] dy dw \\
&= -\nabla_v \cdot (A \nabla_v (\chi(u * \eta^\delta))) + B \cdot \nabla_v (\chi(u * \eta^\delta)) - g_2,
\end{aligned}$$

for some multi-linear operators H_A and H_B arising from the Taylor expansion of A and B , respectively, around the point (x, v) . Finally, the zero order term in (2.10) gives

$$\begin{aligned}
& \chi(x, v) \int c(y, w) u(y, w) \eta^\delta(x-y, v-w) dy dw \\
&= c(\chi(u * \eta^\delta)) - \underbrace{\chi(x, v) \int u(y, w) ((x-y, v-x) \cdot h_c(y, w)) \eta^\delta(x-y, v-w) dy dx}_{=: g_3}
\end{aligned}$$

with h_c from the remainder of the first order Taylor expansion of $c(x, v)$. From the previous we deduce that the function

$$u^\delta := (u * \eta^\delta) \chi \in C_c^\infty(\mathcal{U}),$$

satisfies

$$(A.1) \quad -\nabla_v \cdot (A \nabla_v u^\delta) + \partial_{x^n} u^\delta + v \cdot \nabla_{x'} u^\delta + B \cdot \nabla_v u^\delta + c u^\delta = f^\delta, \quad \forall (x, v) \in \mathcal{U},$$

with

$$f^\delta := (\hat{f} * \eta^\delta) \chi + g_1 + g_2 + g_3 \in C_c^\infty(\mathcal{U}).$$

One verifies that

$$\|g_1 + g_2 + g_3\|_{L^2(\mathcal{U})} \leq C(\|u\|_{L^2(\mathcal{U})} + \epsilon^2 \|\nabla_v u\|_{L^2(\mathcal{U})}),$$

for some constant depending on $\|\sigma\|_{C^1}$, $\|\lambda\|_{C^1}$, $\|\eta^\delta\|_{L^1}$ and $\| |(x, \theta)| \nabla_\theta \eta^\delta \|_{L^1}$.

Since $u, \nabla_v u \in L^2(\mathcal{V})$ and $\chi = 1$ in \mathcal{V} , then $u^\delta \rightarrow u$ and $\nabla_v u^\delta \rightarrow \nabla_v u$, as $\delta \rightarrow 0$, in the L^2 -sense. Furthermore we have the following estimate:

$$(A.2) \quad \begin{aligned} \|f^\delta\|_{L^2(\mathcal{U})} &\leq \|f\|_{L^2(\mathcal{W})} + C(\epsilon^2 \|\nabla_\theta u\|_{L^2(\mathcal{W})} + \|u\|_{L^2(\mathcal{W})}) \\ &\leq \|f\|_{L^2(\mathcal{W})} + C(\epsilon^2 \|\nabla_\theta u\|_{L^2(\mathcal{Q})} + \|u\|_{L^2(\mathcal{Q})}), \end{aligned}$$

uniformly with respect to $\delta \ll 1$, and where the constant $C > 0$ depends on σ and λ through they C^1 -norms.

2) *Higher regularity estimates for smooth compactly supported solutions.* Let's drop the subindex δ for a moment and assume u is smooth and compactly supported. Without lost of generality we assume $\mathcal{U} = X_0 \times \Theta_0$, neighborhood of $(x_0, \theta_0) \in \mathcal{Q}$.

At the core of this proof is the next commutator identity. Denoting the transport operator

$$T = (\partial_{x^n} + v \cdot \nabla_{x'}),$$

we have:

$$\partial_{x^j} = \partial_{v_j} T - T \partial_{v_j}, \quad j = 1, \dots, n-1.$$

Under our new notation (A.1) rewrites as

$$(A.3) \quad \mathcal{L}u := Tu - \nabla_v \cdot (A \nabla_v u) + B \cdot \nabla_v u + cu = f, \quad \forall (x, v) \in \mathcal{U}.$$

We abbreviate the above expression as $Tu = h$ with h of the form

$$h = \nabla_v \cdot (A \nabla_v u) - B \cdot \nabla_v u - cu + f.$$

Denoting $\langle f, g \rangle = \int f g dx dv$ the L^2 -bracket, we have:

$$\begin{aligned} \|D_{x'}^{-1/3} \partial_{x^j} u\|_{L^2}^2 &= \langle D_{x'}^{-2/3} \partial_{x^j} \bar{u}, \partial_{x^j} u \rangle \\ &= \langle D_{x'}^{-2/3} \partial_{x^j} \bar{u}, \partial_{v_j} Tu - T \partial_{v_j} u \rangle \\ &= \langle D_{x'}^{-2/3} \partial_{x^j} \bar{u}, \partial_{v_j} h - T \partial_{v_j} u \rangle \\ &= -\langle \partial_{v_j} D_{x'}^{-2/3} \partial_{x^j} \bar{u}, h \rangle + \langle T D_{x'}^{-2/3} \partial_{x^j} \bar{u}, \partial_{v_j} u \rangle \\ &= -\langle \partial_{v_j} D_{x'}^{-2/3} \partial_{x^j} \bar{u}, h \rangle + \langle D_{x'}^{-2/3} \partial_{x^j} T \bar{u}, \partial_{v_j} u \rangle \\ &= -\langle \partial_{v_j} D_{x'}^{-2/3} \partial_{x^j} \bar{u}, h \rangle - \langle \bar{h}, D_{x'}^{-2/3} \partial_{x^j} \partial_{v_j} u \rangle \\ &= -\langle \partial_{v_j} D_{x'}^{-2/3} \partial_{x^j} \bar{u}, h \rangle - \langle D_{x'}^{-2/3} \partial_{x^j} \partial_{v_j} u, \bar{h} \rangle \\ &= -2\Re \langle D_{x'}^{-2/3} \partial_{x^j} \partial_{v_j} \bar{u}, h \rangle. \end{aligned}$$

We can then bound from above as follows,

$$\begin{aligned} \|D_{x'}^{-1/3} \partial_{x^j} u\|_{L^2}^2 &\leq 2 \|D_{x'}^{-2/3} \partial_{x^j} \partial_{v_j} u\|_{L^2} \|h\|_{L^2} \\ &\leq C \sum_{i=1}^n \|D_{x'}^{-2/3} \partial_{x^i} \nabla_v u\|_{L^2} \|h\|_{L^2} \\ &\leq C \|D_{x'}^{1/3} \nabla_v u\|_{L^2} \|h\|_{L^2}. \end{aligned}$$

therefore,

$$(A.4) \quad \|D_{x'}^{2/3} u\|_{L^2}^2 \leq C \|D_{x'}^{1/3} \nabla_v u\|_{L^2} \|h\|_{L^2}.$$

On the other hand,

$$\mathcal{L} D_{x'}^{1/3} u = D_{x'}^{1/3} f + [\mathcal{L}, D_{x'}^{1/3}] u,$$

which implies

$$(T - \nabla_v \cdot A \nabla_v) D_{x'}^{1/3} u = D_{x'}^{1/3} f + \nabla_v \cdot (D_{x'}^{1/3} A) \nabla_v u - D_{x'}^{1/3} (B \cdot \nabla_v u) - D_{x'}^{1/3} (cu).$$

Multiplying by $D_{x'}^{1/3} \bar{u}$ and integrating gives us

$$\begin{aligned} \int_{\mathbb{R}_+^n \times \mathbb{R}^{n-1}} \langle A \nabla_v D_{x'}^{1/3} u, \nabla_v D_{x'}^{1/3} u \rangle dx dv &= \Re \langle D_{x'}^{1/3} \bar{u}, D_{x'}^{1/3} (f - cu) \rangle + \Re \langle D_{x'}^{1/3} \bar{u}, \nabla_v \cdot (D_{x'}^{1/3} A) \nabla_v u \rangle \\ &= \Re \langle D_{x'}^{2/3} \bar{u}, (f - cu) \rangle - \Re \langle D_{x'}^{1/3} \nabla_v \bar{u}, (D_{x'}^{1/3} A) \nabla_v u \rangle, \end{aligned}$$

where some of the term have vanished due to integration by parts. Then, recalling that A is positive definite, we obtain

$$\begin{aligned}\epsilon^2 \|\nabla_v D_{x'}^{1/3} u\|_{L^2}^2 &\leq C \|D_{x'}^{2/3} u\|_{L^2} \|f\|_{L^2} + \|\nabla_v D_{x'}^{1/3} u\|_{L^2} \|(D_{x'}^{1/3} A) \nabla_v u\|_{L^2} \\ &\leq C \|D_{x'}^{2/3} u\|_{L^2} \|f\|_{L^2} + C \epsilon \|\nabla_v D_{x'}^{1/3} u\|_{L^2} \|f\|_{L^2},\end{aligned}$$

where we used the inequality $\epsilon \|\nabla_v u\| + \|u\|_{L^2} \leq \|f\|_{L^2}$ which is the natural energy estimate of (A.3). We can directly combine the previous inequality with (A.4) to deduce

$$\epsilon^2 \|\nabla_v D_{x'}^{1/3} u\|_{L^2}^2 \leq C \|D_x^{1/3} \nabla_v u\|_{L^2}^{1/2} \|h\|_{L^2}^{1/2} \|f\|_{L^2} + C \epsilon \|\nabla_v D_{x'}^{1/3} u\|_{L^2} \|f\|_{L^2}.$$

Simplifying some terms and applying Young's inequality: $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ for $p^{-1} + q^{-1} = 1$; we get

$$\|\nabla_v D_{x'}^{1/3} u\|_{L^2}^{3/2} \leq C \epsilon^{-2} \|h\|_{L^2}^{1/2} \|f\|_{L^2} + C \epsilon^{-3/2} \|f\|_{L^2}^{3/2},$$

which can be rewritten as

$$\|\nabla_v D_{x'}^{1/3} u\|_{L^2} \leq C (\epsilon^{-4/3} \|h\|_{L^2}^{1/3} \|f\|_{L^2}^{2/3} + \epsilon^{-1} \|f\|_{L^2}).$$

Plugging this into (A.4) gives

$$(A.5) \quad \|D_{x'}^{2/3} u\|_{L^2} \leq C (\epsilon^{-2/3} \|h\|_{L^2}^{2/3} \|f\|_{L^2}^{1/3} + \epsilon^{-1/2} \|f\|_{L^2}^{1/2} \|h\|_{L^2}^{1/2}).$$

If we now multiply (A.3) by $-\nabla_v \cdot A \nabla_v \bar{u}$ and integrate, this yields

$$\Re \langle -\nabla_v \cdot A \nabla_v \bar{u}, Tu \rangle + \int |-\nabla_v \cdot A \nabla_v u|^2 dx dv = \Re \langle -\nabla_v \cdot A \nabla_v \bar{u}, -B \cdot \nabla_v u - cu + f \rangle.$$

However,

$$\begin{aligned}|\Re \langle -\nabla_v \cdot A \nabla_v \bar{u}, Tu \rangle| &\leq |\Re \langle A \nabla_v \bar{u}, \nabla_v Tu \rangle| \\ &= |\Re \langle A \nabla_v \bar{u}, T \nabla_v u + \nabla_{x'} u \rangle| \\ &= |\Re \langle [-T, A] \nabla_v \bar{u}, \nabla_v u \rangle| + |\Re \langle A \nabla_v \bar{u}, \nabla_{x'} u \rangle| \\ &\leq C (\epsilon^2 \|\nabla_v u\|_{L^2}^2 + \|D_{x'}^{1/3} \nabla_v u\|_{L^2} \|D_{x'}^{-1/3} \nabla_{x'} u\|_{L^2}) \\ &\leq C (\epsilon^2 \|\nabla_v u\|_{L^2}^2 + \|D_{x'}^{1/3} \nabla_v u\|_{L^2} \|D_{x'}^{2/3} u\|_{L^2}).\end{aligned}$$

therefore

$$\begin{aligned}\|\nabla_v \cdot A \nabla_v u\|_{L^2}^2 &\leq C \|\nabla_v \cdot A \nabla_v u\|_{L^2} (\epsilon^2 \|\nabla_v u\|_{L^2} + \|u\|_{L^2} + \|f\|_{L^2}) \\ &\quad + C (\epsilon^2 \|\nabla_v u\|_{L^2}^2 + \|D_{x'}^{1/3} \nabla_v u\|_{L^2} \|D_{x'}^{2/3} u\|_{L^2}).\end{aligned}$$

It follows from Hölder inequality in Fourier domain that

$$\|D_{x'}^{1/3} \nabla_v u\|_{L^2} \leq \|D_{x'}^{2/3} u\|_{L^2}^{1/2} \|\Delta_v u\|_{L^2}^{1/2} \leq C \|D_{x'}^{2/3} u\|_{L^2}^{1/2} \|\nabla_v \cdot A \nabla_v u\|_{L^2}^{1/2},$$

which plugged into the previous estimate gives

$$\begin{aligned}\|\nabla_v \cdot A \nabla_v u\|_{L^2}^2 &\leq C \|\nabla_v \cdot A \nabla_v u\|_{L^2} (\epsilon^2 \|\nabla_v u\|_{L^2} + \|u\|_{L^2} + \|f\|_{L^2}) \\ &\quad + C (\epsilon^2 \|\nabla_v u\|_{L^2}^2 + \|\nabla_v \cdot A \nabla_v u\|_{L^2}^{1/2} \|D_{x'}^{2/3} u\|_{L^2}^{3/2}) \\ &\leq C \|\nabla_v \cdot A \nabla_v u\|_{L^2} \|f\|_{L^2} + C \|f\|_{L^2}^2 \\ &\quad + C \|\nabla_v \cdot A \nabla_v u\|_{L^2}^{1/2} (\epsilon^{-1} \|h\|_{L^2} \|f\|_{L^2}^{1/2} + \epsilon^{-3/4} \|f\|_{L^2}^{3/4} \|h\|_{L^2}^{3/4}),\end{aligned}$$

which then implies

$$\begin{aligned} \|\nabla_v \cdot A \nabla_v u\|_{L^2}^2 &\leq C \|f\|_{L^2}^2 + C(\epsilon^{-1} \|h\|_{L^2} \|f\|_{L^2}^{1/2} + \epsilon^{-3/4} \|f\|_{L^2}^{3/4} \|h\|_{L^2}^{3/4})^{4/3} \\ &\leq C(\|f\|_{L^2}^2 + \epsilon^{-4/3} \|h\|_{L^2}^{4/3} \|f\|_{L^2}^{2/3} + \epsilon^{-1} \|f\|_{L^2} \|h\|_{L^2}) \end{aligned}$$

Young's inequality yields

$$\|\nabla_v \cdot A \nabla_v u\|_{L^2} \leq C(\|f\|_{L^2} + \epsilon^2 \|h\|),$$

thus, from the estimate $\|h\|_{L^2} \leq \|\nabla_v \cdot A \nabla_v u\|_{L^2} + C\|f\|_{L^2}$, we finally deduce

$$\|\nabla_v \cdot A \nabla_v u\|_{L^2} \leq C\|f\|_{L^2}.$$

It follows from the local representation of Δ_θ in beam coordinates that

$$\epsilon^2 \|\Delta_\theta u\|_{L^2} \leq C\|f\|_{L^2}.$$

On the other hand, the previous inequality implies $\|h\|_{L^2} \leq C\|f\|_{L^2}$, therefore (A.5) yields

$$\|D_{x'}^{2/3} u\|_{L^2} \leq C(\epsilon^{-2/3} \|f\|_{L^2} + \epsilon^{-1/2} \|f\|_{L^2}) \leq C\epsilon^{-2/3} \|f\|_{L^2}.$$

Considering local coordinates given instead by beam coordinates around $(1, 0 \dots, 0)$ (whose domain contains θ_0), we can repeat the computations above in terms of the pseudodifferential operator $D_{x''}^s = (1 - \Delta_{x''})^{s/2}$, where we decompose the spatial variables as $x = (x^1, x'')$. Consequently, we obtain an analogous inequality for $\|D_{x''}^{2/3} u\|_{L^2}$. It then follows that

$$\|D_{x''}^{2/3} u\|_{L^2} \leq C\epsilon^{-2/3} \|f\|_{L^2},$$

for $D_x^s = (1 - \Delta_x)^{s/2}$, which is obtained by recalling that in Fourier space:

$$(1 + |\xi|^2)^{s/2} \leq (1 + |\xi'|^2 + 1 + |\xi''|^2)^{s/2} \leq (1 + |\xi'|^2)^{s/2} + (1 + |\xi''|^2)^{s/2}.$$

3) *Back to original solution.* We now go back to our original notation and write u^δ for the smooth compactly supported solution and u the original solution to (1.1). From the previous, we can find a sequence $\{\delta_j\}$ so that $\Delta_\theta u^{\delta_j}$ and $D_x^{2/3} u^{\delta_j}$ converge weakly in $L^2(\mathcal{V})$ to $\Delta_\theta u, D_x^{2/3} u \in L^2(\mathcal{V})$, respectively, as $j \rightarrow \infty$. Moreover,

$$\|D_x^{2/3} u\|_{L^2(\mathcal{V})} \leq \liminf \|D_x^{2/3} u^{\delta_j}\|_{L^2(\mathcal{V})} \leq C\|f^{\delta_j}\|_{L^2(\mathcal{U})} \leq C(\|f\|_{L^2(\mathcal{W})} + \|u\|_{\mathcal{K}}),$$

and similarly for $\|\Delta_\theta u\|_{L^2(\mathcal{V})}$. We have proven that $\Delta_\theta u, D_x^{2/3} u \in L_{loc}^2(Q)$. Moreover, for every compact \mathcal{K} and open \mathcal{O} such that $\mathcal{K} \subset \mathcal{O} \subset Q$, there exist a constant $C, C' > 0$ so that

$$\epsilon^2 \|\Delta_\theta u\|_{L^2(\mathcal{K})} \leq C(\|f\|_{L^2(\mathcal{O})} + \epsilon^2 \|\nabla_\theta u\|_{L^2} + \|u\|_{L^2}),$$

and

$$\epsilon^{2/3} \|D_x^{2/3} u\|_{L^2(\mathcal{K})} \leq C'(\|f\|_{L^2(\mathcal{O})} + \epsilon^2 \|\nabla_\theta u\|_{L^2} + \|u\|_{L^2}).$$

Using the equation satisfied by u we also deduce $\theta \cdot \nabla_x u \in L_{loc}^2(Q)$ and

$$\|\theta \cdot \nabla_x u\|_{L^2(\mathcal{K})} \leq C(\|f\|_{L^2(\mathcal{O})} + \epsilon^2 \|\nabla_\theta u\|_{L^2} + \|u\|_{L^2}),$$

for some C depending on the compact \mathcal{K} .

APPENDIX B. EXPLICIT SOLUTION TO (4.15)

Let W be solution to the backward Fermi pencil-beam equation

$$-\tilde{\sigma}\Delta_V W - V \cdot \nabla_{X'} W - \partial_{X^n} W + \tilde{\lambda} W = \Psi, \quad (X, V) \in \mathbb{R}_+^n \times \mathbb{R}^{n-1},$$

for Ψ compactly supported and with vanishing conditions at infinity

$$\lim_{X^n \rightarrow \infty} W = 0.$$

We solve the equation in the whole space $\mathbb{R}^n \times \mathbb{R}^{n-1}$ and claim that W takes the form

$$W(X, V) = \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \tau_{-X^n V}(\mathbf{H}_3(t)) * \tau_{tV} \Psi(t) dt.$$

Let's introduce the following functions:

$$U(X, V) := W(-X, V) \quad F(X, V) := \Psi(-X, V),$$

defined for $X^n \leq 0$. Then U satisfies the equation

$$-\tilde{\sigma}^- \Delta_V U + V \cdot \nabla_{X'} U + \partial_{X^n} U + \tilde{\lambda}^- U = F, \quad (X, V) \in \mathbb{R}^n \times \mathbb{R}^{n-1},$$

with $\tilde{\sigma}^-(t) = \tilde{\sigma}(-t)$ and $\tilde{\lambda}^-(t) = \tilde{\lambda}(-t)$. Let's set

$$\mathbf{U}(\xi, X^n, V) = e^{iX^n(V \cdot \xi) + \int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X'}(U), \quad \text{and} \quad \mathbf{F}(\xi, X^n, V) = e^{iX^n(V \cdot \xi) + \int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X'}(F),$$

therefore

$$\begin{aligned} \partial_{X^n} \mathbf{U} &= [i(V \cdot \xi) + \tilde{\lambda}^-] \mathbf{U} + e^{iX^n(V \cdot \xi) + \int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \partial_{X^n} \mathfrak{F}_{X'}(U) \\ &= \tilde{\sigma}^- e^{iX^n(V \cdot \xi) + \int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \Delta_V (e^{-iX^n(V \cdot \xi) - \int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X'}(U)) + \mathbf{F} \\ &= \tilde{\sigma}^- e^{iX^n(V \cdot \xi)} \Delta_V (e^{-iX^n(V \cdot \xi)} \mathfrak{F}_{X'}(U)) + \mathbf{F} \end{aligned}$$

Denoting

$$\hat{\mathbf{U}}(\xi, X^n, \eta) = \mathfrak{F}_V(\mathbf{U}) \quad \text{and} \quad \hat{\mathbf{F}}(\xi, X^n, \eta) = \mathfrak{F}_V(\mathbf{F}),$$

then taking Fourier Transform on the equation above and using twice the identity

$$\mathfrak{F}_V[e^{it(V \cdot \xi)} h(V)](\eta) = \mathfrak{F}_V[h(V)](\eta - t\xi), \quad t \in \mathbb{R},$$

we obtain

$$\partial_{X^n} \hat{\mathbf{U}} + \tilde{\sigma}^- |\eta - X^n \xi|^2 \hat{\mathbf{U}} = \hat{\mathbf{F}}.$$

Recalling the vanishing condition imposed at infinity, $\lim_{X^n \rightarrow -\infty} U = 0$, an explicit expression for $\hat{\mathbf{U}}$ is given by

$$\hat{\mathbf{U}}(\xi, X^n, \eta) = \int_{-\infty}^{X^n} e^{-\int_t^{X^n} |\eta - s\xi|^2 \tilde{\sigma}^-(s) ds} \hat{\mathbf{F}}(\xi, t, \eta) dt.$$

Since $\hat{\mathbf{U}}(\xi, X^n, \eta) = e^{\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X', V}[U](\xi, X^n, \eta - X^n \xi)$, then

$$\begin{aligned} U(X', X^n, V) &= e^{-\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{U}}(\xi, X^n, \eta + X^n \xi)](X, V) \\ &= e^{-\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X'}^{-1}[e^{-iX^n(V \cdot \xi)} \mathfrak{F}_{X'}^{-1}[\hat{\mathbf{U}}(\xi, X^n, \eta)]](X, V) \\ &= e^{-\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{U}}](X' - X^n V, X^n, V). \end{aligned}$$

Therefore

$$\begin{aligned} U(X', X^n, V) &= e^{-\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \tau_{X^n V} (\mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{U}}]) \\ &= e^{-\int_{-\infty}^{X^n} \tilde{\lambda}^-(s) ds} \tau_{X^n V} \left(\int_{-\infty}^{X^n} \tilde{\mathbf{H}}_3(t) * \mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{F}}(t)] dt \right) \end{aligned}$$

with

$$\tilde{\mathbf{H}}_3(X', X^n, V; t) = \mathfrak{F}_{X', V}^{-1} [e^{-\int_t^{X^n} |\eta + (X^n - s)\xi|^2 \tilde{\sigma}^-(s) ds}].$$

Analogously as for U , we have

$$F(X', t, V) = e^{-\int_{-\infty}^t \tilde{\lambda}^-(s) ds} \mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{F}}](X' - tV, t, V),$$

which implies $\mathfrak{F}_{X', V}^{-1}[\hat{\mathbf{F}}] = e^{\int_{-\infty}^t \tilde{\lambda}^-(s) ds} \tau_{-tV}(F)$, and consequently

$$U(X', X^n, V) = \int_{-\infty}^{X^n} e^{-\int_t^{X^n} \tilde{\lambda}^-(s) ds} \tau_{X^n V} \left(\tilde{\mathbf{H}}_3(X, V; t) * \tau_{-tV}(F(X', t, V)) \right) dt.$$

Going back to the original function, for $X^n > 0$ we obtain

$$\begin{aligned} W(X', X^n, V) &= U(-X', -X^n, V) \\ &= \int_{-\infty}^{-X^n} e^{-\int_t^{-X^n} \tilde{\lambda}^-(s) ds} \tau_{-X^n V} \left(\tilde{\mathbf{H}}_3(-X, V; t) * \tau_{-tV}(F(-X', t, V)) \right) dt \\ &= \int_{X^n}^{\infty} e^{-\int_{-t}^{-X^n} \tilde{\lambda}^-(s) ds} \tau_{-X^n V} \left(\tilde{\mathbf{H}}_3(-X, V; -t) * \tau_{tV}(F(-X', -t, V)) \right) dt \\ &= \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \tau_{-X^n V} \left(\tilde{\mathbf{H}}_3(-X, V; -t) * \tau_{tV}(\Psi(X', t, V)) \right) dt, \end{aligned}$$

We conclude by writing $\mathbf{H}_3(X, V; t) = \tilde{\mathbf{H}}_3(-X, V; -t)$.

APPENDIX C. PROOF OF LEMMA 4.4

Let W be the solution to the backward Fermi pencil-beam equation with vanishing condition at infinity (4.15)-(4.16). It is given explicitly by

$$W(X, V) = \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \tau_{-X^n V} (\mathbf{H}_3(t) * \tau_{tV} \Psi(t)) dt.$$

For $(X'_1, X^n, V_1), (X'_2, X^n, V_2) \in \mathbb{R}^{2(n-1)}$ we have

$$\begin{aligned}
& |W(X'_1, X^n, V_1) - W(X'_2, X^n, V_2)| \\
& \leq \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \int \int \mathbf{H}_3(\tilde{X}', X^n, \tilde{V}; t) \\
& \quad \times |\Psi(X'_1 - \tilde{X}' + (X^n - t)V_1 + t\tilde{V}, t, V_1 - \tilde{V}) - \Psi(X'_2 - \tilde{X}' + (X^n - t)V_2 + t\tilde{V}, t, V_2 - \tilde{V})| d\tilde{X}' d\tilde{V} dt \\
& \leq \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \int \int \mathbf{H}_3(\tilde{X}', X^n, \tilde{V}; t) \\
& \quad \times |(X'_1 - X'_2 + (X^n - t)(V_1 - V_2), V_1 - V_2)| \text{Lip}(\Psi) d\tilde{X}' d\tilde{V} dt \\
& = \text{Lip}(\Psi) \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} \underbrace{\|\mathbf{H}_3(X^n; t)\|_{L^1(\mathbb{R}^{2(n-1)})}}_{=const.} |(X'_1 - X'_2 + (X^n - t)(V_1 - V_2), V_1 - V_2)| dt \\
& \leq C \text{Lip}(\Psi) |(X'_1, V_1) - (X'_2, V_2)| \int_{X^n}^{\infty} e^{-\int_{X^n}^t \tilde{\lambda}(s) ds} |t - X^n| dt \\
& = C \text{Lip}(\Psi) |(X'_1, V_1) - (X'_2, V_2)| \int_0^{\infty} e^{-\int_{X^n}^{\tau+X^n} \tilde{\lambda}(s) ds} \tau d\tau \\
& \leq C \lambda_0^{-2} \text{Lip}(\Psi) |(X'_1, V_1) - (X'_2, V_2)|.
\end{aligned}$$

REFERENCES

- [1] R. ALONSO, & W. SUN, *The radiative transfer equation in the forward-peaked regime*, Comm. Math. Phys., 338(3) (2015), pp. 1233–1286.
- [2] M. ASADZADEH, *The Fokker-Planck operator as an asymptotic limit in anisotropic media*, Math. Comput. Model., 35(9-10) (2002), pp. 1119–1133.
- [3] G. BAL & A. JOLLIVET, *Generalized stability estimates in inverse transport theory*, Inverse Probl. Imaging, 12(1) (2018), pp. 59–90.
- [4] G. BAL, T. KOMOROWSKI & L. RYZHIK, *Kinetic limits for waves in a random medium*. Kinet. Relat. Models, 3(4) (2010), pp. 529–644.
- [5] C. BARDOS, *Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d'approximation application à l'équation de transport*, Ann. Sci. Ecole Norm. Sup. Partis, 3 (1970), pp. 185–233.
- [6] C. BÖRGERS & E. W. LARSEN, *Asymptotic derivation of the Fermi pencil-beam approximation*, Nuclear science and engineering, 123(3) (1996), pp. 343–357.
- [7] C. BÖRGERS & E. W. LARSEN, *On the accuracy of the Fokker-Planck and Fermi pencil beam equations for charged particle transport*, Medical Physics, 23(10) 1996, pp. 1749–1759.
- [8] F. BOUCHUT, *Hypoelliptic regularity in kinetic equations*, J. Math. Pures Appl., 81 (2002), pp. 1135–1159.
- [9] M. BRAMANTI. *An invitation to hypoelliptic operators and Hörmander's vector fields*, New York: Springer, 2014.
- [10] A. CARILLO, *Global weak solutions for the initial-boundary-value problems to the Vlasov-Poisson-Fokker-Planck system*, Math. Meth. Appl. Sci., 21 (1998), pp. 907–938.
- [11] K. CHEN, Q. LI & L. WANG, *Stability of inverse transport equation in diffusion scaling and Fokker-Planck limit*, SIAM J. Appl. Math., 78(5) (2018), pp. 2626–2647.
- [12] E. CUEVA, M. COURDURIER, A. OSSES, V. CASTAÑEDA, B. PALACIOS & S. HÄRTEL, *Mathematical Modeling for 2D Light-Sheet Fluorescence Microscopy image reconstruction*, Inverse Problems, (2020).
- [13] P. DEGOND & S. MAS-GALLIC, *Existence of solutions and diffusion approximation for a model Fokker-Planck equation*, Transport Theor. Stat., 16(4-6) (1987), pp. 589–636.
- [14] L. EYGES, *Multiple scattering with energy loss*, Physical Review, 74(10) (1948), pp. 1534.
- [15] F. GOLSE, C. IMBERT, C. MOUHOT & F. VASSEUR, *Harnack inequality for kinetic Fokker-Planck equation with rough coefficients and application to the Landau equation*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 19(1) (2019), pp. 253–295.

- [16] C. GOMEZ, O. PINAUD, & L. RYZHIK, *Hypoelliptic estimates in radiative transfer*, Comm. Partial Differential Equations, 41(1) (2016), pp. 150–184.
- [17] F. HANSON, I. BENDALL, C. DECKARD & H. HAIDAR, *Off-axis detection and characterization of laser beams in the maritime atmosphere*, Applied optics, 50(18) (2011), pp. 3050–3056.
- [18] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.
- [19] H. J. HWANG, J. JANG, & J. JUNG, *On the kinetic Fokker-Planck equation in a half-space with absorbing barriers*, Indiana Univ. Math. J., 64(6) (2015), pp. 1767–1804.
- [20] H.J. Hwang, J. Jang, & J. Jung, *The Fokker–Planck Equation with Absorbing Boundary Conditions in Bounded Domains*, SIAM J. Math. Anal., 50(2) (2018), pp. 2194–2232.
- [21] C. Imbert, & L. Silvestre, *The weak Harnack inequality for the Boltzmann equation without cut-off*, J. Eur. Math. Soc., 22(2) (2020), pp. 507–592.
- [22] A.D. KIM, & J. B. KELLER, *Light propagation in biological tissue*, JOSA A, 20(1) (2003), pp. 92–98.
- [23] J. J. KOHN, *Pseudo-differential operators and hypoellipticity*, Proc. Symp. Pure Math, 23 (1973), pp. 61–69.
- [24] J. M. LEE, *Introduction to Riemannian manifolds* (Vol. 2), Springer, 2018.
- [25] J. L. Lions, *Equations différentielles opérationnelles et problèmes aux limites* (Vol. 111). Springer-Verlag, 2013.
- [26] G. C. POMRANING, *The Fokker-Planck operator as an asymptotic limit*, Math. Models Methods Appl. Sci, 2(1) (1992), pp. 21–36.
- [27] N. ROY, & F. REID, *Off-axis laser detection model in coastal areas*, Optical Engineering, 47(8) (2008), pp. 086002.
- [28] Q. SHENG, & W. HAN, *Well-posedness of the FokkerPlanck equation in a scattering process*, J. Math. Anal. Appl., 406(2) (2013), pp. 531–536.
- [29] E. M. STEIN & R. SHAKARCHI, *Real analysis: measure theory, integration, and Hilbert spaces*, Princeton University Press, 2009.
- [30] W. WANG, & L. ZHANG, *C^α -regularity of weak solutions of non-homogeneous ultraparabolic equations with drift terms*, preprint, arXiv:1704.05323v2 [math.AP], 2019.