

# ON QUASILINEAR MAXWELL EQUATIONS IN TWO DIMENSIONS

ROBERT SCHIPPA AND ROLAND SCHNAUBELT

ABSTRACT. New sharp Strichartz estimates for the Maxwell system in two dimensions with rough permittivity and non-trivial charges are proved. We use the FBI transform to carry out the analysis in phase space. For this purpose, the Maxwell equations are conjugated to a system of half-wave equations with rough coefficients, for which Strichartz estimates are similarly derived as in previous work by Tataru on scalar wave equations with rough coefficients. We use the estimates to improve the local well-posedness theory for quasilinear Maxwell equations in two dimensions.

## 1. INTRODUCTION AND MAIN RESULTS

The Maxwell equations in three spatial dimensions govern the propagation of electromagnetic fields. We refer to the physics literature with its many excellent accounts (e.g. [10, 20]) for explaining the role of *electric and magnetic fields*  $(E, B) : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \times \mathbb{R}^3$  and *displacement and magnetizing fields*  $(D, H) : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \times \mathbb{R}^3$ . The *electric charges*  $\rho_e : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$  act as sources of the displacement field. In the following space-time coordinates are denoted by  $x = (x_0, x_1, \dots, x_n) = (t, x') \in \mathbb{R} \times \mathbb{R}^n$  and the dual variables in Fourier space by  $\xi = (\xi_0, \xi_1, \dots, \xi_n) = (\tau, \xi') \in \mathbb{R} \times \mathbb{R}^n$ .

In the absence of currents, the Maxwell system in media is given by

$$(1) \quad \begin{cases} \partial_t D = \nabla \times H, & \nabla \cdot D = \rho_e, \\ \partial_t B = -\nabla \times E, & \nabla \cdot B = 0, \\ E(0, \cdot) = E_0, & B(0, \cdot) = B_0. \end{cases}$$

These equations have to be supplemented with material laws linking, e.g.,  $E$  with  $D$  and  $H$  with  $B$ . We consider the *constitutive relations*

$$(2) \quad \begin{aligned} D(x) &= \varepsilon(x)E(x), & \varepsilon : \mathbb{R} \times \mathbb{R}^3 &\rightarrow \mathbb{R}^{3 \times 3}, \\ B(x) &= \mu(x)H(x), & \mu : \mathbb{R} \times \mathbb{R}^3 &\rightarrow \mathbb{R}^{3 \times 3}, \end{aligned}$$

which are linear, pointwise, and instantaneous. The coefficient  $\varepsilon$  is referred to as *permittivity* and  $\mu$  as *permeability*. We aim to describe dispersive properties of electromagnetic fields in possibly anisotropic and inhomogeneous media, so that we allow for  $x$ -dependent and matrix-valued coefficients. In the following we consider  $\mu \equiv 1$  for simplicity. The relations (2) with constant  $\mu$  are frequently used to model phenomena in optics (cf. [28]). We remark that our arguments extend to a variable permeability provided it satisfies the same ellipticity and regularity assumptions

---

2020 *Mathematics Subject Classification*. Primary: 35L45, 35B65, Secondary: 35Q61.

*Key words and phrases*. Maxwell equations, Strichartz estimates, quasilinear wave equations, rough coefficients, half wave equation, FBI transform.

as  $\varepsilon$ . Below  $\mu$  denotes a regularity parameter, which is no longer related to the permeability.

We first focus on rough permittivity  $\varepsilon$  with coefficients in  $C^s$  for  $0 < s \leq 2$  as an intermediate step to quasilinear Maxwell equations, where  $\varepsilon = \varepsilon(E)$ . A prominent example is the *Kerr nonlinearity* given by

$$(3) \quad \varepsilon = \varepsilon(E) = (1 + |E|^2).$$

In this paper the Maxwell system in two spatial dimensions is considered, which can be derived taking  $E$  to be perpendicular to media interfaces (cf. [4, 25]). Actually, if  $\varepsilon_{j3} = \varepsilon_{3j} = 0$  for  $j \in \{1, 2\}$ , if  $E_0, B_0 = H_0$ , and  $\rho_e$  in (1) only depend on  $(x, y) \in \mathbb{R}^2$ , and if the components  $E_{03}, H_{01}$  and  $H_{02}$  vanish, then the solutions  $(E, H)$  to (1) have the same properties. Hence, the resulting Maxwell system in two spatial dimensions is given by

$$(4) \quad \begin{cases} \partial_t D = \nabla_\perp H, & \nabla \cdot D = \rho_e, \\ \partial_t H = -\nabla \times E, \\ D(0, \cdot) = D_0, & H(0, \cdot) = H_0. \end{cases}$$

In the above display we have  $D, E : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$  and  $H, \rho_e : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ , and we set  $\nabla_\perp = (\partial_2, -\partial_1)^t$ . We suppose that  $\varepsilon$  is a matrix-valued function  $\varepsilon : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$  such that for some constants  $\Lambda_1, \Lambda_2 > 0$  and all  $\xi' \in \mathbb{R}^2$  and  $x \in \mathbb{R} \times \mathbb{R}^2$  we have

$$(5) \quad \Lambda_1 |\xi'|^2 \leq \sum_{i,j=1}^2 \varepsilon^{ij}(x) \xi_i \xi_j \leq \Lambda_2 |\xi'|^2, \quad \varepsilon^{ij}(x) = \varepsilon^{ji}(x).$$

Throughout the paper we shall use sum convention and sum over indices appearing twice, e.g.,

$$\varepsilon^{ij}(x) \xi_i \xi_j = \sum_{i,j=1}^2 \varepsilon^{ij}(x) \xi_i \xi_j.$$

Using the matrix

$$(6) \quad P(x, D) = \begin{pmatrix} \partial_t & 0 & -\partial_2 \\ 0 & \partial_t & \partial_1 \\ -\partial_2(\varepsilon_{11} \cdot) + \partial_1(\varepsilon_{21} \cdot) & \partial_1(\varepsilon_{22} \cdot) - \partial_2(\varepsilon_{12} \cdot) & \partial_t \end{pmatrix}$$

with rough symbols, the PDEs in (4) can be rewritten as

$$(7) \quad P(x, D)(D_1, D_2, H) = 0, \quad \nabla \cdot D = \rho_e,$$

where we set

$$\varepsilon^{-1}(x) = (\varepsilon_{ij}(x))_{i,j=1,2}.$$

Let  $P(x, D)(D_1, D_2, H) = (g_1, g_2, h)$ , where  $J = -(g_1, g_2)$  is the electric current and  $h$  has no physical meaning. Then the electric charges in (7) are given by

$$(8) \quad \rho_e(t) = \nabla \cdot D_0 + \int_0^t \nabla \cdot (g_1, g_2) ds.$$

There is a large body of literature for Maxwell equations on smooth space-times, investigating more fundamental decay properties in higher dimensions (cf. [24, 26]). In these works, local energy decay is proved, from which Strichartz estimates can be derived.

Note that well-posedness of (7) with  $\varepsilon = \varepsilon(E)$  in Sobolev spaces  $H^s(\mathbb{R}^2)$  with  $s > 2$  can be established by the energy method for hyperbolic systems (cf. [2, 3, 22]). We revisit the argument in the last section and show how Strichartz estimates

yield improvements. For a detailed account on local well-posedness results for the Maxwell system on spatial domains in  $\mathbb{R}^3$  with nonlinear material laws, we refer to the PhD thesis [34] by M. Spitz. In [35] Spitz showed local well-posedness of the Maxwell equations with perfectly conducting boundary conditions in  $H^3(G)$  for domains  $G \subseteq \mathbb{R}^3$ . This approach neglects dispersive effects.

We are not aware of works on Maxwell equations with rough coefficients in the anisotropic case taking advantage of dispersion. In the isotropic case, i.e.,  $\varepsilon(x) = e(x)1_{3 \times 3}$  with  $e : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$  and  $e \in C^2$ , the second author derived global-in-time Strichartz estimates from local energy decay, jointly with P. D'Ancona [7]. Local-in-time estimates for smooth scalar coefficients were treated in [9].

We remark that in the constant-coefficient case, Liess [21] (see also [23]) showed decay estimates by Fourier analytic methods. Liess inferred that in three spatial dimensions, the time-decay of  $(D, H)$  only resembles the time-decay of the three-dimensional wave equation provided that  $\varepsilon$  has less than three eigenvalues. Otherwise, the time-decay corresponds to the one of solutions to the two-dimensional wave equations. This indicates that in three spatial dimensions additional hypotheses are necessary to obtain the range of the three-dimensional wave equation. This will be subject of future work.

In the present paper, we analyze the dispersive properties of the Maxwell system on  $\mathbb{R}^2$  for rough pointwise material laws in the anisotropic case. We derive Strichartz estimates linking Maxwell equations to half-wave equations. The connection is established by analysis in phase space.

To relate our problem to the scalar wave equation, let  $H$  satisfy (4) and  $\varepsilon$  be time-independent. Differentiating the equation, we infer

$$(9) \quad \partial_t^2 H = \partial_2(\varepsilon_{11}\partial_2 H) - \partial_1(\varepsilon_{12}\partial_2 H) - \partial_2(\varepsilon_{21}\partial_1 H) + \partial_1(\varepsilon_{22}\partial_1 H) =: \Delta_{\varepsilon^{-1}} H.$$

Hence,  $H$  solves a wave equation with rough coefficients, for which Strichartz estimates are known (cf. [37, 38, 39]). However, we aim to derive Strichartz estimates directly for the first-order system (4), as taking additional derivatives typically gives rise to loss in regularity.

To put our results into perspective, we review Strichartz estimates for wave equations. For  $u : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{C}$  let

$$\|u\|_{L_{x_0}^p L_{x'}^q} = \left( \int_{\mathbb{R}} \left( \int_{\mathbb{R}^n} |u(x_0, x')|^q dx' \right)^{p/q} dx_0 \right)^{1/p}.$$

We will frequently omit to indicate space and time integration in  $L^p L^q$ -norms and set  $L^p = L^p L^p$ . Keel and Tao [17] established the sharp range for solutions to the wave equation with constant coefficients in Euclidean space. Let

$$u : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \square u = 0, \quad u(0) = u_0, \quad u_t(0) = u_1.$$

Then the estimate

$$(10) \quad \|u\|_{L^p L^q} \lesssim \|u_0\|_{H^\rho} + \|u_1\|_{H^{\rho-1}}$$

holds provided that

$$2 \leq p, q \leq \infty, \quad \rho = n \left( \frac{1}{2} - \frac{1}{q} \right) - \frac{1}{p},$$

$$\frac{2}{p} + \frac{n-1}{q} \leq \frac{n-1}{2}, \quad (p, q, n) \neq (2, \infty, 3).$$

Tupels  $(\rho, p, q, n)$  satisfying these relations will be called Strichartz pairs. If

$$\frac{2}{p} + \frac{n-1}{q} = \frac{n-1}{2},$$

then  $(\rho, p, q, n)$  will be referred to as sharp Strichartz pair. Also note that the sharp Strichartz pairs imply the other ones by Sobolev's embedding.

An important question is the generalization to variable metrics, i.e.,  $g = g(x)$ . On the one hand, it is known (cf. [15, 16, 27, 29]) that solutions to the variable-coefficient wave equation

$$\square_{g(x)} u = \partial_i g^{ij} \partial_j u = 0,$$

with uniformly hyperbolic  $g \in C^\infty$  have the same dispersive properties, at least locally in time, from which Strichartz estimates (10) can be derived.

In view of non-smooth coefficients, Smith and Sogge [31] pointed out that for merely Hölder-continuous metrics  $g \in C^s$  with  $0 < s < 2$  Strichartz estimates (10) fail. On the other hand, Smith [30] showed that Strichartz estimates remain valid for  $C^2$ -coefficients. In a series of papers [37, 38, 39] Tataru then recovered Strichartz estimates (10) for wave equations with  $C^2$ -coefficients; see also the preceding paper [1] by Bahouri and Chemin and the related work by Klainerman [18]. Tataru derived corresponding sharp estimates with additional derivative loss for  $C^s$ -coefficients as a minor variation of the  $C^2$ -case in [38]. Smith and Tataru proved sharpness in [32].

We show the following theorems for the Maxwell system on  $\mathbb{R}^2$ . To state the results, let

$$(|D|^\alpha f)^\wedge(\xi) = |\xi|^\alpha \hat{f}(\xi), \quad (|D'|^\alpha f)^\wedge(\xi) = |\xi'|^\alpha \hat{f}(\xi).$$

For the sake of simplicity, we suppose the fields to be smooth and understand the Strichartz estimates as a priori estimates. This makes no difference for the application to quasilinear equations.

**Theorem 1.1.** *Let  $\varepsilon : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$  be a matrix-valued function with coefficients in  $C^2$  satisfying (5). Let  $u = (D_1, D_2, H) : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$  with  $\partial_1 D_1 + \partial_2 D_2 = \rho_e$ , and  $P$  as in (6). Then, we find the following estimate to hold:*

$$(11) \quad \| |D|^{-\rho} u \|_{L^p L^q} \lesssim \mu \| u \|_{L^2} + \mu^{-1} \| P u \|_{L^2} + \| |D|^{-\frac{1}{2}} \rho_e \|_{L^2}$$

provided that the right hand-side is finite,  $(\rho, p, q, 2)$  is a Strichartz pair, and

$$\| \partial_x^2 \varepsilon \|_{L^\infty} \leq \mu^4.$$

The space  $C^s = C^s(\mathbb{R}^m)$  for  $s \geq 0$  is equipped with its standard norm. For vectors the norms are given by  $\| u \|_X = \| u_1 \|_X + \| u_2 \|_X + \| u_3 \|_X$ . The additional parameter  $\mu$  is crucial to control the size of coefficients when dealing with quasilinear problems.

Note that (11) implies the estimate with  $\| |D'|^{-\frac{1}{2}} \rho_e \|_{L^2}$  on the right-hand side. Moreover, if  $\| |D|^{-\frac{1}{2}} \rho_e \|_{L^2} \sim \| |D|^{\frac{1}{2}} D \|_{L^2}$ , the estimate for the displacement field  $D = (D_1, D_2)$  in (11) already follows from Sobolev's embedding. Hence, the Strichartz estimates are most relevant for charges possessing the additional regularity  $\| |D|^{-\frac{1}{2}} \rho_e \|_{L^2} \lesssim \| D \|_{L^2}$ , since in this case our results point out that the Maxwell system with  $C^2$ -coefficients exhibits the same dispersive properties as scalar wave equations. This includes the important charge-free case, of course.

Truncating the frequencies of  $\varepsilon_{ij}$  appropriately and using the above theorem, as in [38] we can show Strichartz estimates for  $C^s$ -coefficients if we allow for a loss of derivatives compared to (11).

**Theorem 1.2.** *Let  $\varepsilon : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$  be a matrix-valued function with coefficients in  $C^s$ ,  $0 \leq s < 2$ , satisfying (5). Let  $u = (D_1, D_2, H) : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$  with  $\partial_1 D_1 + \partial_2 D_2 = \rho_e$ . Then, we obtain the estimate*

$$(12) \quad \||D|^{-\rho - \frac{\sigma}{2}} u\|_{L^p L^q} \lesssim \mu \|u\|_{L^2} + \mu^{-1} \|Pu\|_{\dot{H}^{-\sigma}} + \||D|^{-\frac{1}{2} - \frac{\sigma}{2}} \rho_e\|_{L^2}$$

provided that the right hand-side is finite,  $(\rho, p, q, 2)$  is a Strichartz pair,

$$\sigma = \frac{2-s}{2+s}, \quad \text{and} \quad \|\varepsilon^{ij}\|_{\dot{C}^s} \leq \mu^4.$$

The conclusion of Theorem 1.1 remains true with small modifications if the second derivatives of the coefficients belong to  $L^1 L^\infty$ . The motivation for this setup is the quasilinear case  $\varepsilon = \varepsilon(E)$ , as discussed below. For (7) we prove the following variant of Theorem 1.1.

**Theorem 1.3.** *Let  $\varepsilon : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$  be a matrix-valued function with Lipschitz coefficients, satisfying (5) and  $\partial_x^2 \varepsilon \in L^1 L^\infty$ . Let  $u = (D_1, D_2, H) : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$  with  $\partial_1 D_1 + \partial_2 D_2 = \rho_e$ , and  $(\rho, p, q, 2)$  be a Strichartz pair. Then,*

$$(13) \quad \begin{aligned} \||D'|^{-\rho} u\|_{L^p(0, T; L^q)} &\lesssim \mu^{\frac{1}{p}} \|u\|_{L^\infty L^2} + \mu^{-\frac{1}{p'}} \|P(x, D)u\|_{L^1 L^2} \\ &\quad + T^{\frac{1}{2}} \||D'|^{-\frac{1}{2}} \rho_e(0)\|_{L^2(\mathbb{R}^2)} + T^{\frac{1}{2}} \||D'|^{-\frac{1}{2}} \partial_t \rho_e\|_{L^1 L^2}, \end{aligned}$$

whenever the right-hand side is finite, provided that  $\mu \geq 1$ , and

$$T \|\partial_x^2 \varepsilon\|_{L^1 L^\infty} \leq \mu^2.$$

Observe that  $\rho_e(0) = \nabla \cdot D_0$  and  $\partial_t \rho_e = \partial_1 (Pu)_1 + \partial_2 (Pu)_2$  due to (8). Compared to (11), above one thus takes the  $L_x^2$  norm into sum and integral appearing in (8). This happens when passing to  $L^\infty L^2$ - and  $L^1 L^2$ -norms on the right-hand side of (13), see Paragraph 3.4.2. These norms are better suited for the application to quasilinear problems than the  $L^2$ -norms appearing on the right-hand side of (11).

But for these applications one still needs a version for coefficients with less regularity. To state it, let  $(S_\lambda)_{\lambda \in 2^{\mathbb{Z}}}$  denote a homogeneous Littlewood–Paley decomposition in space-time, and  $(S'_\lambda)_{\lambda \in \mathbb{Z}}$  one in the spatial variables only. To avoid problems when summing norms on Littlewood–Paley blocks, the regularity of solutions is measured in homogeneous Besov-type spaces  $\dot{B}_s^{pqr}$  with norms

$$\|u\|_{\dot{B}_s^{pqr}}^r = \sum_\lambda \lambda^{rs} \|S_\lambda u\|_{L^p L^q}^r,$$

and the usual modification for  $r = \infty$ . For the coefficients, following [40, 39] we use the microlocalizable scale of spaces  $\mathcal{X}^s$  given by

$$\|v\|_{\mathcal{X}^s} = \sup_\lambda \lambda^s \|S_\lambda v\|_{L^1 L^\infty},$$

The  $\mathcal{X}^s$ -regularity is an adequate substitute for the  $C^s$ -regularity of the coefficients in our setting. For these regularities we prove the following Strichartz estimate.

**Theorem 1.4.** *Let  $\varepsilon \in \mathcal{X}^s$ ,  $0 \leq s < 2$ , and  $u$ ,  $(\rho, p, q, 2)$  and  $\sigma$  be as in the assumptions of Theorem 1.2. Then, we find the following estimate to hold:*

$$(14) \quad \begin{aligned} \||D|^{-\rho - \frac{\sigma}{p}} u\|_{\dot{B}_0^{pq\infty}} &\lesssim \mu^{\frac{1}{p}} \|u\|_{L^\infty L^2} + \mu^{-\frac{1}{p'}} \||D|^{-\sigma} Pu\|_{L^1 L^2} \\ &\quad + T^{\frac{1}{2}} \||D|^{-\frac{1}{2} - \frac{\sigma}{p}} \rho_e\|_{L^\infty L^2} + T^{\frac{1}{2}} \||D|^{-\frac{1}{2} - \frac{\sigma}{p}} \partial_t \rho_e\|_{L^1 L^2} \end{aligned}$$

for all  $u$  compactly supported in  $[0, T]$ , and  $\mu, T$  satisfying

$$T^s \|\varepsilon\|_{\mathcal{X}^s}^2 \lesssim \mu^{2+s}.$$

Estimates involving the norm of  $P(x, D)u$  in  $L^{p'}L^{q'}$  follow as in [38, 39]. We sketch the proof for the next result with  $C^2$  coefficients.

**Theorem 1.5.** *Let  $\varepsilon \in C^2$  and  $(\rho, p, q, 2)$  be a Strichartz pair. Then, we have*

$$(15) \quad \| |D|^{-\rho} u \|_{L^p L^q} \lesssim \mu \|u\|_{L^2} + \mu^{-1} \|f_1\|_{L^2} + \| |D|^\rho f_2 \|_{L^{p'} L^{q'}} + \| |D|^{-\frac{1}{2}} \rho_e \|_{L^2}$$

whenever

$$P(x, D)u = f_1 + f_2 \quad \text{and} \quad \|\partial_x^2 \varepsilon\|_\infty \leq \mu^4.$$

One can establish versions for  $C^s$ -coefficients with  $0 \leq s < 2$ , cf. [38, Theorem 4], and for  $\mathcal{X}^s$ -coefficients, cf. [39, Corollary 1.6], using arguments from these papers.

On a finite time interval, the homogeneous problem (4) can easily be treated by the above results if  $\partial_t \varepsilon \in L^1 L^\infty$ , since a standard energy estimate yields

$$(16) \quad \|u\|_{L^\infty L^2} \lesssim e^{c\|\partial_t \varepsilon\|_{L^1 L^\infty}} \|u(0)\|_{L^2(\mathbb{R}^2)}.$$

As in [38, Corollary 5], we can also prove estimates with two different Strichartz pairs. We provide such a result on a finite time interval  $(0, T)$ , fixing  $\mu$  and  $T$ , and we further suppose that the solutions are charge-free and  $\varepsilon$  is isotropic. These limitations stem from the use of duality in the proof. Perhaps the latter assumptions can be weakened by modifying the proof of Theorem 1.5 to treat inhomogeneous terms  $\langle D' \rangle^{\tilde{\rho}} f \in L^{\tilde{p}'} L^{\tilde{q}'}$ . This is not pursued presently.

**Corollary 1.6.** *Let  $\varepsilon = eI_{2 \times 2} \in C^s$ ,  $1 \leq s \leq 2$ ,  $P(x, D)u = f$ ,  $\partial_1 u_1 + \partial_2 u_2 = 0$ , and  $(\rho, p, q, 2)$ ,  $(\tilde{\rho}, \tilde{p}, \tilde{q}, 2)$  be Strichartz pairs. Then, we find the following estimate to hold:*

$$(17) \quad \|\langle D' \rangle^{-\rho - \frac{\sigma}{2}} u\|_{L^p(0, T; L^q)} \lesssim_{T, \mu} \|u(0)\|_{L^2} + \|\langle D' \rangle^{\tilde{\rho} + \frac{\sigma}{2}} f\|_{L^{\tilde{p}'}(0, T; L^{\tilde{q}'})}.$$

The proof relies on a now standard application of the Christ–Kiselev lemma, [6]. However, as the time-dependent generators of the Maxwell system are not self-adjoint in  $L^2$ , additional considerations are necessary.

To study quasilinear equations, we use a similar result in the context of Theorem 1.4, cf. Corollary 1.7 in [39]. The quantity  $\|\partial_x \varepsilon\|_{L^2 L^\infty}$  can be controlled for coefficients  $\varepsilon = \varepsilon(E)$  arising in a bootstrap argument.

**Corollary 1.7.** *Assume that  $\|\partial_x \varepsilon\|_{L^2 L^\infty} \lesssim 1$  and for some  $\tilde{s} \in [1, 2)$ , suppose that  $\|\varepsilon\|_{\mathcal{X}^{\tilde{s}}} \lesssim 1$ . Let  $(\rho, p, q, 2)$  be a Strichartz pair. Then the solution  $u$  to*

$$\begin{cases} P(x, D)u = f, & \partial_1 u_1 + \partial_2 u_2 = \rho_e, \\ u(0) = u_0 \end{cases}$$

satisfies

$$\begin{aligned} \|\langle D' \rangle^{-\alpha} u\|_{L^p(0, T; L^q)} &\lesssim_T \|u_0\|_{L^2(\mathbb{R}^2)} + \|f\|_{L^1(0, T; L^2)} \\ &\quad + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{p}} \rho_e(0)\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{p}} \partial_t \rho_e\|_{L^1(0, T; L^2)} \end{aligned}$$

for  $\alpha > \rho + \frac{\sigma}{p}$  and  $\sigma = \sigma(\tilde{s}) = \frac{2 - \tilde{s}}{2 + \tilde{s}}$ .

For the proof of Theorem 1.1 we shall make use of the FBI transform, conjugating the problem to phase space. Roughly speaking, we can diagonalize (4) to two non-degenerate half-wave equations and one degenerate half-wave equation. The degenerate component can be estimated using the divergence of  $(D_1, D_2)$  in such a way that wave Strichartz estimates for  $(D_1, D_2, H)$  hold true. During the conjugation procedure, we encounter pseudo-differential operators with rough symbols. For these we give expansions of composites, which resemble the smooth case. However,

we have to be careful with  $L^2$ -estimates, since we cannot spare several derivatives in the spatial variables. In our opinion the diagonalization procedure is the main novelty of the paper. It quantifies the hyperbolic degeneracy of the Maxwell system through the electric charges and allows to recover wave Strichartz estimates in the charge-free case. The method of proof possibly extends to other first-order systems like Dirac equations with variable coefficients (cf. [5]).

The second key ingredient in the proof of Theorems 1.1 and 1.3 is the following result for the half-wave equation, which we derive varying Tataru's arguments and using his results for the wave equation [38, 39]. We write

$$\tilde{\varepsilon}(x) = (\tilde{\varepsilon}^{ij}(x)) = \begin{pmatrix} \varepsilon_{22}(x) & -\varepsilon_{12}(x) \\ -\varepsilon_{21}(x) & \varepsilon_{11}(x) \end{pmatrix}$$

which is  $\varepsilon$  up to determinant, and let  $\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}$  denote these coefficients with Fourier support truncated to frequencies  $\{|\xi| \leq \lambda^{\frac{1}{2}}\}$ .

**Proposition 1.8.** *Let  $\lambda \in 2^{\mathbb{N}_0}$ ,  $\lambda \gg 1$ , and  $n \geq 2$ . Assume  $\varepsilon = \varepsilon^{ij}(x)$  satisfies  $\varepsilon^{ij} \in C^2$ ,  $\|\partial_x^2 \varepsilon\|_{L^\infty} \leq 1$ , and (5). Let  $Q(x, D)$  denote the pseudo-differential operator with symbol*

$$q(x, \xi) = -\xi_0 + (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij}(x) \xi_i \xi_j)^{1/2}.$$

Moreover, let  $u$  decay rapidly outside the unit cube and  $(\rho, p, q, n)$  be a Strichartz pair. Then, we find the estimates

$$(18) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|Q(x, D) S_\lambda u\|_{L^2}$$

to hold with an implicit constant uniform in  $\lambda$ . For Lipschitz coefficients  $\varepsilon^{ij}$  with  $\|\partial_x^2 \varepsilon\|_{L^1 L^\infty} \leq 1$ , we obtain

$$(19) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|Q(x, D) S_\lambda u\|_{L^2}.$$

At last, we apply the Strichartz estimates to the local well-posedness theory of the system

$$(20) \quad \begin{cases} \partial_t u_1 = \partial_2 u_3, & u(0) = u_0 \in H^s(\mathbb{R}^2; \mathbb{R})^3, \\ \partial_t u_2 = -\partial_1 u_3, & \partial_1 u_1 + \partial_2 u_2 = 0, \\ \partial_t u_3 = \partial_2(\varepsilon^{-1}(u) u_1) - \partial_1(\varepsilon^{-1}(u) u_2), \end{cases}$$

where  $\varepsilon^{-1}(u) = \psi(|u_1|^2 + |u_2|^2)$ ,  $\psi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is smooth, monotone increasing, and  $\psi(0) = 1$ . Observe that the Kerr nonlinearity as given in (3) is covered. One can apply our methods also to matrix-valued  $\varepsilon(E)$  under symmetry constraints providing energy bounds. We remark that one can transform (20) into a system of wave equations taking second derivatives in time. Although it might be possible in principle to apply the previously known Strichartz estimates for wave equations, this approach surely finds its limitations when anisotropic material laws are considered.

By local well-posedness, we mean existence, uniqueness, and continuous dependence of the solutions in  $H^s$  locally in time. We refer to the recent lecture notes by Ifrim and Tataru [14] for explaining the notion of local well-posedness for quasilinear equations in detail. Energy methods, neglecting dispersive properties of (20), give local well-posedness for  $s > 2$  as noted above. For the scalar quasilinear wave equation on  $\mathbb{R}^2$ , in [39] Tataru proved local well-posedness in  $H^s$  for  $s > 11/6$ . We establish the analogous result for the Maxwell system on  $\mathbb{R}^2$ .

**Theorem 1.9.** (20) is locally well-posed for  $s > 11/6$ .

Finally, we show that the derivative loss for Strichartz estimates is sharp for permittivity coefficients in  $C^s$  for  $1 \leq s \leq 2$ . For this purpose, we elaborate on the connection with wave equations with rough coefficients as showed in (9) and use the time-independent counterexamples of Smith and Tataru [32]. Hence, although the improvement in Theorem 1.9 seems little over the energy method, it appears to be the limit of proving well-posedness in  $H^s$  with Strichartz estimates for general coefficients  $\partial_x \varepsilon \in L^1 L^\infty$ . It could still be possible to make further improvements by the arguments of Smith and Tataru [32], see also Klainerman–Rodnianski [19], in the context of quasilinear wave equations. In these works was exploited that the metrical tensor solves a quasilinear wave equation itself. We note that on  $\mathbb{R}^3$  and for isotropic material laws as above, our methods should give an improvement of the regularity level in the local wellposedness theory by  $\frac{1}{3}$  from  $s > \frac{5}{2}$  to  $s > \frac{13}{6}$ , in accordance with [39].

*Outline of the paper.* In Section 2 we recall properties of pseudo-differential operators with rough symbols and of the FBI transform. In Section 3 we first localize the functions in space and frequency and then carry out the conjugation procedure, reducing Theorems 1.1 and 1.3 to dyadic estimates for the half-wave equation. In Section 4 we prove these crucial dyadic estimates stated in Proposition 1.8, following the arguments in [38, 39]. In Section 5 we treat weaker Strichartz estimates, assuming less regularity of the coefficients, as formulated in Theorem 1.2, Theorem 1.4, and Corollary 1.7. Here we also sketch the proof of Theorem 1.5. In Section 6 we improve the local well-posedness for quasilinear Maxwell equations as stated in Theorem 1.9. In Section 7 we elaborate on the link to wave equations and show sharpness of the derivative loss for permittivity coefficients in  $C^s$  for  $1 \leq s \leq 2$ .

## 2. PSEUDO-DIFFERENTIAL OPERATORS WITH ROUGH SYMBOLS AND PROPERTIES OF THE FBI TRANSFORM

This section is devoted to preliminaries regarding the encountered pseudo-differential operators and the FBI transform. As in Tataru’s works [37, 38, 39], we make use of the latter to find suitable conjugates of pseudo-differential operators in phase space. In these references, the key application was to use the conjugate of the rough wave operator as weight in phase space to derive Strichartz estimates. In the present paper, rough symbols additionally come up when conjugating Maxwell equations to a diagonal system of scalar half-wave equations. Thus, we have to analyze the  $L^2$ -boundedness and compositions of rough symbols. The monograph [40] contains many results for symbols which are not smooth in the spatial variables. Here we state the results in the form needed in the present context, and we shall revisit some of the arguments as these will be used in later sections.

**2.1. The FBI transform.** We first recall basic facts about the FBI transform (cf. [8, 37]). For  $\lambda \in 2^{\mathbb{Z}}$ , the FBI transform of an integrable function  $f : \mathbb{R}^m \rightarrow \mathbb{C}$  is defined by

$$T_\lambda f(z) = C_m \lambda^{\frac{3m}{4}} \int_{\mathbb{R}^m} e^{-\frac{\lambda}{2}(z-y)^2} f(y) dy, \quad z = x - i\xi \in T^*\mathbb{R}^m \equiv \mathbb{R}^{2m},$$

$$C_m = 2^{-\frac{m}{2}} \pi^{-\frac{3m}{4}}.$$

We have the isometric mapping property  $T_\lambda : L^2(\mathbb{R}^m) \rightarrow L^2_\Phi(T^*\mathbb{R}^m)$ , where  $\Phi(z) = e^{-\lambda\xi^2}$ . It is natural to write  $z = x - i\xi$  since  $T_\lambda f$  is in fact holomorphic. The



connection with the Fourier transform is emphasized by writing

$$T_\lambda f(z) = C_m \lambda^{\frac{3m}{4}} e^{\frac{\lambda}{2}\xi^2} \int_{\mathbb{R}^m} e^{-\frac{\lambda}{2}(x-y)^2} e^{i\lambda\xi \cdot (x-y)} f(y) dy.$$

An inversion formula for the FBI transform is given by the adjoint in  $L^2_{\mathbb{F}}$ :

$$T_\lambda^* F(y) = C_m \lambda^{\frac{3m}{4}} \int_{\mathbb{R}^{2m}} e^{-\frac{\lambda}{2}(\bar{z}-y)^2} \Phi(z) F(z) dx d\xi.$$

We recall the following identities for conjugating symbols with the FBI transform. In the following we consider symbols  $a(x, \xi) \in C_x^s C_c^\infty$  compactly supported in  $\xi$ . More specifically, we shall assume

$$a(x, \xi) = 0 \text{ for } \xi \notin B(0, 2).$$

Let  $a_\lambda(x, \xi) = a(x, \xi/\lambda)$  denote the scaled symbol supported at frequencies  $\lesssim \lambda$ , and  $A_\lambda = A_\lambda(x, D)$  be the corresponding pseudo-differential operator.

As in [37, 38], the idea is to find an ‘approximate conjugate’  $\tilde{A}_\lambda$  of  $A_\lambda$  such that

$$T_\lambda A_\lambda(y, D) \approx \tilde{A}_\lambda T_\lambda.$$

We record the basic identities

$$\begin{aligned} T_\lambda(yf)(z) &= \left(x + \frac{1}{-i\lambda}(\partial_\xi - \lambda\xi)\right) T_\lambda f, \\ T_\lambda\left(\frac{D}{\lambda}f\right)(z) &= \left(\xi + \frac{1}{\lambda}\left(\frac{1}{i}\partial_x - \lambda\xi\right)\right) T_\lambda f, \end{aligned}$$

yielding the formal asymptotics

$$T_\lambda A_\lambda(x, D) \approx \sum_{\alpha, \beta} (\partial_\xi - \lambda\xi)^\alpha \frac{\partial_x^\alpha \partial_\xi^\beta a(x, \xi)}{|\alpha|! |\beta|! (-i\lambda)^{|\alpha|} \lambda^{|\beta|}} \left(\frac{1}{i}\partial_x - \lambda\xi\right)^\beta T_\lambda.$$

We recall error bounds for the truncated approximations

$$\tilde{a}_\lambda^s = \sum_{|\alpha|+|\beta|<s} (\partial_\xi - \lambda\xi)^\alpha \frac{\partial_x^\alpha \partial_\xi^\beta a(x, \xi)}{|\alpha|! |\beta|! (-i\lambda)^{|\alpha|} \lambda^{|\beta|}} \left(\frac{1}{i}\partial_x - \lambda\xi\right)^\beta.$$

For  $s \leq 1$ , we have

$$\tilde{a}_\lambda^s = a,$$

and for  $1 < s \leq 2$ ,

$$(21) \quad \tilde{a}_\lambda^s = a + \frac{1}{-i\lambda} a_x (\partial_\xi - \lambda\xi) + \frac{1}{\lambda} a_\xi \left(\frac{1}{i}\partial_x - \lambda\xi\right) = a + \frac{2}{\lambda} (\bar{\partial}a) (\partial - i\lambda\xi),$$

where  $\partial = \frac{1}{2}(\partial_x + i\partial_\xi)$  and  $\bar{\partial} = \frac{1}{2}(\partial_x - i\partial_\xi)$ . We will not need higher approximations because for coefficients in  $C^2$  the Strichartz estimates for the Euclidean (half-)wave equation hold true, which are known to be optimal (cf. [17]). To prove Theorem 1.1, it will be enough to use the first-order approximation from the previous display.

Consider the remainder

$$(22) \quad R_{\lambda,a}^s = T_\lambda A_\lambda - \tilde{a}_\lambda^s T_\lambda.$$

In [37, 38] the following approximation result was proved.

**Theorem 2.1** ([38, Theorem 5, p. 393]). *Suppose that  $a \in C_x^s C_c^\infty$ . Then,*

$$\begin{aligned} \|R_{\lambda,a}^s\|_{L^2 \rightarrow L^2_{\mathbb{F}}} &\lesssim \lambda^{-s/2}, \\ \|(\partial_\xi - \lambda)R_{\lambda,a}^s\|_{L^2 \rightarrow L^2_{\mathbb{F}}} &\lesssim \lambda^{1/2-s/2}. \end{aligned}$$

To prove our main results, we use the following multiplier theorem for  $T_\lambda$ .

**Proposition 2.2.** *Let  $1 \leq p, q \leq \infty$ ,  $a \in C_x^s C_c^\infty(\mathbb{R}^m \times \mathbb{R}^m)$  with  $a(x, \xi) = 0$  for  $\xi \notin B(0, 2)$ , and*

$$\sup_{x \in \mathbb{R}^m} \sum_{0 \leq |\alpha| \leq m+1} \|D_\xi^\alpha a(x, \cdot)\|_{L_\xi^1} \leq C.$$

Then, we find the following estimate to hold:

$$\|T_\lambda^* a(x, \xi) T_\lambda f\|_{L_{x_0}^p L_{x'}^q} \lesssim C \|f\|_{L_{x_0}^p L_{x'}^q}.$$

*Proof.* We start with the special case  $b \in L_x^\infty(\mathbb{R}^m)$ ,  $c \in C_c^\infty(\mathbb{R}^m)$ ,  $a(x, \xi) = b(x)c(\xi)$ . It can be treated by a straight-forward kernel estimate. We first compute

$$\begin{aligned} & T_\lambda^*(b(x)c(\xi))T_\lambda f(y) \\ &= C_m^2 \lambda^{\frac{3m}{2}} \int_{\mathbb{R}^{2m}} e^{-\frac{\lambda}{2}(\bar{z}-y)^2} e^{-\lambda\xi^2} b(x)c(\xi) \int_{\mathbb{R}^m} e^{-\frac{\lambda}{2}(z-y')^2} f(y') dy' dx d\xi \\ (23) \quad &= C_m^2 \lambda^{\frac{3m}{2}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} e^{-\frac{\lambda}{2}(x+i\xi-y)^2} e^{-\frac{\lambda}{2}(x-i\xi-y')^2} e^{-\lambda\xi^2} b(x)c(\xi) dx d\xi f(y') dy' \\ &= C_m^2 \lambda^{\frac{3m}{2}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{-\frac{\lambda}{2}((x-y)^2+(x-y')^2)} b(x) dx \int_{\mathbb{R}^m} e^{i\lambda\xi \cdot (y-y')} c(\xi) d\xi f(y') dy' \\ &=: C_m^2 \lambda^{\frac{3m}{2}} \int_{\mathbb{R}^m} B_\lambda(y, y') \hat{c}(\lambda(y-y')) f(y') dy'. \end{aligned}$$

Note that

$$|B_\lambda(y, y')| \leq \|b\|_{L^\infty} \int_{\mathbb{R}^m} e^{-\lambda x^2} dx \lesssim \lambda^{-\frac{m}{2}} \|b\|_{L^\infty},$$

and further,

$$|\hat{c}(\lambda(y-y'))| \leq C_N (1 + \lambda|y-y'|)^{-N} \quad \text{for any } N \in \mathbb{N}.$$

Write  $y = (y_1, y_r) \in \mathbb{R} \times \mathbb{R}^{m-1}$  and likewise for  $y'$ . Two successive applications of Young's inequality imply

$$\begin{aligned} & \lambda^m \left\| \int (1 + \lambda|y-y'|)^{-N} f(y') dy' \right\|_{L_{x_0}^p L_{x'}^q}, \\ & \lesssim \lambda^m \left\| \int (1 + \lambda|y_1 - y'_1|)^{-\frac{N}{2}} (1 + \lambda|y_r - y'_r|)^{-\frac{N}{2}} f(y'_1, y'_r) dy'_1 dy'_r \right\|_{L_{x_0}^p L_{x'}^q}, \\ & \lesssim \lambda \int (1 + \lambda|y_1 - y'_1|)^{-\frac{N}{2}} \|f(y'_1, \cdot)\|_{L_{y'_r}^q} dy'_1 \lesssim \|f\|_{L_{x_0}^p L_{x'}^q}, \end{aligned}$$

for large  $N$ . The two estimates for  $B_\lambda$  and  $\hat{c}$  yield

$$\|T_\lambda^*(b(x)c(\xi))T_\lambda f\|_{L_{x_0}^p L_{x'}^q} \lesssim \|b\|_{L^\infty} \|c\|_{C^N} \|f\|_{L_{x_0}^p L_{x'}^q}.$$

We turn to the case of general  $a(x, \xi)$  according to the assumptions. Let  $\beta \in C_c^\infty(\mathbb{R}^m)$  with  $\beta \equiv 1$  for  $\{|\xi| \leq 2\}$  and  $\text{supp}(\beta) \subseteq B(0, 3)$ . The main reduction is an expansion into the rapidly converging Fourier series

$$a(x, \xi) = \beta(\xi) \sum_{k \in \mathbb{Z}^m} e^{ik\xi} \hat{a}_k(x), \quad \hat{a}_k(x) = \int_{[-\pi, \pi]^n} e^{-ik\xi} a(x, \xi) d\xi, \quad \beta_k(\xi) = e^{ik\xi} \beta(\xi)$$

(cf. [40]). Hence,

$$\|T_\lambda^* a(x, \xi) T_\lambda f\|_{L_{x_0}^p L_{x'}^q} \lesssim \sum_{k \in \mathbb{Z}^m} \|T_\lambda^*(\hat{a}_k(x) \beta_k(\xi)) T_\lambda f\|_{L_{x_0}^p L_{x'}^q}.$$

Integration by parts yields

$$|\hat{a}_k(x)| \lesssim_\ell (1 + |k|)^{-\ell} \sum_{0 \leq |\alpha| \leq \ell} \|D_\xi^\alpha a(x, \cdot)\|_{L_\xi^1}.$$

Take  $\ell = m + 1$  so that  $\sum_{k \in \mathbb{Z}^m} (1 + |k|)^{-\ell} \lesssim_m 1$ . In this case, in (23) we estimate the kernel  $B_\lambda(y, y')$  by

$$|B_\lambda(y, y')| \lesssim_m (1 + |k|)^{-\ell} \lambda^{-\frac{m}{2}} \sup_x \sum_{0 \leq |\alpha| \leq \ell} \|D_\xi^\alpha a(x, \cdot)\|_{L_\xi^1} \lesssim_m (1 + |k|)^{-\ell} C \lambda^{-\frac{m}{2}},$$

and find  $\hat{c}(\lambda(y - y')) = \hat{\beta}(k + \lambda(y - y'))$ . Taking absolute values, we infer

$$T_\lambda^*(\hat{a}_k(x)\beta_k(\xi)T_\lambda f)(y) \lesssim C_a(1 + |k|)^{-\ell} \lambda^m \int |\hat{\beta}(k + \lambda(y - y'))f(y')| dy'.$$

Since

$$\left\| \int |\hat{\beta}(k + \lambda(\cdot - y'))||f(y')| dy' \right\|_{L_{x_0}^p L_{x'}^q} = \left\| \int |\hat{\beta}(\lambda(\cdot - y'))||f(y')| dy' \right\|_{L_{x_0}^p L_{x'}^q},$$

we can finish the proof by

$$\sum_{k \in \mathbb{Z}^m} \|T_\lambda^*(\hat{a}_k(x)e^{ik \cdot \xi}\beta(\xi))T_\lambda f\|_{L^p L^q} \lesssim_m \sum_{k \in \mathbb{Z}^m} (1 + |k|)^{-\ell} C \|f\|_{L^p L^q} \lesssim_m C \|f\|_{L^p L^q}.$$

□

**2.2. Properties of rough symbols.** In this subsection compositions of pseudo-differential operators are recalled and their  $L^2$ -boundedness is quantified. The theory for smooth symbols is vast (cf. [13, 33, 40]). For instance, the  $L^p$ -boundedness of symbols  $a \in S_{1,\delta}^0$ ,  $0 \leq \delta < 1$ , is well-known, see [40, Section 0.11]. Here we give a proof which quantifies in particular the  $L^2$ -boundedness of symbols  $a \in C_x^s C_c^\infty$ , see [40, Chapter 2]. The argument is detailed as it becomes important in later sections.

**Lemma 2.3.** *Let  $1 \leq p, q \leq \infty$ , and  $a \in C_x^s C_c^\infty(\mathbb{R}^m \times \mathbb{R}^m)$  with  $a(x, \xi) = 0$  for  $\xi \notin B(0, 2)$ . Suppose that*

$$\sup_{x \in \mathbb{R}^m} \sum_{0 \leq |\alpha| \leq m+1} \|D_\xi^\alpha a(x, \cdot)\|_{L_\xi^1} \leq C.$$

*Then, we find the following estimate to hold:*

$$\|a(x, D)f\|_{L^p L^q} \lesssim C \|f\|_{L^p L^q}.$$

*Proof.* We first consider the special case of separated variables  $a(x, \xi) = b(x)c(\xi)$ . Hölder's and Young's inequality give

$$\begin{aligned} \|a(x, D)f\|_{L^p L^q} &= \|b(x)c(D)f\|_{L^p L^q} \leq \|b\|_{L^\infty(\mathbb{R}^m)} \|c(D)f\|_{L^p L^q} \\ &\leq \|b\|_{L^\infty(\mathbb{R}^m)} \|\check{c}\|_{L^1(\mathbb{R}^m)} \|f\|_{L^p L^q(\mathbb{R}^m)}. \end{aligned}$$

Integrating by parts to estimate  $\|\check{c}\|_{L^1}$  yields  $\|b\|_{L^\infty(\mathbb{R}^m)} \|\check{c}\|_{L^1} \lesssim C$ .

In the general case, we write

$$a(x, D)f = (2\pi)^{-m} \int_{\mathbb{R}^m} e^{ix\xi} a(x, \xi) \hat{f}(\xi) d\xi = (2\pi)^{-m} \int_{\mathbb{R}^m} e^{ix\xi} a(x, \xi) \beta(\xi) \hat{f}(\xi) d\xi$$

for  $\beta \in C_c^\infty$  with support in  $[-\pi, \pi]^m$  and  $\beta(\xi) \equiv 1$  on  $B(0, 2)$ . We expand  $a(x, \xi)\beta(\xi)$  into the Fourier series in  $\xi$

$$a(x, \xi)\beta(\xi) = \beta(\xi) \sum_{k \in \mathbb{Z}^m} a_k(x) e^{ik\xi},$$

with  $a_k(x) = \int_{[-\pi, \pi]^m} e^{-ik\xi} a(x, \xi) d\xi$ .

We shall estimate every single term  $a_k(x)e^{ik\xi}\beta(\xi)$  via the above argument and then sum over  $k$  using decay from the regularity in  $\xi$ . We have

$$\int a_k(x)e^{ik\xi}\beta(\xi)e^{ix\xi}\hat{f}(\xi)d\xi = a_k(x)(\beta(D)f)(x+k).$$

We find  $|a_k(x)| \leq \|a(x, \cdot)\|_{L^1_\xi}$ , and  $\|\beta(D)f(\cdot+k)\|_{L^pL^q} \lesssim \|f\|_{L^pL^q}$  by Young's inequality and translation invariance. This estimate does not decay in  $k$  sufficiently. For decay in  $k$ , we integrate by parts in  $\xi$  obtaining

$$|a_k(x)| \lesssim_\ell (1+|k|)^{-\ell} \sum_{0 \leq |\alpha| \leq \ell} \|D_\xi^\alpha a(x, \cdot)\|_{L^1_\xi}.$$

For  $\ell = m+1$  this is summable in  $k \in \mathbb{Z}^m$ , and we estimate

$$\begin{aligned} \|a(x, D)f\|_{L^pL^q} &\leq \sum_{k \in \mathbb{Z}^m} \|a_k\|_{L^\infty(\mathbb{R}^m)} \|f\|_{L^pL^q} \\ &\lesssim \sum_{k \in \mathbb{Z}^m} (1+|k|)^{-(m+1)} \sup_{x \in \mathbb{R}^m} \left( \sum_{0 \leq |\alpha| \leq m+1} \|D_\xi^\alpha a(x, \cdot)\|_{L^1_\xi} \right) \|f\|_{L^pL^q} \\ &\lesssim C \|f\|_{L^pL^q}. \end{aligned} \quad \square$$

We turn to compound symbols. As we shall see in Section 3, it suffices to prove dyadic estimates<sup>1</sup>

$$\lambda^{-\rho} \|S_\lambda u\|_{L^pL^q} \lesssim \|S_\lambda u\|_{L^2} + \|P^\lambda S_\lambda u\|_{L^2},$$

where  $S_\lambda$  localizes to frequencies of size  $\lambda \in 2^{\mathbb{N}_0}$  and  $P^\lambda$  denotes the operator

$$P^\lambda(x, D) = \begin{pmatrix} \partial_t & 0 & -\partial_2 \\ 0 & \partial_t & \partial_1 \\ -\partial_2(\varepsilon_{11}^{\lambda^{\frac{1}{2}} \cdot}) + \partial_1(\varepsilon_{21}^{\lambda^{\frac{1}{2}} \cdot}) & \partial_1(\varepsilon_{22}^{\lambda^{\frac{1}{2}} \cdot}) - \partial_2(\varepsilon_{12}^{\lambda^{\frac{1}{2}} \cdot}) & \partial_t \end{pmatrix}.$$

For the components of the permittivity,  $\varepsilon_{ij}^{\lambda^{\frac{1}{2}}}$  means that the frequencies are truncated to size at most  $\lambda^{\frac{1}{2}}$ . Recall the symbol classes for  $m \in \mathbb{R}$ ,  $0 \leq \delta < \rho \leq 1$ :

$$S_{\rho, \delta}^m = \{a \in C^\infty(\mathbb{R}^m \times \mathbb{R}^m) : |\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \lesssim \langle \xi \rangle^{m-\rho+|\beta|+\delta|\alpha|}\}.$$

Hence, the pseudo-differential operators we encounter are smooth in  $x$  and the considered symbols are in  $S_{1, \frac{1}{2}}^m$ . Boundedness  $P(x, D) : H^k(\mathbb{R}^d) \rightarrow H^{k-m}(\mathbb{R}^d)$  is proved in [40, Prop. 0.5E]. Using Lemma 2.3, we will show that the estimates are independent of the dyadic frequency  $\lambda$  when considering Littlewood–Paley pieces.

We recall compositions of pseudo-differential operators. Below we denote

$$\partial_x^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_m}^{\alpha_m} \quad \text{and} \quad D_\xi^\alpha = \partial_\xi^\alpha / (i^{|\alpha|}) \quad \text{for } \alpha \in \mathbb{N}_0^m.$$

**Proposition 2.4** ([40, Proposition 0.3C]). *Given  $P(x, \xi) \in OPS_{\rho_1, \delta_1}^{m_1}$ ,  $Q(x, \xi) \in OPS_{\rho_2, \delta_2}^{m_2}$ , suppose that*

$$0 \leq \delta_2 < \rho \leq 1 \text{ with } \rho = \min(\rho_1, \rho_2).$$

*Then,  $(P \circ Q)(x, D) \in OPS_{\rho, \delta}^{m_1+m_2}$  with  $\delta = \max(\delta_1, \delta_2)$ , and  $P(x, D) \circ Q(x, D)$  satisfies the asymptotic expansion*

$$(P \circ Q)(x, D) = \sum_\alpha \frac{1}{\alpha!} (D_\xi^\alpha P \partial_x^\alpha Q)(x, D) + R,$$

<sup>1</sup>Here we suppose that we are in the charge-free case for the sake of simplicity of exposition.

where  $R : \mathcal{S}' \rightarrow C^\infty$  is a smoothing operator.

When applying this formal expansion in Section 3, we can verify with Lemma 2.3 that the operators coming up in the expansion satisfy acceptable  $L^2$ -bounds. We revisit the proof of Theorem 2.4 to find an explicit form of the remainder  $R$ , after truncating the series expansion. We shall derive

$$(24) \quad (P \circ Q)(x, D) = \sum_{|\alpha| \leq N} \frac{1}{\alpha!} (D_\xi^\alpha P \partial_x^\alpha Q)(x, D) + R_N(x, D)$$

with a remainder  $R_N$  for which we can infer  $L^2$ -bounds decaying in  $\lambda$ . We remark that the estimates almost follow from [40, Proposition 0.5E] and Proposition 2.4. However, the encountered symbols have to be appropriately localized in frequency (see the end of the section), and we thus elect to give more details on the expansion and the error bounds.

We turn to the details. The distribution kernel of  $(P \circ Q)(x, D)$  is given by

$$I(x, y) = \int e^{i\langle x-y, \xi \rangle} (P \circ Q)(x, \xi) d\xi.$$

Here and below, the oscillatory integrals are understood in the sense of distributions (cf. [12, Section VII.8]). For this purpose let  $\rho \in C_c^\infty$  with  $\rho \equiv 1$  in a neighbourhood of 0, set  $\rho_\delta(\xi) = \rho(\delta\xi)$ , and read

$$I(x, y) = \lim_{\delta \rightarrow 0} \int \rho_\delta(\xi) e^{i\langle x-y, \xi \rangle} (P \circ Q)(x, \xi) d\xi.$$

Furthermore,

$$\begin{aligned} (P \circ Q)(x, \xi) &= (2\pi)^{-m} \iint e^{i[\langle x-z, \eta \rangle + \langle z-x, \xi \rangle]} P(x, \eta) Q(z, \xi) d\eta dz \\ &= \left(\frac{\lambda}{2\pi}\right)^m \iint e^{i\lambda\langle x-z, \tilde{\eta} - \tilde{\xi} \rangle} P(x, \lambda\tilde{\eta}) Q(z, \lambda\tilde{\xi}) d\tilde{\eta} d\tilde{\xi}, \end{aligned}$$

where  $\eta = \lambda\tilde{\eta}$  and  $\xi = \lambda\tilde{\xi}$ . This integral is regarded as

$$\begin{aligned} (P \circ Q)(x, \xi) &= \\ &= \left(\frac{\lambda}{2\pi}\right)^m \lim_{\delta \rightarrow 0} \iint \rho_\delta(z', \eta') e^{i\frac{\lambda}{2}\langle (z', \eta'), A(z', \eta') \rangle} P(x, \lambda(\eta' + \tilde{\xi})) Q(x - z', \lambda\tilde{\xi}) dz' d\eta', \end{aligned}$$

where

$$A = \begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}.$$

The phase is stationary at  $(z', \eta') = 0$ . We take a smooth cutoff  $\tilde{\rho} = \rho_{\delta'}$  around the origin. The possibility to choose the cutoff size  $\delta' > \delta$  is used later.

The contribution away from the origin is smoothing:

$$\begin{aligned} \tilde{R}(x, \xi) &= \lim_{\delta \rightarrow 0} \iint e^{i\frac{\lambda}{2}\langle (z', \eta'), A(z', \eta') \rangle} (1 - \tilde{\rho}(\eta', z')) \rho_\delta(\eta', z') \\ &\quad \times P(x, \lambda(\eta' + \tilde{\xi})) Q(x - z', \lambda\tilde{\xi}) dz' d\eta'. \end{aligned}$$

Indeed, in Section 4 we shall see that for the expressions coming up in our analysis, we can show bounds  $O_{L^2}(\lambda^{-N})$  for any  $N$ , independent of  $\delta$ .

For the main contribution, we use Taylor's formula for

$$f(\eta', z') = \tilde{\rho}(\eta', z') P(x, \lambda(\eta' + \tilde{\xi})) Q(x - z', \lambda\tilde{\xi})$$

at the origin, namely

$$f(\eta', z') = \sum_{|\alpha| \leq k} \frac{(D^\alpha f)(0)}{\alpha!} (\eta', z')^\alpha + \sum_{|\beta|=k+1} R_\beta(\eta', z') (\eta', z')^\beta.$$

(Here we can omit  $\rho_\delta$  if  $\delta$  is small enough compared to  $\delta'$ .)

We turn to the first expression. If derivatives act on  $\tilde{\rho}(z', \eta')$  and we evaluate at the origin, then the contribution will vanish. For  $\alpha \in \mathbb{N}_0^{2m}$  we write in the following

$$\alpha = \gamma_1 \cup \gamma_2 \text{ with } \alpha = (\gamma_{11}, \dots, \gamma_{1m}, \gamma_{21}, \dots, \gamma_{2m}).$$

We are left with

$$\iint (\partial_{\eta'}^{\gamma_1} P)(x, \lambda(\eta' + \tilde{\xi})) \partial_{z'}^{\gamma_2} Q(x - z', \lambda\tilde{\xi})|_{(\eta', z')=0} \frac{(\eta', z')^\alpha}{\alpha!} e^{i\frac{\lambda}{2} \langle (z', \eta'), A(z', \eta') \rangle} dz' d\eta'.$$

Since (cf. [33, Section 3.1, p. 100])

$$(2\pi)^{-m} \iint (\eta', z')^\alpha e^{i\frac{\lambda}{2} \langle (z', \eta'), A(z', \eta') \rangle} dz' d\eta' = \lambda^{-m} \begin{cases} \lambda^{-|\alpha|} \frac{\gamma_1!}{i^{|\gamma_1|}}, & \gamma_1 = \gamma_2, \\ 0, & \text{else,} \end{cases}$$

the Taylor polynomial yields the asserted asymptotic expansion.

We turn to the Taylor remainder estimate. We use the integral representation

$$R_\beta(\eta', z') = \frac{|\beta|}{\beta!} \int_0^1 (1-t)^{|\beta|-1} \partial^\beta f(t\eta', z') dt.$$

Write  $\partial^\beta = \partial_{\eta'}^{\beta_1} \partial_{z'}^{\beta_2}$ . By choosing  $\delta' \leq \lambda^{-1}$ , derivatives acting on  $\rho_{\delta'} = \tilde{\rho}$  yield additional negative powers in  $\lambda$ , which makes the resulting expressions better behaved. We thus suppose in the following that the derivatives do not act on  $\tilde{\rho}$ . We analyze the expression

$$\begin{aligned} & \iint \tilde{\rho}(t\eta', tz') \lambda^{|\beta_1|} (\partial_{\eta'}^{\beta_1} P)(x, \lambda(t\eta' + \tilde{\xi})) (\partial_{z'}^{\beta_2} Q)(x - tz', \lambda\tilde{\xi}) \\ & \quad \times e^{i\frac{\lambda}{2} \langle (z', \eta'), A(z', \eta') \rangle} (\eta')^{\beta_1} (z')^{\beta_2} dz' d\eta' \\ & = C \iint \tilde{\rho}(t\eta', tz') (\partial_{\eta'}^{\beta_1} P)(x, \lambda(t\eta' + \tilde{\xi})) (\partial_{z'}^{\beta_2} Q)(x - tz', \lambda\tilde{\xi}) (\partial_{z'}^{\beta_1} e^{i\lambda \langle z', \eta' \rangle}) (z')^{\beta_2} d\eta dz'. \end{aligned}$$

Next, we integrate by parts in  $z'$ . The derivatives can act on  $Q$  or  $(z')^{\beta_2}$  or on  $\tilde{\rho}$ . The latter yields lower-order terms as argued above. We have to use the product rule. The derivatives acting on  $z'$  will be denoted with the multiindex  $\beta_{sub} \leq \beta_1$ , which is supposed to be understood componentwise. We can further suppose that  $\beta_{sub} \leq \beta_2$  because the contribution vanishes otherwise. Hence, we can continue the

above display by

$$\begin{aligned}
 (25) \quad &= \sum_{\beta_{sub} \leq \min(\beta_1, \beta_2)} C_\beta \iint \tilde{\rho}(t\eta', tz') (\partial_{\eta'}^{\beta_1} P)(x, \lambda)(t\eta' + \tilde{\xi}) (\partial_{z'}^{\beta_1 + \beta_2 - \beta_{sub}} Q)(x - tz', \lambda \tilde{\xi}) \\
 &\quad \times e^{i\lambda \langle z', \eta' \rangle} (z')^{\beta_2 - \beta_{sub}} d\eta' dz' + l.o.t. \\
 &= \sum_{\beta_{sub} \leq \min(\beta_1, \beta_2)} \tilde{C}_\beta \iint \tilde{\rho}(t\eta', tz') (\partial_{\eta'}^{\beta_1} P)(x, \lambda)(t\eta' + \tilde{\xi}) (\partial_{z'}^{\beta_1 + \beta_2 - \beta_{sub}} Q)(x - tz', \lambda \tilde{\xi}) \\
 &\quad \times \partial_{\eta'}^{\beta_2 - \beta_{sub}} \left( \frac{e^{i\lambda \langle z', \eta' \rangle}}{\lambda^{|\beta_2 - \beta_{sub}|}} \right) d\eta' dz' + l.o.t. \\
 &= \sum_{\beta_{sub} \leq \min(\beta_1, \beta_2)} C'_\beta \iint \tilde{\rho}(t\eta', tz') (\partial_{\eta'}^{\beta_1 + \beta_2 - \beta_{sub}} P)(x, \lambda)(t\eta' + \tilde{\xi}) \\
 &\quad \times (\partial_{z'}^{\beta_1 + \beta_2 - \beta_{sub}} Q)(x - tz', \lambda \tilde{\xi}) e^{i\lambda \langle z', \eta' \rangle} d\eta' dz' + l.o.t.
 \end{aligned}$$

In Section 3,  $L^2$ -bounds for instances of this expression are a consequence of Lemma 2.3, possibly after choosing the cutoff  $\tilde{\rho}$  differently.

We give a first application, which will be useful in Section 3. Let  $(S'_\lambda)_{\lambda \in 2^{\mathbb{N}_0}}$  denote an inhomogeneous Littlewood–Paley decomposition in  $\mathbb{R}^n$ ,<sup>2</sup>  $\tilde{S}'_\lambda$  denote projections with mildly enlarged support, and suppose that  $(\varepsilon^{ij}) \in C^1$  satisfies (5). For  $\lambda \gg 1$  let  $D_\varepsilon$  and  $\frac{1}{D_\varepsilon}$  denote the pseudo-differential operators given by

$$\begin{aligned}
 (D_\varepsilon f)(x') &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x', \xi' \rangle} \|\xi\|_{\varepsilon(x)} \hat{f}(\xi') d\xi', \\
 \frac{1}{D_\varepsilon} f(x') &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x', \xi' \rangle} \frac{1}{\|\xi\|_{\varepsilon(x)}} \hat{f}(\xi') d\xi',
 \end{aligned}$$

where  $\|\xi\|_{\varepsilon(x)} = \|\xi'\|_{\varepsilon(x)} = (\varepsilon_{\leq \lambda^{\frac{1}{2}}}^{ij} \xi_i \xi_j)^{1/2}$  (summing over  $i, j \in \{1, \dots, n\}$ ) and  $\varepsilon_{\leq \lambda^{\frac{1}{2}}}^{ij}$  are the coefficients with smoothly truncated frequencies at  $\lambda^{\frac{1}{2}}$ . In Section 3 we shall see that choosing  $\lambda = \lambda(\varepsilon) \gg 1$  preserves ellipticity of  $(\varepsilon_{\leq \lambda^{\frac{1}{2}}}^{ij})_{i,j}$ .

The following lemma shows that these operators essentially respect frequency localization as a consequence of the asymptotic expansion by Proposition 2.4.

**Lemma 2.5.** *Let  $\lambda, \mu \in 2^{\mathbb{N}}$ ,  $N \in \mathbb{N}$ , and  $1 \ll \min(\lambda, \mu) \ll \max(\lambda, \mu)$ . Then, we find the following estimate to hold:*

$$(26) \quad \|S'_\mu D_\varepsilon S'_\lambda f\|_{L^2} \lesssim_N (\lambda \vee \mu)^{-N} \|\tilde{S}'_\lambda f\|_{L^2},$$

$$(27) \quad \|S'_\mu \frac{1}{D_\varepsilon} S'_\lambda f\|_{L^2} \lesssim_N (\lambda \vee \mu)^{-N} \|\tilde{S}'_\lambda f\|_{L^2}.$$

*Proof.* We shall focus on the first estimate, as the proof of the second is similar. Firstly, suppose that  $1 \ll \lambda \ll \mu$ . We argue that (26) follows from the expansion in the Kohn–Nirenberg theorem. Note that  $D_\varepsilon S'_\lambda$  has the symbol  $\|\xi\|_{\varepsilon(x)} a_\lambda(\xi')$  and  $S'_\mu$  has the symbol  $a_\mu(\xi')$ . This means that all terms in the asymptotic expansion vanish because the supports in  $\xi$  of the two symbols are disjoint. Furthermore, the estimate for the Taylor remainder follows from the representation (25). In fact, any

<sup>2</sup>We refer to Subsection 3.2 for details.

derivative acting on  $P$  gives a factor  $\mu^{-1}$ , whereas derivatives acting on  $Q$  only lose factors  $\lambda^{1/2}$ . We turn to the estimate of the remainder

$$C\lambda^n \iint e^{i\frac{\lambda}{2}\langle(z',\eta'),A(z',\eta')\rangle} \rho_\delta(\eta', z')(1-\rho_{\delta'}(\eta', z')) a\left(\frac{\lambda}{\mu}(\eta'+\tilde{\xi})\right) \lambda \|\tilde{\xi}\|_{\varepsilon(x-z')} a(\tilde{\xi}) dz' d\eta'.$$

We suppose that  $0 < \delta < \delta' \leq \lambda^{-1}$  with  $\delta'$  fixed. The phase  $\frac{\lambda}{2}\langle(z',\eta'),A(z',\eta')\rangle$  is non-stationary away from the origin. Consequently, we can integrate by parts in  $(z',\eta')$ . This gives factors  $(\lambda|(z',\eta')|)^{-1}$  per integration by parts. When a derivative acts on  $\rho_\delta$  or  $\rho_{\delta'}$ , this gives factors of  $\delta$  or  $\delta'$ , respectively, and when it acts on  $a\left(\frac{\lambda}{\mu}(\eta'+\tilde{\xi})\right)$ , this gives factors  $\frac{\lambda}{\mu}$ , which are all favourable. More care is required when derivatives  $\partial_{z'}$  act on  $\|\tilde{\xi}\|_{\varepsilon(x-z')}$ . Since  $\varepsilon^{ij}$  is regularized and we can only estimate  $\|D|\varepsilon^{ij}\|_{L^\infty} \lesssim 1$ , additional derivatives in  $z'$  give powers of  $\lambda^{\frac{1}{2}}$ . We thus obtain sufficient decay to apply Lemma 2.3 and conclude (26).

We turn to the proof for  $1 \ll \mu \ll \lambda$ . Following along the above lines, we estimate the remainder

$$\begin{aligned} C\lambda^n \iint e^{i\frac{\lambda}{2}\langle(z',\eta'),A(z',\eta')\rangle} (1-\rho_{\delta'}(z',\eta')) \rho_\delta(z',\eta') \\ \times \tilde{a}\left(\frac{\lambda}{\mu}(\eta'+\tilde{\xi})\right) \lambda \|\tilde{\xi}\|_{\varepsilon(x-z')} a(\tilde{\xi}) dz' d\eta', \end{aligned}$$

where  $\text{supp}(\tilde{a}) \subseteq B(0,2)$  and  $\tilde{a} \equiv 1$  on  $B(0,1)$ . Integration by parts in  $(z',\eta')$  yields

- powers of  $(|(z',\eta')|\lambda)^{-1}$  from the non-stationary phase,
- powers of  $\frac{\lambda}{\mu}$  from derivatives acting on  $\tilde{a}\left(\frac{\lambda}{\mu}(\eta'+\tilde{\xi})\right)$ ,
- powers of  $\lambda^{\frac{1}{2}}$  from derivatives acting on  $\|\tilde{\xi}\|_{\varepsilon(x-z')}$ .

We observe that due to the support of  $(1-\rho_{\delta'}(z',\eta'))$  that  $|(z',\eta')| \gtrsim (\delta')^{-1}$ . Since  $\delta' \leq \lambda^{-1}$ , every integration by parts gives a factor of  $(\lambda|(z',\eta')|)^{-\frac{1}{2}}$ .

The estimate for the Taylor remainder becomes more involved, too. Still, taking derivatives of  $P$  as in (25) yields

$$(\partial_{\eta'}^{\beta_1+\beta_2-\beta_{sub}} P)(x, \lambda(t\eta'+\tilde{\xi})) = \left(\frac{1}{\mu}\right)^{|\beta_1+\beta_2-\beta_{sub}|} (\partial_{\xi}^{\beta_1+\beta_2-\beta_{sub}} \tilde{a})\left(\frac{\lambda(t\eta'+\tilde{\xi})}{\mu}\right).$$

In the derivatives  $\partial_{z'}^{\beta_1+\beta_2-\beta_{sub}} Q$  we only lose  $\lambda^{\frac{|\beta_1+\beta_2-\beta_{sub}|}{2}}$ . Choosing  $\mu = \lambda/C$  with  $C$  a large, but fixed constant, the proof is complete. Alternatively, one can argue by taking adjoints (cf. [40, Prop. 0.3B]).  $\square$

We end the section with discussing variants, which will be useful later on. Consider the region

$$\{|\xi_0| \lesssim |(\xi_1, \xi_2)| \sim \lambda\} = A_\lambda \subseteq \mathbb{R}^3.$$

Let  $S'_{\lambda,\tau}$  denote the smooth frequency projection to  $A_\lambda$  and  $\tilde{A}_\lambda$  a mildly enlarged region and  $\tilde{S}'_{\lambda,\tau}$  the corresponding frequency projection. Let  $S^{\tau}_{\gg\lambda}$  be the smooth frequency projection to frequencies  $\{|\xi_0| \gg \lambda\}$ . By the same argument as above, we see

$$\|S^{\tau}_{\gg\lambda} \partial_i^k D_\varepsilon \tilde{S}'_{\lambda,\tau}\|_{L^2 \rightarrow L^2} \lesssim_{k,N} \lambda^{-N}.$$

Together with the above estimates, we see that  $D_\varepsilon S'_{\lambda,\tau} f$  and  $\frac{1}{D_\varepsilon} S'_{\lambda,\tau} f$  are still essentially frequency localized in Fourier space in  $A_\lambda$ . More precisely, we find the following estimate to hold:

$$\|(1-\tilde{S}'_{\lambda,\tau}) \partial_i^k D_\varepsilon S'_{\lambda,\tau} f\|_{L^2} \lesssim_{N,k} \lambda^{-N} \|\tilde{S}'_{\lambda,\tau} f\|_{L^2},$$



and likewise for  $\frac{1}{D_\varepsilon}$ .

### 3. REDUCTION TO DYADIC ESTIMATES FOR THE HALF-WAVE EQUATION

In this section, we show that Theorems 1.1 and 1.3 follow from dyadic estimates for the half-wave equation, given in Proposition 1.8. The key point is to diagonalize the principal symbol, which is carried out first. Furthermore, by commutator and microlocal estimates, we localize in phase space to a region close to the characteristic surface. This is crucial to use the ellipticity of  $\varepsilon$ . We require that  $\varepsilon \in C^1$  as assumed in Theorems 1.1 and 1.3. Observe that  $\tilde{\varepsilon}$  inherits the assumptions on  $\varepsilon$  (up to constants) of these theorems.

**3.1. Diagonalizing the principal symbol.** Firstly, we carry out the diagonalization of the principal symbol. To obtain a better approximation, we consider the operators directly. We remark that diagonalizing the symbol combined with the arguments from [37] allows to prove Strichartz estimates for coefficients in  $C^1$  or with derivative in  $L^p L^\infty$ . However, since the estimates obtained in [37] are not sharp in terms of derivative loss, the corresponding estimates for first-order systems proved this way are not sharp either. Nonetheless, this observation can be useful as it saves error estimates for compounds of pseudo-differential operators. The error analysis is carried out in Paragraph 3.3 in the present context. For more complicated first order systems this might not be easily possible. Here, we carry out the detailed computations in our special case

$$\varepsilon(x) = \begin{pmatrix} \varepsilon^{11}(x) & \varepsilon^{12}(x) \\ \varepsilon^{12}(x) & \varepsilon^{22}(x) \end{pmatrix}.$$

For  $P$  as in (6) we find  $P = Op(\tilde{p}(x, \xi))$  with

$$(28) \quad \tilde{p}(x, \xi) = \begin{pmatrix} i\xi_0 & 0 & -i\xi_2 \\ 0 & i\xi_0 & i\xi_1 \\ -i\xi_2\varepsilon_{11}(x) + i\xi_1\varepsilon_{12}(x) & i\xi_1\varepsilon_{22}(x) - i\xi_2\varepsilon_{12}(x) & i\xi_0 \end{pmatrix} \\ + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ (-\partial_2\varepsilon_{11} + \partial_1\varepsilon_{12})(x) & (\partial_1\varepsilon_{22} - \partial_2\varepsilon_{12})(x) & 0 \end{pmatrix},$$

where the first matrix is the principal symbol  $p(x, \xi)$ . As the operator associated with the second matrix is bounded in  $L^2$  for  $\varepsilon \in C^1$ , it will be neglected. Let

$$\tilde{\varepsilon}(x) = (\tilde{\varepsilon}^{ij}(x)) = \begin{pmatrix} \varepsilon_{22}(x) & -\varepsilon_{12}(x) \\ -\varepsilon_{21}(x) & \varepsilon_{11}(x) \end{pmatrix}$$

denote the adjugate matrix of  $\varepsilon^{-1}$ , i.e.,  $\varepsilon$  up to determinant. We compute the eigenvalues of  $p$  to be  $i\xi_0$ ,  $i(\xi_0 - \|\xi'\|_{\tilde{\varepsilon}})$ , and  $i(\xi_0 + \|\xi'\|_{\tilde{\varepsilon}})$ . Denote

$$(29) \quad d(x, \xi) = \text{diag}(i\xi_0, i(\xi_0 - \|\xi'\|_{\tilde{\varepsilon}}), i(\xi_0 + \|\xi'\|_{\tilde{\varepsilon}})),$$

and set  $\xi_j^* = \xi_j / \|\xi'\|_{\tilde{\varepsilon}}$  for  $j \in \{1, 2\}$ . The corresponding eigenvectors we align as

$$(30) \quad m(x, \xi) = \begin{pmatrix} -\xi_1^*\varepsilon_{22}(x) + \xi_2^*\varepsilon_{12}(x) & \xi_2^* & -\xi_2^* \\ \xi_1^*\varepsilon_{12}(x) - \xi_2^*\varepsilon_{11}(x) & -\xi_1^* & \xi_1^* \\ 0 & 1 & 1 \end{pmatrix}.$$

The inverse matrix is computed to

$$(31) \quad m^{-1}(x, \xi) = \begin{pmatrix} -\xi_1^* & -\xi_2^* & 0 \\ \frac{\xi_2^* \varepsilon_{11}(x) - \xi_1^* \varepsilon_{12}(x)}{2} & \frac{-\xi_1^* \varepsilon_{22}(x) + \xi_2^* \varepsilon_{12}(x)}{2} & \frac{1}{2} \\ \frac{-\xi_2^* \varepsilon_{11}(x) + \xi_1^* \varepsilon_{12}(x)}{2} & \frac{\xi_1^* \varepsilon_{22}(x) - \xi_2^* \varepsilon_{12}(x)}{2} & \frac{1}{2} \end{pmatrix},$$

and hence

$$m(x, \xi)d(x, \xi)m^{-1}(x, \xi) = p(x, \xi).$$

By the arguments from [37] and Proposition 2.2 this decomposition gives (non-sharp) Strichartz estimates for  $\varepsilon \in C^1$ . To prove the sharp result for  $\varepsilon \in C^2$ , we diagonalize  $P$  with pseudo-differential operators. Until the end of this subsection, we suppose that  $\varepsilon \in C^\infty$ . After having reduced to dyadic estimates, we shall see that we can truncate frequencies of  $\varepsilon$ . This will allow to work with smooth symbols. Of course, we have to show bounds independent of the dyadic frequency range.

We now turn to the corresponding operators, starting with

$$(32) \quad \mathcal{D}(x, D) = \text{diag}(\partial_t, \partial_t - iD_{\bar{\varepsilon}}, \partial_t + iD_{\bar{\varepsilon}})$$

induced by  $d(x, \xi)$ . To the eigenvectors in  $m$  we associate the operator

$$(33) \quad \mathcal{M}(x, D) = \begin{pmatrix} \frac{i}{D_{\bar{\varepsilon}}}(\partial_1(\varepsilon_{22}\cdot) - \partial_2(\varepsilon_{12}\cdot)) & \frac{-i}{D_{\bar{\varepsilon}}}\partial_2 & \frac{i}{D_{\bar{\varepsilon}}}\partial_2 \\ \frac{i}{D_{\bar{\varepsilon}}}(\partial_2(\varepsilon_{11}\cdot) - \partial_1(\varepsilon_{12}\cdot)) & \frac{i}{D_{\bar{\varepsilon}}}\partial_1 & \frac{-i}{D_{\bar{\varepsilon}}}\partial_1 \\ 0 & 1 & 1 \end{pmatrix}$$

and to the inverse matrix  $m^{-1}$  we relate

$$(34) \quad \mathcal{N}(x, D) = \begin{pmatrix} i\partial_1 \frac{1}{D_{\bar{\varepsilon}}} & i\partial_2 \frac{1}{D_{\bar{\varepsilon}}} & 0 \\ \frac{i(\varepsilon_{12}\partial_1 - \varepsilon_{11}\partial_2)}{2} \frac{1}{D_{\bar{\varepsilon}}} & \frac{i(\varepsilon_{22}\partial_1 - \varepsilon_{12}\partial_2)}{2} \frac{1}{D_{\bar{\varepsilon}}} & \frac{1}{2} \\ \frac{i(\varepsilon_{11}\partial_2 - \varepsilon_{12}\partial_1)}{2} \frac{1}{D_{\bar{\varepsilon}}} & \frac{i(\varepsilon_{12}\partial_2 - \varepsilon_{22}\partial_1)}{2} \frac{1}{D_{\bar{\varepsilon}}} & \frac{1}{2} \end{pmatrix}.$$

Here  $(\varepsilon_{ij}\cdot)$  denotes the multiplication operator induced by  $\varepsilon_{ij}$ . Note that  $\mathcal{M}$  and  $\mathcal{N}$  are bounded in  $L^2$ . We compute

$$(35) \quad \begin{aligned} (\mathcal{M}\mathcal{D}\mathcal{N})_{11} &= -\frac{1}{D_{\bar{\varepsilon}}}[\partial_1(\varepsilon_{22}\cdot)\partial_t\partial_1 - \partial_2(\varepsilon_{12}\cdot)\partial_t\partial_1 + \partial_2\partial_t(\varepsilon_{11}\cdot)\partial_2 - \partial_2\partial_t(\varepsilon_{12}\cdot)\partial_1] \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{12} &= -\frac{1}{D_{\bar{\varepsilon}}}[\partial_1(\varepsilon_{22}\cdot)\partial_t\partial_2 - \partial_2(\varepsilon_{12}\cdot)\partial_t\partial_2 + \partial_2\partial_t(\varepsilon_{12}\cdot)\partial_2 - \partial_2\partial_t(\varepsilon_{22}\cdot)\partial_1] \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{13} &= -\frac{1}{D_{\bar{\varepsilon}}}\partial_2 D_{\bar{\varepsilon}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{21} &= -\frac{1}{D_{\bar{\varepsilon}}}[\partial_2(\varepsilon_{11}\cdot)\partial_t\partial_1 - \partial_1(\varepsilon_{12}\cdot)\partial_t\partial_1 + \partial_1\partial_t(\varepsilon_{12}\cdot)\partial_1 - \partial_1\partial_t(\varepsilon_{11}\cdot)\partial_2] \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{22} &= -\frac{1}{D_{\bar{\varepsilon}}}[\partial_2(\varepsilon_{11}\cdot)\partial_t\partial_2 - \partial_1(\varepsilon_{12}\cdot)\partial_t\partial_2 + \partial_1\partial_t(\varepsilon_{22}\cdot)\partial_1 - \partial_1\partial_t(\varepsilon_{12}\cdot)\partial_2] \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{23} &= \frac{1}{D_{\bar{\varepsilon}}}\partial_1 D_{\bar{\varepsilon}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{31} &= D_{\bar{\varepsilon}}((\varepsilon_{12}\cdot)\partial_1 - (\varepsilon_{11}\cdot)\partial_2) \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{32} &= D_{\bar{\varepsilon}}((\varepsilon_{22}\cdot)\partial_1 - (\varepsilon_{12}\cdot)\partial_2) \frac{1}{D_{\bar{\varepsilon}}}, \\ (\mathcal{M}\mathcal{D}\mathcal{N})_{33} &= \partial_t. \end{aligned}$$

In Subsection 3.3 we shall see that the difference  $\mathcal{MDN} - P$  is bounded in  $L^2$  with suitable frequency localization.

**3.2. Reductions for  $C^2$ -coefficients.** Next, we carry out the reductions for  $u$  and  $\varepsilon_{ij}$  needed for the proof of Theorem 1.1, that is

- localization to a cube of size 1 and to high frequencies,
- reduction to dyadic estimates,
- truncating frequencies of the coefficients,
- reduction to half-wave equations.

Before these steps, by scaling we can assume that  $|\partial_x^2 \varepsilon^{ij}| \leq 1$  and  $\mu = 1$ .

**3.2.1. Localization to a cube of size 1 and to high frequencies.** We first show that it is sufficient to establish (11) for inhomogeneous norms. Let  $s(\xi)$  be a symbol supported in  $\{1/2 \leq |\xi| \leq 2\}$  such that

$$\sum_j s(2^{-j}\xi) = 1, \quad \xi \in \mathbb{R}^3 \setminus \{0\}.$$

For  $\lambda \in 2^{\mathbb{N}_0}$ , let  $S_\lambda = S(D/\lambda)$  denote the Littlewood–Paley multiplier, which localizes to frequencies of size  $\lambda$  and  $S_0 = 1 - \sum_{j \geq 0} S_{2^j}$ . Write  $u = S_0 u + (1 - S_0)u$ .

Sobolev’s embedding and the Hardy–Littlewood–Sobolev inequality yield

$$\| |D|^{-\rho} S_0 u \|_{L^p L^q} \lesssim \| S_0 u \|_{L^2} \lesssim \| u \|_{L^2}.$$

For the contribution of  $(1 - S_0)u$  in (11) we observe that

$$\partial_k(\varepsilon_{ij}(1 - S_0)u) = \partial_k(\varepsilon_{ij}u) - \partial_k(\varepsilon_{ij}S_0u)$$

and

$$\begin{aligned} \|\partial_k(\varepsilon_{ij}S_0u)\|_{L^2} &\lesssim \|\varepsilon_{ij}\partial_k S_0u\|_{L^2} + \|(\partial_k \varepsilon_{ij})S_0u\|_{L^2} \\ &\lesssim \|\varepsilon_{ij}\|_{L^\infty} \|S_0u\|_{L^2} + \|\partial_k \varepsilon_{ij}\|_{L^\infty} \|S_0u\|_{L^2} \\ &\lesssim \|\varepsilon_{ij}\|_{C^1} \|S_0u\|_{L^2}. \end{aligned}$$

Hence,  $\|P(1 - S_0)u\|_{L^2} \lesssim \|Pu\|_{L^2} + \|\varepsilon_{ij}\|_{C^1} \|u\|_{L^2}$ .

This means that low frequencies can always be estimated by the Hardy–Littlewood–Sobolev inequality and Sobolev’s embedding, so that we can assume that  $u$  has only large frequencies. It suffices to prove

$$\|(1 + |D|^2)^{-\frac{\rho}{2}} u\|_{L^p L^q} \lesssim \|u\|_{L^2} + \|Pu\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2}} \rho_e\|_{L^2}.$$

Take a smooth partition of unity

$$1 = \sum_{j \in \mathbb{Z}^{n+1}} \chi_j(x), \quad \chi_j(x) = \chi(x - j), \quad \text{supp}(\chi) \subseteq B(0, 2),$$

and let  $\rho_j = \partial_1(\chi_j u_1) + \partial_2(\chi_j u_2)$ . By considering commutators, we first note

$$\sum_j \left( \|\chi_j u\|_{L^2}^2 + \|P(\chi_j u)\|_{L^2}^2 \right) \lesssim \|u\|_{L^2}^2 + \|Pu\|_{L^2}^2.$$

On the other hand, since  $p, q \geq 2$  we have

$$(36) \quad \|(1 + |D|^2)^{-\frac{\rho}{2}} u\|_{L^p L^q}^2 \lesssim \sum_j \|(1 + |D|^2)^{-\frac{\rho}{2}} \chi_j u\|_{L^p L^q}^2.$$

To prove the latter estimate, we write

$$(1 + |D|^2)^{-\frac{\rho}{2}} u = \sum_{j,k} \chi_j(x) (1 + |D|^2)^{-\frac{\rho}{2}} \chi_k(x) u.$$

If  $|j-k| \geq 100$ , for the distributional kernel of  $\chi_j(1+|D|^2)^{-\frac{\rho}{2}}\chi_k$  we find the bound

$$\begin{aligned} (2\pi)^3 |G(x, y)| &= \left| \int e^{i(x-y, \xi)} \chi_j(x) \chi_k(y) (1+|\xi|^2)^{-\frac{\rho}{2}} d\xi \right| \\ &\lesssim_N (1+|j-k|)^{-N} (1+|x-y|)^{-N} \end{aligned}$$

since  $|x-y| \gtrsim |j-k| \geq 1$ . Hence, we can estimate  $\chi_j(1+|D|^2)^{-\frac{\rho}{2}}\chi_k u$  for  $|j-k| \geq 10n$  using Young's inequality and obtain

$$\|(1+|D|^2)^{-\frac{\rho}{2}} u\|_{L^p L^q}^2 \lesssim \|u\|_{L^2}^2 + \left\| \sum_{j,k:|j-k| \leq 10n} \chi_k (1+|D|^2)^{-\rho/2} \chi_j u \right\|_{L^p L^q}^2.$$

Let  $\tilde{\chi}_j = \sum_{|k-j| \leq 10n} \chi_k$ . Due to the pointwise bound

$$\sum_j \tilde{\chi}_j (1+|D|^2)^{-\rho/2} \chi_j u \lesssim \left( \sum_j |\tilde{\chi}_j (1+|D|^2)^{-\rho/2} \chi_j u|^2 \right)^{1/2},$$

and Minkowski's inequality for  $p, q \geq 2$ , we conclude the proof of (36). It remains to show

$$(37) \quad \sum_j \|\langle D \rangle^{-\frac{1}{2}} \rho_j\|_{L^2}^2 \lesssim \|\langle D \rangle^{-\frac{1}{2}} \rho_e\|_{L^2}^2 + \|u\|_{L^2}^2.$$

This fact is a consequence of the inequality  $\sum_j \|\chi_j v\|_{H^{-1/2}}^2 \leq \|v\|_{H^{-1/2}}^2$  which is dual to

$$\|u\|_{H^{\frac{1}{2}}}^2 \leq \sum_j \|\chi_j u\|_{H^{\frac{1}{2}}}^2.$$

The above estimate follows by interpolation from its elementary variants in  $L^2$  and  $H^1$ .

This concludes the reduction to compact support, and we suppose in the following that  $u$  is supported in the unit cube.

**3.2.2. Reduction to dyadic estimates.** Here we shall see that it is enough to prove

$$(38) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|P S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}$$

for  $\lambda \geq 1$ . To reduce to the above display, we have to show the commutator estimate

$$\sum_{\lambda=2^j \geq 1} \|[P, S_\lambda]u\|_{L^2}^2 \lesssim \|u\|_{L^2}^2.$$

Set  $\tilde{S}_\lambda = \sum_{|j| \leq 2} S_{2^j \lambda}$  as a mildly enlarged version of  $S_\lambda$ , and write  $[P, S_\lambda] = [P, S_\lambda] \tilde{S}_\lambda - S_\lambda P(1 - \tilde{S}_\lambda)$ . We need the inequalities

$$\begin{aligned} \|[P, S_\lambda]v\|_{L^2} &\lesssim \|v\|_{L^2}, \\ \|S_\lambda P(1 - \tilde{S}_\lambda)u\|_{L^2} &\lesssim \lambda^{-\delta} \|u\|_{L^2} \end{aligned}$$

for some  $\delta > 0$ . Since the commutator

$$[\partial_k(\varepsilon_{ij \cdot}), S_\lambda]v = \partial_k(\varepsilon_{ij} S_\lambda v) - S_\lambda \partial_k(\varepsilon_{ij} v) = \partial_k[\varepsilon_{ij}, S_\lambda]v$$

has the kernel  $K(x, y) = \partial_k(\varepsilon_{ij}(x) - \varepsilon_{ij}(y)) \hat{s}(\lambda(x-y)) \lambda^{n+1}$ , the  $L^2$ -boundedness follows from  $\varepsilon \in C^2$ . For the second term note that only frequencies of  $\varepsilon_{ij}$  of size  $\lambda$  and higher matter as

$$S_\lambda \partial_k(\varepsilon_{ij}(1 - \tilde{S}_\lambda)u) = S_\lambda \partial_k(\varepsilon_{ij}^{\geq \lambda}(1 - \tilde{S}_\lambda)u).$$

We thus find

$$\|S_\lambda \partial_k(\varepsilon_{ij}^{\geq \lambda}(1 - \tilde{S}_\lambda)u)\|_{L^2} \lesssim \lambda \|\varepsilon_{ij}^{\geq \lambda}\|_{L^\infty} \|u\|_{L^2} \lesssim \lambda^{-1} \|\varepsilon_{ij}\|_{C^2} \|u\|_{L^2},$$

due to the standard estimate

$$\sup_{\lambda} \lambda^2 \|\varepsilon_{ij}^{\geq \lambda}\|_{L^\infty} \lesssim \|\varepsilon_{ij}\|_{C^2}.$$

Observe that  $S_\lambda u$  is no supported in the unit cube anymore. Instead it is rapidly decreasing, which suffices for the arguments in Section 4, cf. (55).

3.2.3. *Truncating the coefficients of  $P$  at frequency  $\lambda^{\frac{1}{2}}$ .* We check that it is enough to prove (38) when the coefficients have Fourier transform supported in  $\{|\xi| \leq \lambda^{\frac{1}{2}}\}$ . We stress that the uniform ellipticity for  $\varepsilon$  after Fourier truncation will still be crucial. For that purpose, we observe that  $\|\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}}\|_{L^\infty} \lesssim \lambda^{-1} \|\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}}\|_{C^2}$ , and thus

$$\varepsilon_{ij}^{\leq \lambda^{\frac{1}{2}}} \xi_i \xi_j = \varepsilon_{ij} \xi_i \xi_j - \varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}} \xi_i \xi_j \geq (\Lambda_1 - C\lambda^{-1}) \|\xi'\|^2 \geq \frac{1}{2} \Lambda_1 \|\xi'\|^2$$

for sufficiently large  $\lambda \geq 1$ .

The error term coming from frequency truncation in (38) is estimated by

$$\begin{aligned} \|\partial_k(\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}} S_\lambda u)\|_{L^2} &\lesssim \|\partial_k(\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}})\|_{L^\infty} \|S_\lambda u\|_{L^2} + \|\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}}\|_{L^\infty} \|\partial_k S_\lambda u\|_{L^2} \\ &\lesssim \|\varepsilon_{ij}\|_{C^1} \|S_\lambda u\|_{L^2} + \lambda \|\varepsilon_{ij}^{\geq \lambda^{\frac{1}{2}}}\|_{L^\infty} \|S_\lambda u\|_{L^2}, \end{aligned}$$

which is bounded by the first term on the right-hand side of (38). In the following we write  $\varepsilon_{ij}^{\lambda^{\frac{1}{2}}} = \varepsilon_{ij}^{\leq \lambda^{\frac{1}{2}}}$ . Consequently, it is enough to prove

$$(39) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|P^\lambda S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2},$$

with the operator

$$P^\lambda = \begin{pmatrix} \partial_t & 0 & -\partial_2 \\ 0 & \partial_t & \partial_1 \\ -\partial_2(\varepsilon_{11}^{\lambda^{\frac{1}{2}}}\cdot) + \partial_1(\varepsilon_{12}^{\lambda^{\frac{1}{2}}}\cdot) & \partial_1(\varepsilon_{22}^{\lambda^{\frac{1}{2}}}\cdot) - \partial_2(\varepsilon_{12}^{\lambda^{\frac{1}{2}}}\cdot) & \partial_t \end{pmatrix},$$

whose coefficients are truncated at frequency  $\lambda^{\frac{1}{2}}$ . If there is no possibility of confusion, the frequency projection of the coefficients will be implicit in the following, and we write  $P$  instead of  $P^\lambda$ .

3.2.4. *Reduction to half-wave equations.* We shall consider the two refined regions  $\{|\xi_0| \gg |(\xi_1, \xi_2)|\}$  and  $\{|\xi_0| \lesssim |(\xi_1, \xi_2)|\}$ . As noticed above, the first region is away from the characteristic surface. To deal with it, we first apply Theorem 2.1 obtaining

$$\|T_\lambda \left( \frac{P(x, D)}{\lambda} S_\lambda u \right) - p(x, \xi) T_\lambda S_\lambda u\|_{L_{\mathbb{F}}^2} \lesssim \lambda^{-\frac{1}{2}} \|S_\lambda u\|_{L^2}.$$

Suppose that  $u$  has frequencies in  $\{|\xi_0| \gg |(\xi_1, \xi_2)|\}$ . By Theorem 2.1, the same is true for  $v_\lambda = T_\lambda S_\lambda u$  up to an acceptable error  $\lambda^{-\frac{1}{2}} \|S_\lambda u\|_{L^2}$ . For such  $v_\lambda$  it is easy to see that

$$\|v_\lambda\|_{L_{\mathbb{F}}^2} \lesssim \|p(x, \xi) v_\lambda\|_{L_{\mathbb{F}}^2}$$

because

$$p(x, \xi) = m(x, \xi) d(x, \xi) m^{-1}(x, \xi) \quad \text{and} \quad |d_{ii}| \gtrsim |\xi_0| \gtrsim 1.$$

Using also the  $L^2$ -mapping properties of  $T_\lambda$  and the triangle inequality, we deduce

$$\|S_\lambda u\|_{L^2} = \|v_\lambda\|_{L_{\mathbb{F}}^2} \lesssim \lambda^{-\frac{1}{2}} \|S_\lambda u\|_{L^2} + \lambda^{-1} \|P(x, D) S_\lambda u\|_{L^2}.$$

Sobolev's embedding thus yields

$$\lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \lambda^{-\rho} \lambda^{\frac{1}{2} + \rho} \|S_\lambda u\|_{L^2} \lesssim \|S_\lambda u\|_{L^2} + \|P(x, D)S_\lambda u\|_{L^2}.$$

So (39) is true if the frequencies of  $u$  are confined to  $\{|\xi_0| \gg |(\xi_1, \xi_2)|\}$ .

We turn to the main contribution coming from  $\{|\xi_0| \lesssim |(\xi_1, \xi_2)|\}$  where the characteristic surfaces are contained. In the following we assume that the space-time Fourier transform of  $u$  is supported in this region. We first prove

$$\begin{aligned} \lambda^{-\rho} \|S_\lambda w\|_{L^p L^q} &\lesssim \|S_\lambda w\|_{L^2} + \|\mathcal{D}S_\lambda w\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2} \\ &\lesssim \|S_\lambda u\|_{L^2} + \|\mathcal{D}S_\lambda w\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}, \end{aligned}$$

where  $w = \tilde{S}_\lambda \mathcal{N} S_\lambda u$  and as above

$$\mathcal{D} = \begin{pmatrix} i\partial_t & 0 & 0 \\ 0 & i\partial_t - D_{\tilde{\varepsilon}} & 0 \\ 0 & 0 & i\partial_t + D_{\tilde{\varepsilon}} \end{pmatrix}.$$

The estimates of the second and third component of  $S_\lambda w$  are a consequence of Proposition 1.8 to be established in Section 4. Here the charge  $\rho_e$  does not enter. For the first component, as in Subsection 3.3 below one shows that the product  $\partial_1 \frac{1}{D_{\tilde{\varepsilon}}} S_\lambda$  has the symbol  $i\xi_j^* s_\lambda(\xi)$  up to an error bounded by  $c\lambda^{-1/2} \|S_\lambda u\|_{L^2}$ . Combined with Theorem 2.1, we deduce

$$\|T_\lambda w_1 - [m^{-1}(x, \xi) T_\lambda S_\lambda u]_1\|_{L_{\mathbb{F}}^2} \lesssim \lambda^{-\frac{1}{2}} \|S_\lambda u\|_{L^2}.$$

By  $\partial_1 u_1 + \partial_2 u_2 = \rho_e$  and Theorem 2.1, we find

$$\begin{aligned} &\|[m^{-1}(x, \xi) T_\lambda S_\lambda u]_1\|_{L_{\mathbb{F}}^2} \\ &\leq \left\| \frac{1}{D_{\tilde{\varepsilon}}} S_\lambda \rho_e \right\|_{L^2} + \|T_\lambda \left( \frac{1}{D_{\tilde{\varepsilon}}} \partial_1 S_\lambda u_1 + \frac{1}{D_{\tilde{\varepsilon}}} \partial_2 S_\lambda u_2 \right) - i\xi_1^* T_\lambda u_1 - i\xi_2^* T_\lambda u_2\|_{L_{\mathbb{F}}^2} \\ &\lesssim \lambda^{-1} \|S_\lambda \rho_e\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda u\|_{L^2}. \end{aligned}$$

Recall that the ultimate estimate for the first term is a consequence of Lemma 2.3 and the previous frequency localization  $\{|\xi'| \sim |\xi|\}$ .

Hence, using Sobolev's embedding and the  $L^2$ -mapping properties of the FBI-transform, the contribution of  $S_\lambda w_1$  can be estimated by

$$\lambda^{-\rho} \|S_\lambda w_1\|_{L^p L^q} \lesssim \lambda^{\rho + \frac{1}{2} - \rho} \|S_\lambda w_1\|_{L^2} = \lambda^{\frac{1}{2}} \|T_\lambda w_1\|_{L_{\mathbb{F}}^2} \lesssim \|S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2},$$

which gives the estimate for  $w_1$ . So far, we have proved that

$$\lambda^{-\rho} \|\tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|\mathcal{D} \tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}$$

For (38), we yet have to show that

$$(40) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \lambda^{-\rho} \|\tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^p L^q} + \|S_\lambda u\|_{L^2},$$

$$(41) \quad \|\mathcal{D} \tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^2} \lesssim \|\tilde{S}_\lambda \mathcal{M} \tilde{S}_\lambda \mathcal{D} \tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^2} + \|S_\lambda u\|_{L^2},$$

and (42) below. We start with the proof of (41). Let  $\tilde{w} = \mathcal{D} \tilde{S}_\lambda \mathcal{N} S_\lambda u$ . We can as well consider  $\tilde{w} = \tilde{S}_\lambda \mathcal{D} \tilde{S}_\lambda \mathcal{N} S_\lambda u$  as  $\mathcal{D}$  and  $\mathcal{N}$  respect frequency localization up to negligible errors.

Theorem 2.1 and calculations as in the next subsection yield

$$\|T_\lambda^* m(x, \xi) T_\lambda \tilde{S}_\lambda \tilde{w} - \tilde{S}_\lambda \mathcal{M} \tilde{S}_\lambda \tilde{w}\|_{L^2} \lesssim \lambda^{-\frac{1}{2}} \|\tilde{S}_\lambda \tilde{w}\|_{L^2}.$$

By the triangle inequality we infer

$$\begin{aligned}\|\tilde{S}_\lambda \tilde{w}\|_{L^2} &= \|T_\lambda^* m^{-1}(x, \xi) T_\lambda T_\lambda^* m(x, \xi) T_\lambda \tilde{S}_\lambda \tilde{w}\|_{L^2} \\ &\lesssim \|T_\lambda^* m(x, \xi) T_\lambda \tilde{S}_\lambda \tilde{w}\|_{L^2} \\ &\lesssim \|\tilde{S}_\lambda \mathcal{M} \tilde{S}_\lambda \tilde{w}\|_{L^2} + \lambda^{-\frac{1}{2}} \|\tilde{S}_\lambda \tilde{w}\|_{L^2},\end{aligned}$$

where  $\lambda^{-\frac{1}{2}} \|\tilde{S}_\lambda w\|_{L^2}$  can be absorbed into the left hand-side for  $\lambda$  large enough.

For the proof of (40), we write in a similar way

$$\tilde{S}_\lambda \mathcal{N} S_\lambda u = T_\lambda^* m^{-1}(x, \xi) T_\lambda S_\lambda u + E S_\lambda u$$

with  $\|E\|_{L^2 \rightarrow L^2} \lesssim \lambda^{-\frac{1}{2}}$ . Proposition 2.2 and Sobolev's embedding then imply

$$\begin{aligned}\lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} &\lesssim \lambda^{-\rho} \|T_\lambda^* m^{-1}(x, \xi) T_\lambda S_\lambda u\|_{L^p L^q} \\ &\lesssim \lambda^{-\rho} \|\tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^p L^q} + \lambda^{\frac{1}{2}} \|E S_\lambda u\|_{L^2} \\ &\lesssim \lambda^{-\rho} \|\tilde{S}_\lambda \mathcal{N} S_\lambda u\|_{L^p L^q} + \|S_\lambda u\|_{L^2}.\end{aligned}$$

**3.3. Estimates of the error terms.** The purpose of this paragraph is to prove

$$(42) \quad \|\mathcal{M} \mathcal{D} \mathcal{N} S'_{\lambda, \tau} - P S'_{\lambda, \tau}\|_{L^2 \rightarrow L^2} \lesssim 1,$$

where  $S'_{\lambda, \tau}$  was introduced at the end of Section 2 and  $\mathcal{M}$ ,  $\mathcal{D}$ ,  $\mathcal{N}$  in (32)–(34). We show the estimate for  $C^1$ -coefficients  $\tilde{\varepsilon}$  which are Fourier truncated at frequencies  $\lambda^{\frac{1}{2}}$ . This will make the estimate applicable for our proofs of Theorems 1.1 and 1.3. For the sake of brevity, we write  $S'_\lambda$  for  $S'_{\lambda, \tau}$  in the following.

**Proposition 3.1.** *Let  $\varepsilon \in C^1$  and suppose that (5) is satisfied. With the notations from the previous sections, we find (42) to hold.*

*Proof.* We use Lemma 2.5 to include frequency projections between  $\mathcal{M}$ ,  $\mathcal{D}$ , and  $\mathcal{N}$ .<sup>3</sup> We prove (42) componentwise. In detail we shall analyze  $(\mathcal{M} \mathcal{D} \mathcal{N})_{11}$ ,  $(\mathcal{M} \mathcal{D} \mathcal{N})_{13}$ , and  $(\mathcal{M} \mathcal{D} \mathcal{N})_{31}$  as the claim follows for the other components by the same means.

*Estimate for  $(\mathcal{M} \mathcal{D} \mathcal{N})_{11}$ :* Firstly, suppose that  $\tilde{\varepsilon}$  is time-independent. (Later we see that the argument extends to time-dependent  $\tilde{\varepsilon}$ .) We shall show

$$\left\| \frac{1}{D_{\tilde{\varepsilon}}} \tilde{S}'_\lambda (\partial_i (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij}(\cdot) \partial_j) \tilde{S}'_\lambda \frac{1}{D_{\tilde{\varepsilon}}} S'_\lambda - S'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim \lambda^{-1}.$$

The above display implies  $\|(\mathcal{M} \mathcal{D} \mathcal{N})_{11} S'_\lambda - \partial_i S'_\lambda\|_{L^2 \rightarrow L^2} \lesssim 1$ . In the proof come up error terms such that

$$\left\| \frac{1}{D_{\tilde{\varepsilon}}} \tilde{S}'_\lambda ((\partial_1 \tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{11}(\cdot)) \partial_1) \tilde{S}'_\lambda \frac{1}{D_{\tilde{\varepsilon}}} S'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim \frac{1}{\lambda},$$

which are straightforward. Thus, it is enough to analyze

$$-\frac{1}{D_{\tilde{\varepsilon}}} \tilde{S}'_\lambda (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij} \partial_i \partial_j) \tilde{S}'_\lambda \frac{1}{D_{\tilde{\varepsilon}}} S'_\lambda.$$

This we break into three symbols

$$\frac{1}{\|\xi\|_{\tilde{\varepsilon}}} \tilde{a}_\lambda(\xi), \quad (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij} \xi_i \xi_j \tilde{a}_\lambda(\xi)), \quad \frac{1}{\|\xi\|_{\tilde{\varepsilon}}} a_\lambda(\xi).$$

<sup>3</sup>Strictly speaking, we should always enlarge the frequency projection a little bit when applying Lemma 2.5. This is not recorded to lighten the notation.

We compute the expansion of the first and second symbol

$$(43) \quad \frac{1}{\|\xi\|_{\tilde{\varepsilon}}^2} \tilde{\varepsilon}^{ij} \xi_i \xi_j \tilde{a}_\lambda(\xi) + \sum_{1 \leq |\alpha| \leq N'} \frac{1}{\alpha!} (D_\xi^\alpha \frac{1}{\|\xi\|_{\tilde{\varepsilon}}}) \partial_x^\alpha (\tilde{\varepsilon}^{ij} \xi_i \xi_j \tilde{a}_\lambda(\xi)) + R_{N'},$$

with  $\|R_{N'}\|_{L^2 \rightarrow L^2} \lesssim_{N'} \lambda^{-N}$  for some  $N \geq 1$  depending on  $N'$ , cf. (25).

Denote the operators in the expansion by  $E^\alpha$ . We find from collecting powers of  $\lambda$ , namely  $\lambda^{\frac{|\alpha|-1}{2}} \lambda^2$  for factors with derivatives in  $x$  and  $\lambda^{-1-|\alpha|}$  for those with derivatives in  $\xi$ , that

$$\|E^\alpha\|_{L^2 \rightarrow L^2} \lesssim_\alpha \lambda^{\frac{1}{2} - \frac{|\alpha|}{2}}.$$

Consequently, the error is bounded in  $L^2$ . Together with  $\frac{1}{D_\varepsilon} S'_\lambda$  it allows us to estimate the  $L^2$ -operator norm by  $\frac{1}{\lambda}$ , which is acceptable.

The leading-order symbol in (43) is given by  $\|\xi\|_{\tilde{\varepsilon}}^{-1} \tilde{\varepsilon}^{ij} \xi_i \xi_j \tilde{a}_\lambda(\xi)$ . It is homogeneous of degree 1. We calculate the composite with  $\frac{1}{D_\varepsilon} S'_\lambda$ . Like in the previous computation, we can estimate the lower-order terms in  $L^2$  by  $\frac{1}{\lambda}$ . The leading-order term is given by

$$\frac{1}{\|\xi\|_{\tilde{\varepsilon}}^2} \tilde{\varepsilon}^{ij} \xi_i \xi_j \tilde{a}_\lambda(\xi) = \tilde{a}_\lambda(\xi).$$

Consequently,

$$-\frac{1}{D_\varepsilon} \tilde{S}'_\lambda \Delta_{\tilde{\varepsilon}} \tilde{S}'_\lambda \frac{1}{D_\varepsilon} S'_\lambda = S'_\lambda + E_{N'} S'_\lambda \quad \text{with} \quad \|E_{N'}\|_{L^2 \rightarrow L^2} \lesssim \lambda^{-1}.$$

Finally, for time-dependent  $\tilde{\varepsilon}$ , we compute as above

$$\tilde{S}'_\lambda \partial_t \frac{1}{D_\varepsilon} S'_\lambda = \tilde{S}'_\lambda \frac{1}{D_\varepsilon} \partial_t S'_\lambda + O_{L^2}(\lambda^{-1}),$$

and in the same venue,

$$\left\| \frac{1}{D_\varepsilon} \partial_i \tilde{S}'_\lambda \partial_t (\tilde{\varepsilon}^{ij} \xi_j) \partial_j \tilde{S}'_\lambda \frac{1}{D_\varepsilon} S'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim 1.$$

*Estimate for  $(MDN)_{13}$ :* We have to prove

$$\|(MDN)_{13} S'_\lambda - P_{13} S'_\lambda\|_{L^2 \rightarrow L^2} = \left\| -\frac{1}{D_\varepsilon} \partial_2 D_\varepsilon S'_\lambda + \partial_2 S'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim 1.$$

First note that

$$\left\| -\frac{1}{D_\varepsilon} \partial_2 \tilde{S}'_\mu D_\varepsilon \tilde{S}'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim_N (\mu \vee \lambda)^{-N}$$

for  $\mu \ll \lambda$  or  $\mu \gg \lambda$  because of the estimates

$$\|S'_\mu D_\varepsilon S'_\lambda\|_{L^2 \rightarrow L^2} \lesssim_N (\mu \vee \lambda)^{-N} \quad \text{and} \quad \left\| \frac{1}{D_\varepsilon} \partial_i \tilde{S}'_\mu \right\|_{L^2 \rightarrow L^2} \lesssim \mu,$$

see Lemma 2.5. Hence, it is enough to show

$$\left\| -\frac{1}{D_\varepsilon} \partial_2 \tilde{S}'_\lambda D_\varepsilon S'_\lambda + \partial_2 S'_\lambda \right\|_{L^2 \rightarrow L^2} \lesssim 1.$$

The first operator is composed of the two operators with symbols

$$p(x, \xi) = -\frac{1}{\|\xi\|_{\tilde{\varepsilon}}} (i\xi_2) \tilde{a}_\lambda(\xi) \quad \text{and} \quad q(x, \xi) = \|\xi\|_{\tilde{\varepsilon}} a_\lambda(\xi).$$

We find

$$(P \circ Q)(x, D) = Op(-i\xi_2 a_\lambda(\xi)) + \sum_{1 \leq |\alpha| \leq N'} \frac{1}{\alpha!} Op(D_\xi^\alpha p(x, \xi) \partial_x^\alpha q(x, \xi)) + R_{N'}$$



with  $\|R_{N'}\|_{L^2 \rightarrow L^2} \lesssim_{N'} \lambda^{-N}$ . The operators  $E^\alpha$  in the sum over  $\alpha$  can be bounded by collecting powers of  $\lambda$  and exploiting homogeneity, where one obtains  $\lambda^{-|\alpha|}$  from  $D_\xi^\alpha p$  and  $\lambda^{\frac{|\alpha|+1}{2}}$  from  $D_x^\alpha q$ . In total, we find the inequality  $\|E^\alpha\|_{L^2 \rightarrow L^2} \lesssim \lambda^{\frac{1-|\alpha|}{2}}$ , and conclude the  $L^2$ -boundedness of  $\sum_{1 \leq |\alpha| \leq N'} E^\alpha$ . The proof is complete since

$$Op(i\xi_2 a_\lambda(\xi)) = \partial_2 S'_\lambda.$$

*Estimate for  $(\mathcal{MDN})_{31}$ :* Below we show

$$\|[(\mathcal{MDN})_{31} + \partial_2(\varepsilon_{11}^{\lambda^{1/2}} \cdot) - \partial_1(\varepsilon_{12}^{\lambda^{1/2}} \cdot)] S'_\lambda\|_{L^2 \rightarrow L^2} \lesssim 1.$$

First,  $((\partial_2 \varepsilon_{11}) - (\partial_1 \varepsilon_{12}))(\cdot)$  is bounded in  $L^2$ . As above the contribution of  $\tilde{S}'_\mu \frac{1}{D_\varepsilon} \tilde{S}'_\lambda$  can be neglected for  $\lambda \ll \mu$  or  $\mu \ll \lambda$ . It remains to verify

$$\| -D_\varepsilon \tilde{S}'_\lambda \varepsilon_{11}^{\lambda^{1/2}} \partial_2 \tilde{S}'_\lambda \frac{1}{D_\varepsilon} S'_\lambda + \varepsilon_{11}^{\lambda^{1/2}} \partial_2 S'_\lambda \|_{L^2 \rightarrow L^2} \lesssim 1$$

since the terms containing  $\varepsilon_{12}^{\lambda^{1/2}}$  can be estimated in the same venue.

The first operator we perceive as composition of the operators associated with the symbols

$$-\|\xi\|_\varepsilon \tilde{a}_\lambda(\xi), \quad \varepsilon_{11}^{\lambda^{1/2}} (-i\xi_2) \tilde{a}_\lambda(\xi), \quad \frac{1}{\|\xi\|_\varepsilon} a_\lambda(\xi).$$

For the composite of the first and second symbol, we find

$$Op(\varepsilon_{11}^{\lambda^{1/2}} (i\xi_2) \|\xi\|_\varepsilon \tilde{a}_\lambda) + \sum_{1 \leq |\alpha| \leq N'} \frac{1}{\alpha!} Op(D_\xi^\alpha (-\|\xi\|_\varepsilon \tilde{a}_\lambda) \partial_x^\alpha \varepsilon_{11}^{\lambda^{1/2}} (-i\xi_2) \tilde{a}_\lambda(\xi)) + R_{N'}$$

with  $\|R_{N'}\|_{L^2 \rightarrow L^2} \lesssim \lambda^{-N}$ . The operators  $E^\alpha$  satisfy

$$\|E^\alpha\|_{L^2 \rightarrow L^2} \lesssim_\alpha \lambda^{1-|\alpha|} \lambda^{\frac{|\alpha|-1}{2}} \lambda = \lambda^{\frac{3-|\alpha|}{2}}.$$

The error in  $L^2 \rightarrow L^2$  is thus bounded by  $\lambda$ . Since  $\|\frac{1}{D_\varepsilon} \tilde{S}'_\lambda\|_{L^2 \rightarrow L^2} \lesssim \frac{1}{\lambda}$ , this gives an  $L^2$ -bounded contribution.

Next, we compute the composite  $C$  of

$$Op(\varepsilon_{11}^{\lambda^{1/2}} (i\xi_2) \|\xi\|_\varepsilon \tilde{a}_\lambda(\xi)) \quad \text{and} \quad \frac{1}{D_\varepsilon} S'_\lambda.$$

We find

$$C = \varepsilon_{11}^{\lambda^{1/2}} (x) \partial_2 S'_\lambda + \sum_{1 \leq |\alpha| \leq N'} \frac{1}{\alpha!} Op(D_\xi^\alpha (\varepsilon_{11}^{\lambda^{1/2}} i\xi_2 \|\xi\|_\varepsilon \tilde{a}_\lambda(\xi)) \partial_x^\alpha (\frac{1}{\|\xi\|_\varepsilon} a_\lambda(\xi))) + R_{N'},$$

where  $\|R_{N'}\|_{L^2 \rightarrow L^2} \lesssim_{N'} \lambda^{-N}$ . The error estimate

$$\sum_{1 \leq |\alpha| \leq N'} \|E^\alpha\|_{L^2 \rightarrow L^2} \lesssim 1$$

is routine by now. The proof is complete.  $\square$

**3.4. Reductions for  $\partial_x^2 \varepsilon \in L^1 L^\infty$ .** As in Subsection 3.2 we carry out the following steps to reduce Theorem 1.3 to the dyadic estimates

$$(44) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|PS_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}$$

for  $\lambda \gtrsim 1$ , where the Fourier support of  $\tilde{\varepsilon}$  contained in  $\{|\xi| \leq \lambda^{1/2}\}$  and  $u$  essentially supported in the unit cube and its space-time Fourier transform is supported in  $\{|\xi_0| \lesssim |(\xi_1, \xi_2)|\}$ . These steps are

- reduction to the case  $\mu = 1$ ,

- reduction to a cube of size 1 and to large frequencies,
- estimate away from the characteristic surface,
- reduction to dyadic estimates,
- truncating the coefficients at frequency  $\lambda^{\frac{1}{2}}$ .

At this point, we can use the estimate from Section 3.3 to complete the reduction from Theorem 1.3 to Proposition 1.8.

3.4.1. *Reduction to the case  $\mu = 1$ .* For this we can follow the argument from [39, Section 3] closely. We omit the details.

3.4.2. *Reduction to a cube of size 1 and to large frequencies.* If  $\mu = 1$ , then we can choose  $T = 1$  by rescaling. It is enough to show

$$(45) \quad \begin{aligned} \|\langle D' \rangle^{-\rho} u\|_{L^p(0,1;L^q)} &\lesssim \|u\|_{L^\infty L^2} + \|P(x, D)u\|_{L^1 L^2} \\ &\quad + \|\langle D' \rangle^{-\frac{1}{2}} \rho_e(0)\|_{L^2(\mathbb{R}^2)} + \|\langle D' \rangle^{-\frac{1}{2}} \partial_t \rho_e\|_{L^1 L^2} \end{aligned}$$

because the Hardy-Littlewood-Sobolev and Bernstein's inequality yield

$$\|\langle D' \rangle^{-\rho} S'_0 u\|_{L^p(0,1;L^q)} \lesssim \|S'_0 u\|_{L^\infty L^2} \lesssim \|u\|_{L^\infty L^2}$$

for the low frequencies. For high frequencies, inequalities (45) and (13) with  $\mu = 1$  are equivalent. We next reduce (45) to the estimate

$$(46) \quad \|\langle D' \rangle^{-\rho} u\|_{L^p(0,2;L^q)} \lesssim \|u\|_{L^2} + \|P(x, D)u\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2}} \rho_e\|_{L^2}.$$

Indeed, if one applies this inequality to solutions  $u$  of the homogenous problem with a cut-off in time, one obtains

$$\|\langle D' \rangle^{-\rho} u\|_{L^p(0,1;L^q)} \lesssim \|u\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2}} \rho_e\|_{L^2} \lesssim \|u(0)\|_{L^2(\mathbb{R}^2)} + \|\langle D' \rangle^{-\frac{1}{2}} \rho_e(0)\|_{L^2(\mathbb{R}^2)}$$

using also the basic energy estimate. Combined with Duhamel's formula and Minkowski's inequality, this estimate implies (45). Regarding the role of the charge, we observe that for free solutions  $(u_1, u_2, u_3)$  we have  $\partial_t \rho_e = 0$ . However, a free solution emanating from  $Pu(s)$  gives charges  $\partial_t \rho_e$  by (8). Since we use Duhamel's formula in the proof, the additional term  $\|\langle D' \rangle^{-\frac{1}{2}} \partial_t \rho_e\|_{L^1 L^2}$  appears.

Inequality (46) respects the finite speed of propagation and, as in Paragraph 3.2.1, we can decompose  $u$  in components supported in cubes of sidelength 1. These estimates sum up to (46).

3.4.3. *Estimate away from the characteristic surface.* Next, we argue that it is enough to prove the stronger estimate

$$(47) \quad \|\langle D \rangle^{-\rho} u\|_{L^p L^q} \lesssim \|u\|_{L^2} + \|P(x, D)u\|_{L^2} + \|\langle D \rangle^{-\frac{1}{2}} \rho_e\|_{L^2}.$$

Apparently, for  $u$  having space-time Fourier transform in the region  $\{|\tau| \lesssim |\xi'|\}$ , (47) implies (46). But for  $\{|\xi'| \ll |\tau|\}$ ,  $P$  is an elliptic operator with Lipschitz coefficients of order 1, which gains one derivative (cf. Paragraph 3.2.4), and (46) follows from Sobolev's embedding.

3.4.4. *Reduction to a dyadic estimate.* We replace (47) by the stronger estimate

$$(48) \quad \|\langle D \rangle^{-\rho} u\|_{L^p L^q} \lesssim \|u\|_{L^2} + \|P(x, D)u\|_{L^r L^2} + \|\langle D \rangle^{-\frac{1}{2}} \rho_e\|_{L^2}$$

with  $1 < r < 2$ , to use a result from [39]. We now show that (48) follows from

$$(49) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|P(x, D)S_\lambda u\|_{L^r L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}.$$

By Littlewood–Paley theory, here one only has to prove the commutator bound

$$\sum_{\lambda \in 2^{\mathbb{N}_0}} \|[P, S_\lambda]u\|_{L^r L^2}^2 \lesssim \|u\|_{L^2}^2.$$

We rewrite  $P$  in non-divergence form, where the error terms are easily estimated in  $L^2$ . It thus suffices to prove

$$\sum_{\lambda \geq 1} \|[\varepsilon_{ij}, S_\lambda] \partial_j u\|_{L^r L^2}^2 \lesssim \|u\|_{L^2}^2,$$

which is [39, Equ. (3.8)]. At last, similar as in Paragraph 3.4.2 the inequality (49) is replaced by

$$(50) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|PS_\lambda u\|_{L^1 L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}.$$

3.4.5. *Truncating the coefficients at frequencies  $\lambda^{\frac{1}{2}}$ .* Finally, the Fourier coefficients of  $\tilde{\varepsilon}$  are truncated to the region  $\{|\xi| \lesssim \lambda^{1/2}\}$ . For  $\lambda > 0$ , let

$$(51) \quad U_\lambda = \sum_{j \geq 0} S_{2^j \lambda}.$$

We estimate the contribution of  $U_{c\lambda^{1/2}} \varepsilon_{ij} =: \varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}}$  in  $P$ , with  $c \leq 1$ . To pass to the summand  $\|PS_\lambda u\|_{L^1 L^2}$ , we compute

$$\begin{aligned} \|\partial_k(\varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}} S_\lambda u)\|_{L^1 L^2} &\leq \|(\partial_k \varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}}) S_\lambda u\|_{L^1 L^2} + \|\varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}} \partial_k S_\lambda u\|_{L^1 L^2} \\ &\lesssim \|\partial_x \varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}}\|_{L^1 L^\infty} \|S_\lambda u\|_{L^\infty L^2} + \lambda \|\varepsilon_{ij}^{\geq c\lambda^{\frac{1}{2}}}\|_{L^1 L^\infty} \|S_\lambda u\|_{L^\infty L^2} \\ &\lesssim \|S_\lambda u\|_{L^\infty L^2} \end{aligned}$$

by means of the assumptions on  $\varepsilon$ . Hence, (50) is a consequence of

$$(52) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|P^\lambda S_\lambda u\|_{L^1 L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2},$$

where  $P^\lambda$  denotes  $P$  with Fourier-truncated  $\varepsilon$ , where we often drop the superscript.

By  $L^2$ -wellposedness, the energy inequality and the estimate away from the characteristic surface, similar as above we see that it is enough to prove

$$\lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|P^\lambda S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2},$$

which is (44). As in the second part of Paragraph 3.2.4 and using Proposition 3.1, we can now reduce Theorem 1.3 to (19) as stated in Proposition 1.8.

#### 4. PROOF OF THE HALF-WAVE ESTIMATE

This section is devoted to the proof of Proposition 1.8. We follow the strategy of [38, 39] to establish the estimates

$$(53) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|Q(x, D)S_\lambda u\|_{L^2},$$

$$(54) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^\infty L^2} + \|Q(x, D)S_\lambda u\|_{L^2},$$

where

$$Q(x, D) = Op(q(x, \xi)), \quad q(x, \xi) = \xi_0 - (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij} \xi_i \xi_j)^{1/2}.$$

Furthermore, as pointed out in the previous section, we can suppose that  $u$  has space-time Fourier transform in  $\{|\xi_0| \lesssim |(\xi_1, \xi_2)|\}$  and is essentially supported in the unit cube. For estimate (53), we suppose that  $\|\partial_x^2 \tilde{\varepsilon}\|_{L^\infty} \lesssim 1$  and for (54) we suppose that  $\|\partial_x^2 \tilde{\varepsilon}\|_{L^1 L^\infty} \lesssim 1$ . We start with the proof of (53).

#### 4.1. Proof for $C^2$ -coefficients.

4.1.1. *Reduction to a neighborhood of the characteristic surface.* Let  $v_\lambda = T_\lambda S_\lambda u$ . The map  $v_\lambda$  is concentrated in the region

$$U = \{|x| \leq 2, \quad \frac{1}{4} \leq |\xi| \leq 4\}.$$

More precisely, we find

$$(55) \quad \|v_\lambda\|_{L^2_{\mathbb{F}}(U^c)} \lesssim e^{-c\lambda} \|S_\lambda u\|_{L^2},$$

see [38, p. 397]. Hence, it suffices to obtain estimates for  $v_\lambda$  in  $U$ .

For such  $v_\lambda$ , Theorem 2.1 with  $s = 2$  yields

$$(56) \quad v_\lambda \in L^2_{\mathbb{F}}, \quad (\lambda q + 2(\bar{\partial}q)(\partial - i\lambda\xi))v_\lambda \in L^2_{\mathbb{F}}$$

$$(57) \quad \lambda^{-1/2}(\partial_\xi - \lambda\xi)v_\lambda \in L^2_{\mathbb{F}}, \quad \lambda^{-1/2}(\partial_\xi - \lambda\xi)(\lambda q + 2(\bar{\partial}q)(\partial - i\lambda\xi))v_\lambda \in L^2_{\mathbb{F}},$$

cf. [38, Eq. (23), (24)] and (21). In the above display  $f \in L^2_{\mathbb{F}}$  means that  $f$  is uniformly bounded in  $\lambda$  in terms of the right-hand side of (53).

We take the second estimate from (56) and the first from (57) to infer

$$\lambda^{\frac{1}{2}}q v_\lambda \in L^2_{\mathbb{F}},$$

using that  $i(\partial_\xi - \lambda\xi) = \partial - i\lambda\xi$  by holomorphy. Consequently, we can suppose that  $v_\lambda$  is supported in a small neighbourhood of  $K = \{q = 0\}$ . Away from the characteristic set, the gain of  $\lambda^{\frac{1}{2}}$  is enough to conclude the reduced estimate by Sobolev's embedding.

As observed in [38], we can replace  $q$  by a  $C^1$ -multiple of it using the bound

$$q(\partial - i\lambda\xi)v_\lambda \in L^2_{\mathbb{F}},$$

see [38, Eq. (26)]. It is precisely this computation by which in [38] a wave symbol like  $p(x, \xi) = \xi_0^2 - \bar{\varepsilon}^{ij}(x)\xi_i\xi_j$  can be replaced by one of the form

$$q(x, \xi) = \xi_0 - t(x, \xi'),$$

where  $t$  is 1-homogeneous in  $\xi'$ . This is achieved by factorizing  $p$  into the symbols of two half-wave equations and considering separated neighborhoods of  $K \cap U$ . In the present context we can thus adopt the arguments and use several results from [38].

For the sake of completeness, we sketch this reasoning in the following. Since  $v_\lambda$  is holomorphic, estimate (56) yields that

$$\begin{aligned} [i(q_x \partial_\xi - q_\xi \partial_x) + \lambda(q - i\xi \cdot q_x - \xi \cdot q_\xi)]v_\lambda &\in L^2_{\mathbb{F}}, \\ [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - \xi \cdot q_\xi - i\xi \cdot q_x)]v_\lambda &\in L^2_{\mathbb{F}}. \end{aligned}$$

see [38, p. 398]. For  $w = \Phi^{1/2}v_\lambda$  we deduce

$$(58) \quad [(q_x \partial_\xi - q_\xi \partial_x) - i\lambda(q - \xi \cdot q_\xi)]w \in L^2$$

$$(59) \quad [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - i\xi \cdot q_x)]w \in L^2.$$

The first equation is an ODE along the Hamiltonian flow, which is used to derive estimates on the cone; the second equation is an ODE along the gradient curves, which is needed to obtain estimates away from the cone.

4.1.2. *Estimates on the cone.* The first bounds in (56) and (57) translate to  $w$  as

$$w \in L^2, \quad \lambda^{-\frac{1}{2}} \partial_\xi w \in L^2.$$

The trace theorem then implies

$$\lambda^{-\frac{1}{4}} w \in L^2(K \cap U).$$

Starting from (57), we also find

$$\lambda^{-\frac{1}{2}} \partial_\xi [(q_x \partial_\xi - q_\xi \partial_x) - i\lambda(q - q_\xi \cdot \xi)] w \in L^2,$$

Note that the scalar function  $q - q_\xi \cdot \xi$  vanishes for 1-homogeneous  $q$ . Again by the trace theorem we obtain

$$\lambda^{-\frac{1}{4}} H_q w := \lambda^{-\frac{1}{4}} (q_x \partial_\xi - q_\xi \partial_x) w \in L^2(K \cap U)$$

Only these  $L^2$  estimates for  $w$  and  $H_q w$  are used later on, cf. [38, Eq. (30), (31)].

4.1.3. *Estimates away from the cone.* To simplify the analysis, we replace  $q$  by the distance function  $r$  to  $K$  which solves the eikonal equation

$$|\nabla_{x,\xi} r| = 1, \quad r = 0 \text{ in } K.$$

$r$  is a  $C^1$ -function, which yields a  $C^1$ -diffeomorphism  $K \cap U \times (-\varepsilon, \varepsilon) \rightarrow \tilde{U}$ , where  $\tilde{U}$  is a neighborhood of  $K \cap U$ . It is the inverse of

$$(x_0, \xi_0, r) \rightarrow (x, \xi) = (x_0, \xi_0) + r \frac{\nabla_{x,\xi} q(x_0, \xi_0)}{|\nabla_{x,\xi} q(x_0, \xi_0)|},$$

which is a local diffeomorphism from  $K \cap U \times [-\varepsilon, \varepsilon]$  onto a neighborhood of  $K$ , see [38, Eq. (32)].

As seen on p. 400 of [38], the quotient  $r/q$  is Lipschitz so that we can replace  $q$  by  $r$  in (59). We introduce  $C^2$ -coordinates on  $K \cap U$ , denoted by  $\zeta$ . Hence  $(\zeta, r)$  are new coordinates in  $U$  near  $K$ . In these coordinates, equation (59) becomes

$$Jw := [\partial_r + \lambda(r - ir_x(\zeta)(\xi(\zeta) + rr_\xi(\zeta)))] w = f \in L^2.$$

Here we consider  $\xi, x, r_x$  and  $r_\xi$  as functions of  $\zeta \in K \cap U$ , which is partly suppressed below. We split this into the homogeneous and inhomogeneous problems

$$\begin{aligned} Jw_1 &:= f \in L^2, & w_1 &= 0 \text{ on } K, \\ Jw_2 &:= f \in L^2, & w_2 &= w \text{ on } K, \end{aligned}$$

cf. [38, Eq. (34), (35)]. In [38, Eq. (36)] it was shown that

$$\lambda^{\frac{1}{2}} \|w_1\|_{L^2} \lesssim \|f\|_{L^2},$$

and hence  $w_2 \in L^2$ .

Below we use the transformation  $dxd\xi = h(r, \zeta) drd\zeta$ , where  $h$  is strictly positive,  $h(0, \zeta) = 1$  and  $d\zeta$  denotes the Lebesgue measure on  $K \cap U$ . The function  $\tilde{w}_2 = (1 - \frac{1}{h})w_2$  solves  $J\tilde{w}_2 = (\partial_r \frac{1}{h})w_2 \in L^2$  and  $\tilde{w}_2 = 0$  on  $K$ . Hence,  $\tilde{w}_2$  can be estimated as  $w_1$ . Corresponding to  $w = (w_1 + \tilde{w}_2) + \frac{1}{h}w_2$ , we split  $S_\lambda u = u_1 + u_2$ ; i.e.,  $u_2 = T_\lambda^*(\Phi^{\frac{1}{2}} \frac{1}{h} w_2)$ . The above estimate yields  $\|u_1\|_{L^2} = O(\lambda^{-1/2})$ , and thus the claim for  $u_1$  follows from Sobolev's embedding. Passing to  $\frac{1}{h}w_2$  normalizes  $h$  in the following computations.

4.1.4. *Reduction to oscillatory integral estimates.* It remains to analyze  $u_2$ . Using the ODE for  $w_2$ , we can write

$$\begin{aligned} u_2(y) &= c_n \lambda^{\frac{3(n+1)}{4}} \int e^{-\frac{\lambda}{2}(y-x-rr_x-i(\xi+rr_\xi))^2} e^{-\frac{\lambda}{2}(\xi+rr_\xi)^2} e^{-\frac{\lambda}{2}r^2} e^{i\lambda(rr_x\xi+\frac{1}{2}r^2r_xr_\xi)} w(\zeta) dr d\zeta \\ &= c_n \lambda^{-\frac{1}{2}} \lambda^{\frac{3(n+1)}{4}} \int_K e^{i\lambda\xi(y-x)} e^{-\frac{\lambda}{2}\omega_\zeta(y-x)^2} a(\zeta) w(\zeta) d\zeta, \end{aligned}$$

see [38, Eq. (38)], where

$$\omega_\zeta(y-x) = (y-x)^2 - \frac{[(r_x + ir_\xi) \cdot (y-x)]^2}{r_\xi^2 + 2r_x^2 + ir_x \cdot r_\xi}, \quad a(\zeta) = c(r_\xi^2 + 2r_x^2 + ir_x r_\xi)^{-1/2}.$$

As noted on p. 402 of [38], the coefficients of the quadratic form  $\omega$  are continuous in  $x$  and smooth in  $\xi$ , and we have  $\Re\omega > 0$ .

Because of Paragraph 4.1.2, it suffices to show

$$(60) \quad \|V_\lambda w\|_{L^p L^q} \lesssim \lambda^{\rho+1/4} (\|w\|_{L^2(K)} + \|H_q w\|_{L^2(K)})$$

for all  $w$  supported in  $K \cap U$ , where

$$V_\lambda w = \lambda^{\frac{3(n+1)}{4}} \int_K e^{i\lambda\xi(y-x)} e^{-\frac{\lambda}{2}\omega_\zeta(y-x)^2} w(\zeta) d\zeta.$$

The oscillatory integral estimate (60) is proved in Theorem 6 of [38] for symbols  $q$  which are 1-homogeneous in  $\xi'$  and have the form

$$(61) \quad q(x, \xi) = \xi_0 - t(x, \xi')$$

using  $C^1$ -equivalence, which is precisely the present concern (see (63)). We state this theorem. It involves the Hamilton flow  $(x_t, \xi_t)$  for  $q$  starting at  $(x, \xi)$ , i.e.,

$$(62) \quad \begin{aligned} \partial_t x_t &= q_x(x_t, \xi_t), & x(0) &= x, \\ \partial_t \xi_t &= q_\xi(x_t, \xi_t), & \xi(0) &= \xi. \end{aligned}$$

**Proposition 4.1.** *Let  $V_\lambda$  as above and  $b(x, \xi)$  be a smooth compactly supported function, which vanishes near the origin and is 1 in  $\{1/4 \leq |\xi| \leq 4\}$ . Let  $L$  denote the pseudo-differential operator along the Hamilton flow given by*

$$Lw(x, \xi) = \int_0^\infty e^{-t} w(x_t, \xi_t) dt.$$

(One has  $L = (H_q + 1)^{-1}$ .) Then,

$$\|V_\lambda b(x, \xi) L\|_{L^2(K) \rightarrow L^p L^q} \lesssim \lambda^{\rho+1/4}.$$

We will not revisit in detail the technical complex interpolation argument from [38, p. 402–408], which is in fact carried out for phase functions as in (61). The final ingredient to finish the proof is to estimate the kernel  $H^t$  of the operator

$$Z^t = V_\lambda b F^t \delta_{q(x, \xi)=0} a V_\lambda^*,$$

where  $F^t$  denotes the translation by  $t$  along the Hamilton flow. Such kernel bounds are provided by Theorem 7 of [38] which we recall here.

**Proposition 4.2.** *The kernels  $H^t$  satisfy*

$$|H^t(y, \tilde{y})| \lesssim \lambda^{n+1} e^{-c\lambda(\tilde{y}_0 - y_0 - t)^2} (1 + \lambda|y - \tilde{y}|)^{-\frac{n-1}{2}}.$$

To derive this result, the Hamilton flow has to be analyzed.

4.1.5. *The regularity of the Hamilton flow.* We revisit the proof of the regularity of the Hamilton flow. We note that the Fourier truncation yields

$$|\partial_x^\alpha \tilde{\varepsilon}| \leq c_\alpha \lambda^{\frac{|\alpha|-2}{2}}, \quad |\alpha| \geq 2.$$

A word of clarification regarding the following estimates: Seemingly, the half wave symbol

$$(63) \quad q(x, \xi) = \xi_0 - (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij}(x) \xi_i \xi_j)^{1/2}$$

is less regular than the wave symbol

$$p(x, \xi) = \xi_0^2 - \tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij}(x) \xi_i \xi_j.$$

For the following estimates, however, it suffices to estimate the Hamilton flow in a small neighborhood  $N$  of  $K \cap U$  for times  $t \in [0, 1]$ . In this region we have  $|\xi'| \sim 1$  and  $|\xi_0| \lesssim 1$ . It follows that in this range of  $(\xi_0, \xi')$  the Hamilton flow for  $H_q$  satisfies the same estimates as for  $H_p$ . This is again reflected through the  $C^1$ -equivalence of  $p$  and  $q$ . In fact,

$$\|H_p w\|_{L^2(K \cap U)} \approx \|H_q w\|_{L^2(K \cap U)}.$$

We record Lemmas 9 and 10 of [38] that estimate the flow and give an expansion.

**Lemma 4.3.** *In  $N$ , the Hamilton flow generated by  $q$  from (63) satisfies*

$$\begin{aligned} |\partial_\xi^\alpha x_t| &\leq c_\alpha t (1 + t\sqrt{\lambda})^{|\alpha|-1}, \quad |\alpha| \geq 1, \\ |\partial_\xi^\alpha \xi_t| &\leq c_\alpha (1 + t\sqrt{\lambda})^{|\alpha|-1}, \quad |\alpha| \geq 1. \end{aligned}$$

**Lemma 4.4.** *For the Hamilton flow in  $N$  we have the representation*

$$\begin{aligned} x_t &= x + tq_\xi + t^2 g(t, x, \xi), \\ \xi_t &= \xi + th(t, x, \xi), \end{aligned}$$

where  $g$  and  $h$  are bounded by

$$|\partial_\xi^\alpha h(t, x, \xi)|, |\partial_\xi^\alpha g(t, x, \xi)| \leq c_\alpha (1 + t\sqrt{\lambda})^{|\alpha|-1}, \quad |\alpha| \geq 1.$$

By the above, the kernel estimate in Proposition 4.2 can be shown by (non)-stationary phase arguments. For the details we refer to [38, p. 412-415]. This finishes the proof of (53) and thus of Theorem 1.1.

4.2. **Proof for  $L^1 L^\infty$  - coefficients.** Next, we show that the estimates remain true if the assumption  $\varepsilon \in C^2$  is replaced by  $\partial_x^2 \varepsilon \in L^1 L^\infty$ . We start with error estimates for conjugating with the FBI transform taken from Theorem 2.3 of [39].

**Theorem 4.5.** *Let  $\lambda \gtrsim 1$  and  $\partial_x^2 \varepsilon \in L^1 L^\infty$ . Set  $a(x, \xi) = (\tilde{\varepsilon}_{\lambda^{1/2}}^{ij}(x) \xi_i \xi_j)^{1/2} s_\lambda(\xi)$ . Then, we find the following estimates to hold:*

$$\begin{aligned} \|\Phi^{1/2} R_{\lambda, a}\|_{L^\infty L^2 \rightarrow L^2} &\lesssim \lambda^{-3/4}, \\ \|\Phi^{1/2} R_{\lambda, a}\|_{L^\infty L^2 \rightarrow L_{x_0}^1 L_{x', \xi'}^2(K)} &\lesssim \lambda^{-3/4}. \end{aligned}$$

(On  $K$  one uses its surface measure  $d\zeta$ .) We turn to the proof of (54) with the notation given there. We again use the FBI transform and write

$$w = \Phi^{1/2} T_\lambda S_\lambda u, \quad S_\lambda u = T_\lambda^* \Phi^{1/2} w.$$

As in the previous subsection,  $w$  is essentially supported in  $U$ , see (55). Recalling (21) and (22), we have

$$(\lambda q + 2(\bar{\partial} q)(\partial - i\lambda\xi)) \Phi^{-1/2} w = \Phi^{-1/2} g$$

with

$$g = \Phi^{1/2}(\lambda R_{\lambda,q} S_\lambda u + T_\lambda Q(x, D) S_\lambda u).$$

The error estimates from Theorem 4.5 allow us to bound  $g$  in terms of the right-hand side of (54). This is stated in the next result which is Lemma 3.1 of [39].

**Lemma 4.6.** *With the above notation we have*

$$\begin{aligned} \|g\|_{L^2} &\lesssim \lambda^{1/4} \|S_\lambda u\|_{L^\infty L^2} + \|Q(x, D) S_\lambda u\|_{L^2}, \\ \|g|_K\|_{L^1_{x_0} L^2_{x', \xi'}} &\lesssim \lambda^{1/4} \|S_\lambda u\|_{L^\infty L^2} + \|Q(x, D) S_\lambda u\|_{L^2}. \end{aligned}$$

4.2.1. *Gradient and Hamilton flow equations.* As in (58) and (59), we find the equations

$$(64) \quad [(q_x \partial_\xi - q_\xi \partial_x) - i\lambda(q - \xi \cdot q_\xi)]w = -ig,$$

$$(65) \quad [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - i\xi \cdot q_x)]w = g,$$

see [39, Eq. (3.15), (3.16)]. Again, the first equation is an ODE along the Hamilton flow of  $q$ , while the second is an ODE along the gradient curves of  $q$ . Moreover, (64) is used to derive estimates on the cone  $K$ , and (65) off the cone. The  $L^1 L^\infty$ -bound on  $\partial_x^2 \varepsilon$  is needed for the analysis of the Hamilton flow. We first sketch the estimate away from  $K$ , which works for Lipschitz coefficients.

Taking the inner product with  $qw$  of (65) and integrating by parts, we obtain

$$\lambda \|qw\|_2^2 = \frac{1}{2} \langle (|\nabla q|^2 + q\Delta q)w, w \rangle + \Re \langle qw, g \rangle.$$

The boundedness of  $U$  and  $\nabla q$  and the inequality  $|\Delta q| \lesssim \lambda^{1/2}$  yield

$$\lambda^{1/2} \|qw\|_2 \lesssim \|w\|_2 + \lambda^{-1/2} \|g\|_2.$$

Away from the cone, i.e.,  $|q| \geq c$ , this inequality gains half of a derivative allowing us to prove Strichartz estimates by Sobolev's embedding as before. In the following we thus suppose that  $w$  and  $g$  are supported in a neighborhood  $N$  of  $K \cap U$ .

4.2.2. *Gradient flow decomposition.* We use the gradient flow equation to decompose  $w$  into two parts,

$$w = w_1 + w_2,$$

where  $w_j$  solves the inhomogeneous and the homogenous equations

$$(66) \quad [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - iq_x \cdot \xi)]w_1 = -ig, \quad w_1|_K = 0,$$

$$(67) \quad [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - iq_x \cdot \xi)]w_2 = 0, \quad w_2|_K = w.$$

We split  $S_\lambda u$  correspondingly, i.e.,

$$u_i = T_\lambda^* \Phi^{1/2} b(x, \xi) w_i,$$

where  $b$  is a smooth cut-off for  $N$ . The estimate for  $w_1$  follows as on p. 431 of [39].

To estimate  $w_2$ , we have to analyze the regularity of the gradient flow. Let  $(x, \xi)$  be initial data on  $K \cap U = \{q = 0\} \cap U$ . The flow  $(x_t, \xi_t)$  is given by (62). Due to the support of  $w$ , it is enough to analyze the regularity of the gradient flow in  $U$ . Here we have the same estimates as in Theorem 3.2 in [39].

**Theorem 4.7.** *For  $q(x, \xi)$  given by (61), the following estimates hold on  $U$ :*

$$\begin{aligned} |\partial_x^\alpha \partial_\xi^\beta x_t| &\leq c_{\alpha, \beta} \lambda^{\frac{|\alpha| - 1}{2}} e^{c_{\alpha, \beta} \sqrt{\lambda} |t|}, \quad |\alpha| + |\beta| > 0, \\ |\partial_x^\alpha \partial_\xi^\beta (\xi_t - \xi)| &\leq c_{\alpha, \beta} \lambda^{\frac{|\alpha| - 1}{2}} e^{c_{\alpha, \beta} \sqrt{\lambda} |t|}. \end{aligned}$$



The proof from [39] for the symbol  $p$  works also for  $q$  since in  $U$  the derivatives of  $q$  and  $p$  satisfy the same bounds.

4.2.3. *Reduction to oscillatory integral estimate.* In the next result  $u_2$  is expressed in terms of trace of  $w$  on the cone.

**Proposition 4.8.** *Assume  $q$  is of the form (61). Then, we have*

$$u_2 = \lambda^{-1/2} V_\lambda w|_K,$$

where  $V_\lambda$  is the integral operator

$$V_\lambda w = \lambda^{\frac{3(n+1)}{4}} \int_K e^{i\lambda\xi(x-y)} G(x, y, \xi) w dx d\xi$$

with  $G$  satisfying

$$|\partial_x^\alpha \partial_\xi^\beta G(x, y, \xi)| \leq c_{\alpha, \beta} \lambda^{\frac{|\alpha|}{2}} e^{-c\lambda(x-y)^2}.$$

The proof of the corresponding Theorem 3.3 in [39] is based on the estimates from the previous lemma. Hence, the argument can be transferred to the half-wave symbol, as long as the estimates for the gradient flow apply. This is guaranteed by the support condition of  $w_2$ . The estimate for  $u_2$  is thus reduced to the oscillatory integral bound

$$\|V_\lambda w\|_{L^p L^q} \lesssim \lambda^{\rho+1/4} \|H_q w\|_{L_{x_0}^1 L_{x', \xi'}^2(K)}$$

for  $w$  supported in  $K \cap U$ .

Such an estimate is proved in Theorem 3.4 of [39]. The analysis from [39] applies due to the Lipschitz equivalence of  $p$  and  $q$  on  $U$ . We record this result.

**Theorem 4.9.** *Let  $a(x, \xi)$  be a smooth compactly supported function, which is 0 near  $\xi = 0$  and 1 in  $\{1/4 \leq |\xi'| \leq 4, |\xi_0| \leq 4\}$ . Then,*

$$\|V_\lambda a(x, \xi) L\|_{L^2(K \cap \{x_0=0\}) \rightarrow L^p L^q} \lesssim \lambda^{\rho+1/4},$$

where  $L$  is the transport operator along the Hamilton flow given by

$$(Lw)(x, \xi) = \begin{cases} 0, & \text{if } x_0 < 0, \\ w(x_t, \xi_t), & \text{if } x_{t0} = 0, x_0 \geq 0. \end{cases}$$

The above estimate is a consequence of the regularity of the Hamilton flow. As proved in [39, p. 434–436], the estimates for  $C^2$  coefficients stated in Lemma 4.3 and 4.4 remain valid. As pointed out above, these estimates are a consequence of estimates for the derivatives, which are equivalent for the wave and half-wave symbol in a suitable neighbourhood of the cone away from the origin. This finishes the proof of (54) and thus of Theorem 1.3.

## 5. ESTIMATES FOR LESS REGULAR COEFFICIENTS

In this section we first prove the weaker estimates stated in Theorems 1.2 and 1.4 for  $\varepsilon$  without a second derivative. Since the results are clear for  $s = 0$  by Sobolov's embedding, we assume that  $s \in (0, 2)$ .

*Proof of Theorem 1.2.* We follow the argument from [38]. By rescaling we can assume that  $\|\varepsilon^{ij}\|_{\dot{C}^s} \leq 1$  and  $\mu = 1$ . Note this transfers to  $\tilde{\varepsilon}$ , up to a constant. As in Paragraph 3.2.1, the low frequencies are estimated by Sobolev's embedding. Here we need that  $\|\partial_k(\varepsilon_{ij} S_0 u)\|_{H^{-\sigma}} \lesssim \|u\|_{L^2}$ , which is true since  $1 - \sigma < s$ . Hence we will suppose that the Fourier transform of  $u$  vanishes near the origin. We can then localize  $u$  to a cube of sidelength 1. Indeed, take a smooth partition of unity

$(\chi_j)_{j \in \mathbb{Z}^3}$ , localizing to cubes with sidelength 1. Compared to Paragraph 3.2.1, the inequality

$$\begin{aligned} & \sum_j (\|\chi_j u\|_{L^2}^2 + \|P(x, D)(\chi_j u)\|_{H^{-\sigma}}^2 + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{2}} \rho_j\|_{L^2}^2) \\ & \lesssim \|u\|_{L^2}^2 + \|P(x, D)u\|_{H^{-\sigma}}^2 + \|\langle D \rangle^{-\frac{1}{2} - \frac{\sigma}{2}} \rho_e\|_{L^2}^2 \end{aligned}$$

needs a bit more care. The estimate for the first term is clear. The summands in the second look like

$$\partial_i(\varepsilon_{kl} \chi_j u) = \chi_j \partial_i(\varepsilon_{kl} u) + (\partial_i \chi_j) \varepsilon_{kl} u.$$

Here the second term on the right-hand side can be estimated in  $L^2$ . For the first one we note that

$$\sum_j \|\chi_j v\|_{H^{-\sigma}}^2 \lesssim \|v\|_{H^{-\sigma}}^2$$

follows by duality and interpolation from  $\|\sum_j \chi_j u_j\|_{H^k}^2 \lesssim \sum_j \|u_j\|_{H^k}^2$  for  $k \in \{0, 1\}$ . The third term can be estimated as previously.

We next claim that (12) is a consequence of the dyadic estimate

$$(68) \quad \lambda^{-\rho} \lambda^{-\frac{\sigma}{2}} \|S_\lambda u\|_{L^p L^q} \lesssim \|S_\lambda u\|_{L^2} + \|P^\lambda S_\lambda u\|_{H^{-\sigma}} + \lambda^{-\frac{1}{2} - \frac{\sigma}{2}} \|S_\lambda \rho_e\|_{L^2},$$

where  $P^\lambda$  is the operator  $P$  after Fourier-truncation of  $\varepsilon_{ij}$  at  $\nu = \lambda^{\frac{2}{2+s}}$ . To recover (12) from the above display, it suffices to establish

$$\sum_{\lambda \in 2^{\mathbb{N}_0}} \|(P^\lambda S_\lambda - S_\lambda P)u\|_{H^{-\sigma}}^2 \lesssim \|u\|_{L^2}^2.$$

Observe that also the frequencies of  $P^\lambda S_\lambda u$  are localized to  $\lambda$ , since the coefficients are truncated at  $\nu \ll \lambda$ . Taking out  $\partial_k$ , we thus have to show

$$\sum_{\lambda \in 2^{\mathbb{N}_0}} \lambda^{2-2\sigma} \|(\varepsilon_{ij}^{\lesssim \nu} S_\lambda - S_\lambda \varepsilon_{ij})u\|_{L^2}^2 \lesssim \|u\|_{L^2}^2.$$

which follows from

$$\begin{aligned} & \|[\varepsilon_{ij}^{\lesssim \nu}, S_\lambda]v\|_{L^2} \lesssim \lambda^{\sigma-1-\delta} \|v\|_{L^2}, \\ & \sum_{\lambda \in 2^{\mathbb{N}_0}} \lambda^{2-2\sigma} \|S_\lambda(\varepsilon_{ij}^{\lesssim \nu} v)\|_{L^2}^2 \lesssim \|v\|_{L^2}^2. \end{aligned}$$

for some  $\delta > 0$ . These estimates can be proved as on p. 417 of [38] using  $1 - \sigma < s$ .

Finally, since the frequencies of  $P^\lambda S_\lambda u$  are located at  $\lambda$ , inequality (68) becomes

$$\lambda^{-\rho} \|S_\lambda u\|_{L^p L^q} \lesssim \lambda^{\frac{\sigma}{2}} \|S_\lambda u\|_{L^2} + \lambda^{-\frac{\sigma}{2}} \|P^\lambda S_\lambda u\|_{L^2} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2}.$$

This is a special case of (11) with  $\mu := \lambda^{\sigma/2}$ . Because the Fourier transform of the coefficients  $\varepsilon_{ij}^{\lesssim \nu}$  of  $P^\lambda$  are supported in  $\{|\xi| \leq \nu = \lambda^{\frac{2}{2+s}}\}$ , we find

$$\|\partial_x^2 \varepsilon_{ij}^{\lesssim \nu}\|_{L^\infty} \lesssim \|\partial_x^2 \varepsilon_{ij}^{\lesssim \nu}\|_{L^\infty} \lesssim \nu^{2-s} \|\varepsilon_{ij}\|_{\dot{C}^s} \lesssim \lambda^{2\sigma} \|\varepsilon^{ij}\|_{\dot{C}^s} \lesssim \lambda^{2\sigma}.$$

(Note that  $(\varepsilon_{ij}^{\lesssim \nu}) := (\varepsilon_{ij}^{\lesssim \nu})^{-1}$  enters the statement of Theorem 1.1 for  $P^\lambda$ .) Therefore, Theorem 1.1 with  $\mu = \lambda^{\sigma/2}$  yields (68), and Theorem 1.2 is proven.  $\square$

Next, we turn to the proof of Theorem 1.4.

*Proof of Theorem 1.4.* To show Theorem 1.4, we note that the proof of Theorem 1.3 yields moreover the estimate

$$\begin{aligned} \| |D|^{-\rho} u \|_{L^p(0,T;L^q)} &\lesssim \mu^{\frac{1}{p}} \|u\|_{L^\infty L^2} + \mu^{-\frac{1}{p'}} \|P(x,D)u\|_{L^1 L^2} \\ &\quad + T^{\frac{1}{2}} \left( \sum_{\lambda \geq 1} \lambda^{-1} \|S_\lambda \rho_\varepsilon\|_{L^\infty L^2}^2 + \sum_{\lambda \geq 1} \lambda^{-1} \|S_\lambda \partial_t \rho_\varepsilon\|_{L^1 L^2}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

By rescaling we can suppose that  $T = 1$  and  $\|\varepsilon\|_{\mathcal{X}^s}^2 \leq \mu^{2+s}$ . We have to transfer the condition on  $\varepsilon$  to  $\tilde{\varepsilon}$ . For this we characterize  $\mathcal{X}^s$  by differences as for usual Besov spaces, cf. Theorems 2.36 and 2.37 of [2]. Then one can proceed as for  $\varepsilon \in C^s$ . After reducing the claim to dyadic estimate and truncating the coefficients at frequency  $\mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}}$ , we will infer the estimates from Theorem 1.3.

We start with a first frequency localization. The low frequencies  $\lambda \leq \max\{1, \mu^{-\frac{1}{\sigma}}\}$  are estimated by Sobolev's embedding as above. For  $\lambda \geq \max\{1, \mu^{-\frac{1}{\sigma}}\}$ , we truncate  $\varepsilon_{ij}$  at frequencies  $\lambda/16$  and let  $P^\lambda$  denote  $P$  with these coefficients  $\varepsilon_{ij}^\lambda$ . We claim that it suffices to prove the dyadic estimates

$$(69) \quad \begin{aligned} \lambda^{-\rho - \frac{\sigma}{p}} \|S_\lambda u\|_{L^p L^q} &\lesssim \mu^{\frac{1}{p}} \|S_\lambda u\|_{L^\infty L^2} + \lambda^{-\sigma} \mu^{-\frac{1}{p'}} \|P^\lambda S_\lambda u\|_{L^1 L^2} \\ &\quad + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S_\lambda \rho_\varepsilon\|_{L^\infty L^2} + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S_\lambda \partial_t \rho_\varepsilon\|_{L^1 L^2}. \end{aligned}$$

Indeed, this inequality implies (14) provided that

$$\| |D|^{-\sigma} (S_\lambda P - P^\lambda S_\lambda) u \|_{L^1 L^2}^2 \lesssim \mu^2 \|\tilde{S}_\lambda u\|_{L^\infty L^2}^2.$$

holds. Factoring out the derivatives, we thus have to show

$$(70) \quad \|(S_\lambda \varepsilon_{ij} - \varepsilon_{ij}^\lambda S_\lambda) u\|_{L^1 L^2}^2 \lesssim \mu^2 \lambda^{2(\sigma-1)} \|\tilde{S}_\lambda u\|_{L^\infty L^2}^2.$$

Because  $u$  is compactly supported in time, it suffices to take compactly supported  $\varepsilon_{ij}$ . Then, we know that

$$\|\varepsilon_{ij}\|_{L^1 L^\infty} \lesssim 1, \quad \|\varepsilon_{ij}\|_{\mathcal{X}^s} \lesssim \mu^{\frac{2+s}{2}}$$

(using also the boundedness of  $\varepsilon_{ij}$ ), which gives

$$\|S_\lambda \varepsilon_{ij}\|_{L^1 L^\infty} \lesssim \min(\mu^{\frac{2+s}{2}} \lambda^{-s}, 1).$$

As  $1 - \sigma = 2s/(2+s)$ , we infer

$$\begin{aligned} \|\varepsilon_{ij}\|_{\dot{B}_{1-\sigma}^{1-\infty}} &= \sum_{\lambda} \lambda^{1-\sigma} \|S_\lambda \varepsilon_{ij}\|_{L^1 L^\infty} \\ &\lesssim \sum_{\lambda \lesssim \mu^{\frac{2+s}{2s}}} \lambda^{1-\sigma} + \sum_{\lambda \gtrsim \mu^{\frac{2+s}{2s}}} \lambda^{1-\sigma} \mu^{\frac{2+s}{2}} \lambda^{-s} \lesssim \mu. \end{aligned}$$

At this point, (70) follows from inequality (4.4) of [39].

Next, we further restrict the Fourier support of  $\varepsilon_{ij}^\lambda$  to  $\mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}}$ . This is possible since we can control the error in the second term in (69) by means of

$$\begin{aligned} \|\partial_x (U_{\mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}}} \varepsilon_{ij}^\lambda) S_\lambda u\|_{L^1 L^2} &\lesssim \lambda \|U_{\mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}}} \varepsilon_{ij}\|_{L^1 L^\infty} \|S_\lambda u\|_{L^\infty L^2} \\ &\lesssim \lambda (\mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}})^{-s} \|\varepsilon_{ij}\|_{\mathcal{X}^s} \|S_\lambda u\|_{L^\infty L^2} \\ &\lesssim \mu \lambda^\sigma \|S_\lambda u\|_{L^\infty L^2}, \end{aligned}$$

where the ultimate inequality follows from the  $\mathcal{X}^s$ -bound for  $\varepsilon$  and  $U_\nu$  was defined in (51). To apply Theorem 1.3, we also note that

$$\|\partial_x^2(S_{\lesssim \mu^{\frac{1}{2}} \lambda^{\frac{2}{2+s}} \varepsilon_{ij}^\lambda})\|_{L^1 L^\infty} \lesssim \mu^{\frac{2-s}{2}} \lambda^{\frac{2(2-s)}{2+s}} \|\varepsilon_{ij}\|_{\mathcal{X}^s} \lesssim (\mu \lambda^\sigma)^2,$$

where  $\mu \lambda^\sigma \geq 1$ . Now, (69) is a consequence of Theorem 1.3.  $\square$

Finally, we discuss how the inhomogeneous estimates given in Theorem 1.5 can be deduced from Theorem 1.1 and the analysis in [38].

*Proof of Theorem 1.5.* By rescaling it is enough to consider the case

$$\mu = 1, \quad \|\partial_x^2 \varepsilon\|_\infty \leq 1.$$

Then we can modify the arguments in the proof of Theorem 1.1 to reduce the desired inequality to  $u$  with support in a cube of size 1, to localize it to a dyadic estimate at frequency  $\lambda \gtrsim 1$  and for the region  $\{|\xi_0| \lesssim |\xi'|\}$ , and to truncate the coefficients at frequency  $\lambda^{\frac{1}{2}}$ . Since the diagonalization  $P = \mathcal{M}\mathcal{D}\mathcal{N}$  essentially respects frequency localization and  $L^p$ -properties, we finally reduce to the inequality

$$(71) \quad \lambda^{-\rho} \|S_\lambda u\|_{L^p L^{q'}} \lesssim \|S_\lambda u\|_{L^2} + \|\tilde{f}_1^\lambda\|_{L^2} + \lambda^\rho \|\tilde{f}_2^\lambda\|_{L^{p'} L^{q'}} + \lambda^{-\frac{1}{2}} \|S_\lambda \rho_e\|_{L^2},$$

where  $\mathcal{D}(x, D)S_\lambda u = \tilde{f}_1^\lambda + \tilde{f}_2^\lambda$ . The degenerate component of  $\mathcal{D}$  is handled as in Paragraph 3.2.4. We sketch the reduction to the above display. Using Proposition 3.1, we can write up to error terms of order  $O_{L^2}(\lambda^{-N})$ :<sup>4</sup>

$$PS_\lambda = \mathcal{M}\mathcal{D}\mathcal{N}S_\lambda + E_\lambda$$

with  $\|E_\lambda\|_{L^2 \rightarrow L^2} \lesssim 1$ . Let  $v$  denote the original function and  $S_\lambda \mathcal{N}S_\lambda v = \tilde{S}_\lambda u$ . We let  $\mathcal{M}S_\lambda \mathcal{D}S_\lambda \mathcal{N}S_\lambda v = f_1^\lambda + f_2^\lambda + g^\lambda$  with  $\|g^\lambda\|_{L^2} \lesssim \|S_\lambda v\|_{L^2}$ . This requires

$$\mathcal{M}\tilde{f}_1^\lambda = f_1^\lambda + g^\lambda, \quad \mathcal{M}\tilde{f}_2^\lambda = f_2^\lambda.$$

Multiplication with  $\mathcal{N}$  gives (cf. Proposition 2.4)

$$\tilde{f}_1^\lambda + E_1^\lambda \tilde{f}_1^\lambda = \mathcal{N}f_1^\lambda + \mathcal{N}g^\lambda, \quad \tilde{f}_2^\lambda + E_2^\lambda \tilde{f}_2^\lambda = \mathcal{N}f_2^\lambda,$$

where we have

$$\|E_1^\lambda\|_{L^2 \rightarrow L^2} \lesssim \lambda^{-1} \quad \text{and} \quad \|E_2^\lambda\|_{L^{p'} L^{q'} \rightarrow L^{p'} L^{q'}} \lesssim \lambda^{-1}$$

by Lemma 2.3. Hence, the equations for  $\tilde{f}_1^\lambda$  and  $\tilde{f}_2^\lambda$  can be solved in  $L^2$  and  $L^{p'} L^{q'}$ , respectively, by the Neumann series, yielding the estimates

$$\|\tilde{f}_1^\lambda\|_{L^2} \lesssim \|f_1^\lambda\|_{L^2} + \|S_\lambda v\|_{L^2}, \quad \|\tilde{f}_2^\lambda\|_{L^{p'} L^{q'}} \lesssim \|f_2^\lambda\|_{L^{p'} L^{q'}}.$$

We abbreviate  $\tilde{f}_i = \tilde{f}_i^\lambda$  in the following.

As before, let

$$v_\lambda = T_\lambda S_\lambda u_i, \quad i = 2, 3, \quad \text{and} \quad q(x, \xi) = \xi_0 - (\tilde{\varepsilon}_{\lambda^{\frac{1}{2}}}^{ij}(x) \xi_i \xi_j)^{1/2}.$$

An application of Theorem 2.1 shows that (cf. [38, p. 418])

$$v_\lambda \in L_{\Phi}^2, \quad [\lambda q + 2(\bar{\partial}q)(\partial - i\lambda\xi)]v_\lambda - T_\lambda \tilde{f}_2 \in L_{\Phi}^2.$$

Now set

$$w = \Phi^{1/2} v_\lambda, \quad g = \Phi^{1/2} T_\lambda \tilde{f}_2.$$

<sup>4</sup>We suppress frequency localization and slightly enlarged variants between the pseudo-differential operators to lighten the notation.

The properties of the ODE along the Hamilton flow yield

$$\lambda^{-1/4}w \in L^2(K \cap U), \quad \lambda^{-1/4}(H_q w - g) \in L^2(K \cap U).$$

As above, we use the ODE along the gradient flow of  $q$  to decompose  $w = w_1 + w_2$ , where

$$\begin{aligned} [(q_x \partial_\xi + q_\xi \partial_x) + \lambda(q - \xi \cdot q_\xi)]w_1 &= f, & w_1|_K &= 0, \\ [(q_x \partial_x + q_\xi \partial_\xi) + \lambda(q - i\xi \cdot q_x)]w_2 &= 0, & w_2|_K &= w. \end{aligned}$$

The function  $w_2$  can further be decomposed into “good” and “bad” parts,

$$w_2 = w_2^g + w_2^b,$$

whose traces on the cone  $K \cap U$  satisfy

$$\lambda^{-1/4}H_q w_2^g \in L^2, \quad H_q w_2^b = g.$$

The map  $w_2^g$  can be treated as  $w_2$  in Paragraph 4.1.4. In a similar way  $w_2^b$  can be controlled by  $g$ , and thus  $\tilde{f}_2$ , as indicated on p. 419 of [38]. Analogously, we decompose  $w_1$  as

$$w_1 = w_1^g + w_1^b,$$

and solve

$$(q_x \partial_x + q_\xi \partial_\xi + \lambda(q - iq_x \cdot \xi))w_1^g \in L^2, \quad w_1^g|_K = 0,$$

respectively,

$$(q_x \partial_x + q_\xi \partial_\xi + \lambda(q - iq_x \cdot \xi))w_1^b = g, \quad w_1^b|_K = 0.$$

Again,  $w_1^g$  can be treated as in Paragraph 4.1.4. For the estimate of  $w_1^b$  we refer to [38, pp. 419–422]. The argument applies due to the Lipschitz equivalence of wave and half-wave symbol in the phase space region of interest, which was already discussed above.  $\square$

We next point out how homogeneous Strichartz estimates yield inhomogeneous Strichartz estimates. For this we invoke the following consequence of the Christ–Kiselev lemma (cf. [6]).

**Lemma 5.1** ([11, Lemma 8.1]). *Let  $X$  and  $Y$  be Banach spaces and for all  $s, t \in \mathbb{R}$  let  $K(s, t) : X \rightarrow Y$  be an operator-valued kernel from  $X$  to  $Y$ . Suppose we have the estimate*

$$\left\| \int_{\mathbb{R}} K(s, t) f(s) ds \right\|_{L^q(\mathbb{R}, Y)} \leq A \|f\|_{L^p(\mathbb{R}, X)}$$

for some  $A > 0$  and  $1 \leq p < q \leq \infty$ , and  $f \in L^p(\mathbb{R}; X)$ . Then, we have

$$\left\| \int_{s < t} K(s, t) f(s) ds \right\|_{L^q(\mathbb{R}, Y)} \leq C_{p,q} A \|f\|_{L^p(\mathbb{R}, X)}.$$

We are ready for the proof of Corollary 1.6.

*Proof of Corollary 1.6.* By well-posedness in  $L^2$  for  $C^1$ -coefficients, let  $(U(t, s))_{t, s \in \mathbb{R}}$  denote the propagator of  $P$  in  $L^2$ .  $U(t, s)u_0$  denotes the solution at time  $t$  to

$$\begin{cases} Pu = 0, \\ u(s) = u_0 \in L^2. \end{cases}$$

We let  $Pu(t) = (\partial_t 1_{3 \times 3} - A(t))u(t)$  so that  $A(t)$  denotes the time-dependent generator. The full solution is given by Duhamel's formula

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)f(s)ds.$$

Let  $\mathcal{T} : L^2 \rightarrow L^p([0, T], L^q)$ ,  $\mathcal{T}f = \langle D' \rangle^{-\rho - \frac{\sigma}{2}} U(t, 0)f$ . The estimates from Theorem 1.2 applied to homogeneous solutions for  $\varepsilon \in C^s$ ,  $1 \leq s \leq 2$ , yield

$$\|\mathcal{T}u_0\|_{L^p(0, T; L^q)} \lesssim_\mu \|U(t, 0)u_0\|_{L^2} \lesssim_{T, \|\varepsilon\|_{C^1}} \|u_0\|_{L^2}.$$

Indeed, the full derivatives can be replaced by the purely spatial derivatives as a consequence of microlocal estimates since the coefficients of  $P$  belong to  $C^1$ , and the energy estimate holds true because  $\|\partial_x \varepsilon\|_{L^1 L^\infty} \lesssim_T \|\varepsilon\|_{C^1}$ . We aim to estimate the Duhamel term

$$\langle D' \rangle^{-\rho - \frac{\sigma}{2}} \int_0^t U(t, s)f(s)ds = \langle D' \rangle^{-\rho - \frac{\sigma}{2}} U(t, 0) \int_0^t U(0, s)f(s)ds.$$

By invoking Lemma 5.1, it suffices to estimate

$$(72) \quad \|\langle D' \rangle^{-\rho - \frac{\sigma}{2}} U(t, 0) \int_0^T U(0, s)f(s)ds\|_{L^p L^q} \lesssim \|\langle D' \rangle^{\tilde{\rho} + \frac{\sigma}{2}} f\|_{L^{\tilde{p}'} L^{\tilde{q}'}}.$$

To prove boundedness of (72), we use the  $\mathcal{T}\mathcal{T}^*$ -argument with duality with respect to  $L^2$ . Note that

$$\mathcal{T}^* F = \int_0^T U(s, 0)^* (\langle D' \rangle^{-\rho - \frac{\sigma}{2}} f(s)) ds.$$

Hence, to derive estimates for  $\int_0^T U(0, s)f(s)ds$ , it suffices to show

$$\|\langle D' \rangle^{-\tilde{\rho} - \frac{\sigma}{2}} U(0, t)^* g\|_{L^{\tilde{p}}(0, T; L^{\tilde{q}})} \lesssim \|g\|_{L^2(\mathbb{R}^2)}.$$

Since  $L^2$ -duality respects divergence-free vector fields,  $(u_1, u_2)$  is divergence-free, and so is  $(f_1, f_2)$ , we can suppose that  $(g_1, g_2)$  is divergence-free.

We compute the time-dependent generator for  $U(0, t)^*$  by

$$\partial_t \langle U(0, t)^* u, v \rangle = \partial_t \langle u, U(0, t)v \rangle = \langle u, -U(0, t)A(t)v \rangle = -\langle A(t)^* U(0, t)^* u, v \rangle$$

in the general case  $\varepsilon = (\varepsilon^{ij})_{i,j=1,2}$ . We thus find for the full operator

$$P^* = \begin{pmatrix} -\partial_t & 0 & \varepsilon_{11}\partial_2 - \varepsilon_{12}\partial_1 \\ 0 & -\partial_t & -\varepsilon_{22}\partial_1 + \varepsilon_{12}\partial_2 \\ \partial_2 & -\partial_1 & -\partial_t \end{pmatrix} = -P^t + E$$

with  $\|E\|_{L^2 \rightarrow L^2} \lesssim 1$  for Lipschitz coefficients. Hence, it suffices to prove estimates for  $P^t$ . These are a consequence of the proof of Theorem 1.2. The only difference happens in the diagonalization: Transposing (29)-(31) yields

$$(m^{-1})^t(x, \xi)d(x, \xi)m^t(x, \xi) = p^t(x, \xi).$$

Note

$$m^t(x, \xi) = \begin{pmatrix} -\xi_1^* \varepsilon_{22}(x) + \xi_2^* \varepsilon_{12}(x) & \xi_1^* \varepsilon_{12}(x) - \xi_2^* \varepsilon_{11}(x) & 0 \\ \xi_2^* & -\xi_1^* & 1 \\ -\xi_2^* & \xi_1^* & 1 \end{pmatrix},$$

and analogously for the operators in (32)-(34). The estimate for  $P^t$  corresponding of Theorem 1.2 is

$$\|\langle D' \rangle^{-\tilde{\rho} - \frac{\sigma}{2}} u\|_{L^{\tilde{p}} L^{\tilde{q}}} \lesssim_{T, \mu} \|u\|_{L^2} + \|P^t u\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{2}} \rho_e^*\|_{L^2}$$

with  $\rho_e^* = \nabla \cdot (\tilde{\varepsilon}^{-1} \tilde{u})$ ,  $\tilde{u} = (u_1, u_2)$ , and the adjugate  $\tilde{\varepsilon}^{-1} = \varepsilon \cdot \det(\varepsilon^{-1})$  of  $\varepsilon^{-1}$ .

Consequently, for  $U(0, t)^* u_0$  we find

$$\begin{aligned} \|\langle D' \rangle^{-\tilde{\rho}-\frac{\sigma}{2}} U(0, t)^* u_0\|_{L^{\tilde{p}} L^{\tilde{q}}} &\lesssim \|U(0, t)^* u_0\|_{L^2} + \|(P^* - E)U(0, t)^* u_0\|_{L^2} \\ &\quad + \|\langle D' \rangle^{-\frac{1}{2}-\frac{\sigma}{2}} \nabla \cdot (\tilde{\varepsilon}^{-1} U(0, t)^* u_0)\|_{L^2} \\ &\lesssim \|u_0\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2}-\frac{\sigma}{2}} \rho_e^*\|_{L^2}, \end{aligned}$$

which follows from the energy estimate in  $L^2$ . For low frequencies, the energy estimate also allow to dominate the second term in the last line by the first. For high frequencies, we can take advantage that  $\tilde{u}_0$  is divergence free, that  $\rho_e^*(t) = \nabla \cdot (\tilde{\varepsilon}^{-1} \tilde{u}_0)$ , and that  $\varepsilon$  is isotropic. It follows

$$\|S'_{\geq 1}(\langle D' \rangle^{-\frac{1}{2}-\frac{\sigma}{2}} \nabla \cdot (\tilde{\varepsilon}^{-1} \tilde{u}))\|_{L^2} \lesssim_{\|\varepsilon\|_{C^1}} \|u(0)\|_{L^2}.$$

We have proved for divergence-free  $(f_1, f_2)$  that

$$\|\langle D' \rangle^{-\rho-\frac{\sigma}{2}} \int_0^t U(t, s) f(s) ds\|_{L^p(0, T; L^q)} \lesssim \|\langle D' \rangle^{\tilde{\rho}+\frac{\sigma}{2}} f\|_{L^{\tilde{p}}(0, T; L^{\tilde{q}})},$$

where  $(\rho, p, q, 2)$ ,  $(\tilde{\rho}, \tilde{p}, \tilde{q}, 2)$  are Strichartz pairs. Note that the case  $p = \tilde{p} = 2$  cannot be covered by Lemma 5.1, but  $p = 2$  is not wave-admissible in two dimensions anyway. We finish the proof by Duhamel's formula and the triangle inequality.  $\square$

In the following we derive Strichartz estimates only with spatial derivatives for  $\|\partial_x \varepsilon\|_{L^2 L^\infty} \lesssim 1$ . This will become useful in the quasilinear case where we shall control  $\|\nabla_{x'} u\|_{L^4 L^\infty}$  and have  $\|\partial_x(\varepsilon(u))\|_{L^2 L^\infty} \lesssim_{\|u\|_{L^\infty}} \|\nabla_{x'} u\|_{L^2 L^\infty} \lesssim_T \|\nabla_{x'} u\|_{L^4 L^\infty}$ . Also note that

$$\|\partial_x \varepsilon\|_{L^2 L^\infty} \lesssim \|\varepsilon\|_{\mathcal{X}^s}$$

by Sobolev's embedding for  $s > \frac{3}{2}$ . Actually, this condition is only needed in our proof of the microlocal estimate, for which we use the following commutator estimate for the FBI transform, also due to Tataru [37]. Let  $X = L^2 L^\infty$  and  $X^1 = \{u \in X : \partial_x u \in X\}$ .

**Theorem 5.2** ([37, Theorem 2]). *Assume that  $a \in X^1 C_c^\infty$ . Then,*

$$\|R_{\lambda, a}\|_{L^\infty L^2 \rightarrow L^2_{\mathbb{F}}} = \|T_\lambda A_\lambda - a T_\lambda\|_{L^\infty L^2 \rightarrow L^2_{\mathbb{F}}} \lesssim_{\|a\|_{X^1 C_c^\infty}} \lambda^{-\frac{1}{2}}.$$

We can now show Corollary 1.7.

*Proof of Corollary 1.7.* Let  $S_\lambda^{\ll \tau}$  denote the smooth frequency localization to regions  $\{|\xi| \sim \lambda \wedge |\xi'| \ll \lambda\}$ . In the following we derive a favorable estimate of  $\|\langle D' \rangle^{-\rho-\frac{\sigma}{p}} S_\lambda^{\ll \tau} u\|_{L^p L^q}$  for  $\lambda \gg 1$ ,  $(\rho, p, q, 2)$  a Strichartz pair, and  $p \neq \infty$ . As symbol at unit frequencies we consider

$$p(x, \xi) = i \begin{pmatrix} \xi_0 & 0 & -\xi_2 \\ 0 & \xi_0 & \xi_1 \\ -\varepsilon_{11}(x)\xi_2 + \varepsilon_{21}(x)\xi_1 & \varepsilon_{22}(x)\xi_1 - \varepsilon_{12}(x)\xi_2 & \xi_0 \end{pmatrix} s_0(\xi)$$

with  $s_0 \in C_c^\infty(B(0, 2))$ ,  $s_0 = 1$  for  $1 \sim |\xi| \sim |\xi_0| \gg |\xi'|$  such that  $S_\lambda^{\ll \tau} = s_0(\xi/\lambda)$ .

By Hölder's inequality and Sobolev embedding, we find

$$\begin{aligned}
\|\langle D' \rangle^{-\rho - \frac{\sigma}{p}} S_\lambda^{\llcorner \tau} u\|_{L^p L^q} &\lesssim \lambda^{\frac{1}{2} - \frac{\sigma}{p}} \|S_\lambda^{\llcorner \tau} u\|_{L^2} \\
&= \lambda^{\frac{1}{2} - \frac{\sigma}{p}} \|T_\lambda^* p^{-1}(x, \xi) T_\lambda T_\lambda^* p(x, \xi) T_\lambda S_\lambda^{\llcorner \tau} u\|_{L^2} \\
&\lesssim \lambda^{\frac{1}{2} - \frac{\sigma}{p}} \|p(x, \xi) T_\lambda S_\lambda^{\llcorner \tau} u\|_{L^2} \\
&\lesssim \lambda^{\frac{1}{2} - \frac{\sigma}{p}} \|R_{\lambda, p}\|_{L^\infty L^2 \rightarrow L^2} \|S_\lambda^{\llcorner \tau} u\|_{L^\infty L^2} \\
&\quad + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} (\|P(x, D) S_\lambda^{\llcorner \tau} u\|_{L^2} + \|\partial_x \varepsilon\|_{L^2 L^\infty} \|S_\lambda^{\llcorner \tau} u\|_{L^\infty L^2}).
\end{aligned}$$

Since  $\|R_{\lambda, p}\|_{L^\infty L^2 \rightarrow L^2} \lesssim \|\partial_x \varepsilon\|_{L^2 L^\infty} \lambda^{-\frac{1}{2}}$  by Theorem 5.2, the first term is acceptable. The third term comes from recovering divergence form and is clearly admissible. Also note that we omitted an error term from localizing  $T_\lambda S_\lambda^{\llcorner \tau} u$  to  $\{|\xi_0| \sim 1, |\xi'| \ll 1\}$  as this gains arbitrarily many derivatives. We turn to the second term. For this purpose let  $P^\lambda(x, D)$  denote  $P$  with coefficients frequency truncated at  $\lambda$ . The coefficients of  $P^\lambda$  have Lipschitz norm  $\lesssim \lambda^{\frac{1}{2}}$ . Note that

$$\begin{aligned}
\|P(x, D) S_\lambda^{\llcorner \tau} u\|_{L^2} &\leq \|P^\lambda(x, D) S_\lambda^{\llcorner \tau} u\|_{L^2} + \|P^{\gtrsim \lambda}(x, D) S_\lambda^{\llcorner \tau} u\|_{L^2} \\
&\lesssim \|P^\lambda(x, D) S_\lambda^{\llcorner \tau} u\|_{L^2} + \|\partial_x \varepsilon\|_{L^2 L^\infty} \|u\|_{L^\infty L^2}.
\end{aligned}$$

Moreover, a kernel estimate yields

$$\begin{aligned}
\|P^\lambda S_\lambda^{\llcorner \tau} u\|_{L^2} &\leq \|S_\lambda^{\llcorner \tau} P^\lambda u\|_{L^2} + \|[P^\lambda, S_\lambda^{\llcorner \tau}]u\|_{L^2} \\
&\lesssim \|S_\lambda^{\llcorner \tau} P^\lambda u\|_{L^2} + \lambda^{\frac{1}{2}} \|u\|_{L^2} \\
&\lesssim \|S_\lambda^{\llcorner \tau} P u\|_{L^2} + \|S_\lambda^{\llcorner \tau} P^{\gtrsim \lambda} u\|_{L^2} + \lambda^{\frac{1}{2}} \|u\|_{L^2}.
\end{aligned}$$

For the first term we use Bernstein's inequality in time

$$\|S_\lambda^{\llcorner \tau} P u\|_{L^2} \lesssim \lambda^{\frac{1}{2}} \|S_\lambda^{\llcorner \tau} P u\|_{L^1 L^2}$$

and note that it is still summable in  $\lambda$  due to the additional factor  $\lambda^{-\frac{\sigma}{p}}$  and since we only want to estimate  $\|\langle D' \rangle^{-\alpha} u\|_{L^p L^q}$  for  $\alpha > \rho + \frac{\sigma}{p}$  in Corollary 1.7. Likewise, the third term can be summed. For the second term we note

$$\|S_\lambda^{\llcorner \tau} P^{\gtrsim \lambda} u\|_{L^2} \lesssim \|\partial_x \varepsilon\|_{L^2 L^\infty} \|u\|_{L^\infty L^2}.$$

This handles the part with spatial frequencies much smaller than temporal ones.

For the dyadic frequency blocks  $\{|\tau| \sim |\xi'| \sim \lambda\}$  we can use Theorem 1.4 with  $P^\lambda$  to recover the dyadic estimate (69):

$$\begin{aligned}
\lambda^{-\rho - \frac{\sigma}{p}} \|S_\lambda S'_\lambda u\|_{L^p L^q} &\lesssim \|S_\lambda u\|_{L^\infty L^2} + \lambda^{-\sigma} \|P^\lambda S_\lambda S'_\lambda u\|_{L^1 L^2} \\
&\quad + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S'_\lambda \rho_e(0)\|_{L^2} + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S'_\lambda \partial_t \rho_e\|_{L^1 L^2} \\
&\lesssim \|u\|_{L^\infty L^2} + \lambda^{-\sigma} \|S_\lambda P u\|_{L^1 L^2} \\
&\quad + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S'_\lambda \rho_e(0)\|_{L^2} + \lambda^{-\frac{1}{2} - \frac{\sigma}{p}} \|S'_\lambda \partial_t \rho_e\|_{L^1 L^2}.
\end{aligned}$$

We used the same commutator considerations as in the proof of Theorem 1.4 and the fundamental theorem of calculus together with Minkowski's inequality. By the energy estimate, we conclude for  $\alpha > \rho + \frac{\sigma}{p}$

$$\begin{aligned}
\|\langle D' \rangle^{-\alpha} u\|_{L^p(0, T; L^q)} &\lesssim \|u_0\|_{L^2} + \|P u\|_{L^1 L^2} \\
&\quad + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{p}} \rho_e(0)\|_{L^2} + \|\langle D' \rangle^{-\frac{1}{2} - \frac{\sigma}{p}} \partial_t \rho_e\|_{L^1 L^2},
\end{aligned}$$

and the proof is complete.  $\square$



## 6. LOCAL WELL-POSEDNESS FOR A 2D QUASILINEAR MAXWELL SYSTEM

This section is devoted to the proof of Theorem 1.9. Recall that the system under consideration is given by

$$(73) \quad \begin{cases} \partial_t u_1 &= \partial_2 u_3, & u(0) = u_0 \in H^s(\mathbb{R}^2; \mathbb{R})^3, \\ \partial_t u_2 &= -\partial_1 u_3, & \partial_1 u_1 + \partial_2 u_2 = 0, \\ \partial_t u_3 &= \partial_2(\varepsilon^{-1}(u)u_1) - \partial_1(\varepsilon^{-1}(u)u_2), \end{cases}$$

where  $\varepsilon^{-1}(u) = \psi(|u_1|^2 + |u_2|^2)$  and  $\psi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is smooth, monotone increasing and satisfies  $\psi(0) = 1$ . We denote  $\tilde{u} = (u_1, u_2)$ .

Without using dispersive properties, energy methods yield local well-posedness in  $H^s(\mathbb{R}^2)$  for  $s > 2$ . To improve on this by Strichartz estimates, we follow the arguments from Ifrim–Tataru [14]. Let  $A = \sup_{0 \leq t' \leq t} \|u(t')\|_{L_{x'}^\infty}$  and  $B(t) = \|\nabla_{x'} u(t)\|_{L_{x'}^\infty}$ . In the following we take local existence of smooth solutions for granted, which follows by classical arguments, e.g., parabolic regularization. We focus on proving estimates in rough norms. The argument consists of three steps:

- 1) We derive energy estimates

$$(74) \quad E^s(u(t)) \lesssim e^{c(A) \int_0^t B(t') dt'} E^s(u(0)),$$

for smooth solutions  $u$ , where  $E^s(u) \approx_A \|u\|_{H^s}$  and  $s \geq 0$ .

- 2) We show  $L^2$ -Lipschitz bounds for differences of solutions  $v = u^1 - u^2$

$$(75) \quad \|v(t)\|_{L^2}^2 \lesssim e^{c(A) \int_0^t B(t') dt'} \|v(0)\|_{L^2}^2,$$

where  $u^1$  and  $u^2$  solve (73) and

$$A = \sup_{0 \leq t' \leq t} \|u^1(t')\|_{L_{x'}^\infty} + \sup_{0 \leq t' \leq t} \|u^2(t')\|_{L_{x'}^\infty},$$

$$B(t) = \|\nabla_{x'} u^1(t)\|_{L_{x'}^\infty} + \|\nabla_{x'} u^2(t)\|_{L_{x'}^\infty}.$$

- 3) We conclude the proof of Theorem 1.9, using frequency envelopes. These were introduced by Tao in the context of wave maps [36] and turned out as very useful to treat quasilinear evolution equations.

We start with energy estimates.

**Proposition 6.1.** *Let  $s \geq 0$ . Then, we find (74) to hold. Let  $s > 11/6$ . For  $u_0 \in H^s$ , there is a time  $T = T(\|u_0\|_{H^s})$  such that  $T$  is lower semicontinuous and*

$$\sup_{t \in [0, T]} \|u(t)\|_{H^s} \lesssim \|u_0\|_{H^s}.$$

*Proof.* We consider energy norms

$$\|u\|_{E^s}^2 = \langle \langle D' \rangle^s u, C(u) \langle D' \rangle^s u \rangle \approx_A \|u\|_{H^s}^2,$$

where we define the smooth map  $C$  in (78) below. We rewrite (73) as

$$(76) \quad \partial_t u = \mathcal{A}^j(u) \partial_j u,$$

with the coefficient matrices

$$\mathcal{A}^1(u) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ -2\psi'(|\tilde{u}|^2)u_1u_2 & -2\psi'(|\tilde{u}|^2)u_2^2 - \psi(|\tilde{u}|^2) & 0 \end{pmatrix},$$

$$\mathcal{A}^2(u) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 2\psi'(|\tilde{u}|^2)u_1^2 + \psi(|\tilde{u}|^2) & 2\psi'(|\tilde{u}|^2)u_1u_2 & 0 \end{pmatrix}.$$

For the time derivative we find

$$(77) \quad \frac{d}{dt} \|u\|_{E^s}^2 = \langle \langle D' \rangle^s (\mathcal{A}^j(u) \partial_j u), C(u) \langle D' \rangle^s u \rangle + \langle \langle D' \rangle^s u, C(u) \langle D' \rangle^s (\mathcal{A}^j(u) \partial_j u) \rangle \\ + \langle \langle D' \rangle^s u, C'(u) (\partial_t u) \langle D' \rangle^s u \rangle,$$

at time  $t$ , which is suppressed. Because of (76), the last term is estimated by

$$|\langle \langle D' \rangle^s u, \tilde{C}(u) (\partial_t u) \langle D' \rangle^s u \rangle| \lesssim c(A) B \|u\|_{H^s}^2.$$

The first term in (77) can be expressed as

$$\begin{aligned} & \langle \langle D' \rangle^s (\mathcal{A}^j(u) \partial_j u), C(u) \langle D' \rangle^s u \rangle \\ &= \langle \mathcal{A}^j(u) (\langle D' \rangle^s \partial_j u) + (\langle D' \rangle^s \mathcal{A}^j(u) - \mathcal{A}^j(u) \langle D' \rangle^s) \partial_j u, C(u) \langle D' \rangle^s u \rangle \\ &= \langle \mathcal{A}^j(u) (\langle D' \rangle^s \partial_j u), C(u) \langle D' \rangle^s u \rangle + II. \end{aligned}$$

Paraproduct/Moser estimates yield

$$II \leq \| \langle \langle D' \rangle^s \mathcal{A}^j(u) - \mathcal{A}^j(u) \langle D' \rangle^s \rangle \partial_j u \|_{L^2} \| C(u) \langle D' \rangle^s u \|_{L^2} \lesssim_A B \|u\|_{H^s}^2.$$

Integrating by parts, the other summand becomes

$$\begin{aligned} & \langle \mathcal{A}^j(u) \langle D' \rangle^s \partial_j u, C(u) \langle D' \rangle^s u \rangle_{L^2} \\ &= - \langle \langle D' \rangle^s u, \partial_j (\mathcal{A}_j(u) * C(u) \langle D' \rangle^s u) \rangle_{L^2} \\ &= - \langle \langle D' \rangle^s u, \mathcal{A}^j(u) * C(u) \langle D' \rangle^s \partial_j u \rangle_{L^2} + \mathcal{O}_A(B \|u\|_{H^s}^2). \end{aligned}$$

The second summand in (77) can be treated analogously. We seek to cancel the highest-order terms. This gives the condition

$$\mathcal{A}^j(u) * C(u) = C(u) * \mathcal{A}^j(u).$$

Solving the system of equations and setting  $C_{33} = 1$ , we find

$$(78) \quad C(u) = \begin{pmatrix} \psi + 2\psi' \cdot u_1^2 & 2\psi' \cdot u_1 u_2 & 0 \\ 2\psi' \cdot u_1 u_2 & \psi + 2\psi' \cdot u_2^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We thus conclude

$$\frac{d}{dt} \|u(t)\|_{E^s}^2 \lesssim_A B(t) \|u(t)\|_{E^s}^2,$$

and Grönwall's lemma implies (74). To check that

$$\|u\|_{H^s}^2 \approx_A \|u\|_{E^s}^2,$$

it is enough to show that  $C$  is uniformly elliptic. This follows from Young's inequality in the computation

$$\begin{aligned} & \left\langle \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \begin{pmatrix} \psi + 2\psi' \cdot u_1^2 & 2\psi' \cdot u_1 u_2 \\ 2\psi' \cdot u_1 u_2 & \psi + 2\psi' \cdot u_2^2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\rangle \\ & \geq (\psi + 2\psi' \cdot u_1^2) \xi_1^2 - 4\psi' \cdot \left( \frac{u_1^2 \xi_1^2}{2} + \frac{u_2^2 \xi_2^2}{2} \right) + (\psi + 2\psi' \cdot u_2^2) \xi_2^2 \\ & = \psi \cdot |\xi|^2. \end{aligned}$$

The second claim of the proposition is a consequence of Sobolev's embedding for  $s > 2$ . To improve on this, we use Strichartz estimates and show

$$(79) \quad \|\nabla_{x'} u\|_{L^4(0,T;L^\infty)} \lesssim \|u_0\|_{H^s}$$

for  $s > 11/6$ . To bootstrap, we require that  $\|\nabla_{x'} u\|_{L^4(0,T_0;L^\infty)} \leq K$  for a fixed number  $K > 0$  and a maximally defined time  $T_0 > 0$ . Take  $T \in (0, T_0)$  with

$\|\partial_x \varepsilon\|_{L^2(0,T;L^\infty)} \lesssim_A T^{1/4} K \leq 1$  and  $\|\partial_x \varepsilon\|_{L^1(0,T;L^\infty)} \lesssim_A T^{3/4} K \leq 1$ . We thus have uniform constants in the energy inequality (74) and in the Strichartz estimate

$$(80) \quad \|\langle D' \rangle^{-\alpha} w\|_{L^p(0,T;L^q)} \lesssim \|w_0\|_{L^2} + \|P(x, D)w\|_{L^1 L^2},$$

for  $\alpha > \rho + \frac{1}{3p}$  from Corollary 1.7 with  $\tilde{s} = 1$ , if  $\partial_1 w_1 + \partial_2 w_2 = 0$ . For low frequencies, Bernstein's inequality and (74) yield

$$\|S'_{\lesssim 1} \nabla_{x'} u\|_{L^4 L^\infty} \lesssim_A T^{1/4} \|u_0\|_{L^2}.$$

For high frequencies, we define the auxiliary function  $v = \langle D' \rangle^s u$  satisfying  $\partial_1 v_1 + \partial_2 v_2 = 0$ . Similar as above, a fixed time estimate gives

$$(81) \quad \|P(x, u, D)v(t)\|_{L^2} = \| [P(x, u, D), \langle D' \rangle^s] u(t) \|_{L^2} \lesssim_A \|\nabla u(t)\|_{L_x^\infty} \|u(t)\|_{H^s},$$

where  $P(x, u, D)$  has coefficients  $\varepsilon(u)^{-1}$ . For the Strichartz pair  $(3/4, 4, \infty, 2)$ , inequality (80) implies

$$\|S'_{\geq 1} \nabla_{x'} u\|_{L^4 L^\infty} \lesssim \|\langle D' \rangle^{1-s} v\|_{L^4 L^\infty} \lesssim \|v_0\|_{L^2} + \|P(x, u, D)v\|_{L^1 L^2},$$

since  $s > 1 + \rho + \frac{1}{3p} = \frac{11}{6}$  by our assumption. By virtue of (81) and (74), we conclude

$$\begin{aligned} \|S'_{\geq 1} \nabla_{x'} u\|_{L^4 L^\infty} &\lesssim_A \|u_0\|_{H^s} + \|\nabla_{x'} u\|_{L^1 L_x^\infty} \|u\|_{L^\infty H^s} \\ &\lesssim \|u_0\|_{H^s} + T^{4/3} \|\nabla_{x'} u\|_{L^4 L_x^\infty} \|u\|_{L^\infty H^s} \\ &\lesssim \|u_0\|_{H^s} + T^{4/3} \|\nabla_{x'} u\|_{L^4 L_x^\infty} \|u_0\|_{H^s} \end{aligned}$$

on  $[0, T]$  also using the equivalence  $\|u(t)\|_{E^s} \approx_A \|u(t)\|_{H^s}$ . Starting with a sufficiently large  $K$ , we can now fix a small  $T_1 = T_1(\|u_0\|_{H^s}) \in (0, T_0)$  such that  $\|\nabla_{x'} u\|_{L^4(0, T_1; L^\infty)} \lesssim \|u_0\|_{H^s} < K$ .  $\square$

We turn to the  $L^2$ -bound for differences.

**Proposition 6.2.** *Let  $u^1$  and  $u^2$  be two solutions to (73) with finite  $A$  and  $B$ , and set  $v = u^1 - u^2$ . Then, we find (75) to hold. Moreover, if  $s > 11/6$ , there is a time  $T = T(\|u^i(0)\|_{H^s})$  such that  $T$  is lower semicontinuous and*

$$(82) \quad \sup_{t \in [0, T]} \|v(t)\|_{L^2} \lesssim_{\|u^i(0)\|_{H^s}} \|v(0)\|_{L^2}.$$

*Proof.* We observe that  $v$  solves the equation

$$\begin{aligned} \partial_t v &= \mathcal{A}^j(u^1) \partial_j v + [\mathcal{A}^1(u^1) - \mathcal{A}^1(u^2)] \partial_1 u^2 + [\mathcal{A}^2(u^1) - \mathcal{A}^2(u^2)] \partial_2 u^2 \\ &= \mathcal{A}^j(u^1) \partial_j v + \mathcal{B}^j(u^1, u^2)(v, \partial_j u^2). \end{aligned}$$

To prove the claim, we work with the equivalent norm

$$\|v\|_{2, u^1}^2 = \langle v, C(u^1)v \rangle_2$$

with  $C$  from the previous proof. For fixed  $t$ , we calculate

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|v\|_{2, u^1}^2 &= \langle \mathcal{A}^j(u^1) \partial_j v, C(u^1)v \rangle + \langle v, C(u^1) \mathcal{A}^j(u^1) \partial_j v \rangle \\ &\quad + \langle \mathcal{B}^j(u^1, u^2)(v, \partial_j u^2), C(u^1)v \rangle + \langle v, C(u^1) \mathcal{B}^j(u^1, u^2)(v, \partial_j u^2) \rangle \\ &\quad + \mathcal{O}_A(B \|v\|_2^2), \end{aligned}$$

where the error accounts for the contribution of the time derivative of  $C(u^1)$ . Furthermore,

$$|\langle \mathcal{B}^j(u^1, u^2)(v, \partial_j u^2), C(u^1)v \rangle| \leq \|\mathcal{B}^j(u^1, u^2)(v, \partial_j u^2)\|_2 \|C(u^1)v\|_2 \lesssim_A B \|v\|_2^2.$$

The key estimate

$$|\langle \mathcal{A}^j(u^1)\partial_j v, C(u^1)v \rangle + \langle v, C(u^1)\mathcal{A}^j(u^1)\partial_j v \rangle| \lesssim_A B \|v\|_2^2$$

is carried out as in the proof of Proposition 6.1, employing our choice of  $C$ . Estimate (79) now implies (82).  $\square$

In the third step, we show continuous dependence by means of the frequency envelope argument detailed in Ifrim–Tataru [14]. The envelopes represent the dyadically localized Sobolev energy. We use the following tailored version, where  $P_k$  is the standard Littlewood–Paley projector at frequency  $2^k$ .

**Definition 6.3.**  $(c_k)_{k \geq 0} \in \ell^2$  is called a *frequency envelope* for a function  $u$  in  $H^s$  if we have the following properties.

a) Energy bound:

$$\|P_k u\|_{H^s} \leq c_k.$$

b) Slowly varying: There is  $\delta > 0$  such that for all  $j, k \in \mathbb{N}$

$$\frac{c_k}{c_j} \lesssim 2^{-\delta|j-k|}.$$

The envelopes are called *sharp* if they also  $\|u\|_{H^s}^2 \approx \sum_k c_k^2$  for a family of functions.

As noted in [14] such envelopes always exist. The idea is to show that for a solution not only the  $H^s$ -norm is propagated, but also the frequency envelope for a time depending on the control parameters. This allows to infer continuous dependence. We give the details.

*Regularization.* Let  $u_0 \in H^s(\mathbb{R}^2)$  and  $(c_k)_{k \geq 0}$  be a sharp frequency envelope for  $u_0$  in  $H^s$ . The regularized initial data  $u_0^n = P_{\leq n} u_0$  for  $n \in \mathbb{N}$  have the following properties.

i) Uniform bounds:

$$\|P_k u_0^n\|_{H^s} \lesssim c_k,$$

ii) High frequency bounds:

$$\|u_0^n\|_{H^{s+j}} \lesssim 2^{jn} c_n,$$

iii) Difference bounds:

$$\|u_0^{n+1} - u_0^n\|_{L^2} \lesssim 2^{-sn} c_n,$$

iv) Limit as  $n \rightarrow \infty$ :

$$u_0 = \lim_{n \rightarrow \infty} u_0^n \text{ in } H^s.$$

The regularized initial data give rise to a family of smooth solutions.

*Uniform bounds:* Propositions 6.1 and 6.2 now yield a time interval of length  $T = T(\|u_0\|_{H^s})$  on which the solutions exist, and also  $L^2$ -bounds their differences:

i) High frequency bounds:

$$\|u^n\|_{C([0,T], H^{s+j})} \lesssim 2^{nj} c_n,$$

ii) Difference bounds:

$$\|u^{n+1} - u^n\|_{C([0,T], L^2)} \lesssim 2^{-sn} c_n.$$

Interpolation implies

$$\|u^{n+1} - u^n\|_{C([0,T],H^m)} \lesssim c_n 2^{-(s-m)n}$$

for  $m \geq 0$ . By the  $L^2$ -bound, we have

$$\|u - u^n\|_{C([0,T],L^2)} \lesssim 2^{-sn}.$$

This gives convergence in  $L^2$  since

$$u - u^n = \sum_{j=n}^{\infty} (u^{j+1} - u^j).$$

The frequency localization of the summands at  $2^{j+1}$  and the error bound show that

$$\|u - u^n\|_{C([0,T],H^s)} \lesssim c_{\geq n} = \left( \sum_{j \geq n} c_j^2 \right)^{1/2},$$

which gives convergence in  $C([0,T],H^s)$ . We can now show Theorem 1.9.

*Proof of Theorem 1.9.* Consider a sequence of initial data

$$u_{0k} \rightarrow u_0 \text{ in } H^s(\mathbb{R}^2), \quad s > \frac{11}{6}.$$

Propositions 6.1 and 6.2 show that the solutions  $u_k$  with initial values  $u_{0,k}$  exist on a common time interval  $[0, T]$  with  $T = T(\|u_0\|_{H^s})$  and that  $u_k$  converges to  $u$  in  $C([0, T], L^2)$ . For the solutions  $u_k^n$  and  $u^n$  with regularized initial data  $u_{0k}^n$  and  $u^n$ , we have  $u_{0k}^n \rightarrow u_0^n$  in  $H^s$  and hence

$$u_k^n \rightarrow u^n \text{ in } C([0, T], H^s)$$

as  $k \rightarrow \infty$  by the above reasoning. We then derive

$$\begin{aligned} \|u_k - u\|_{C([0,T],H^s)} &\lesssim \|u_k^n - u^n\|_{C([0,T],H^s)} + \|u^n - u\|_{C([0,T],H^s)} \\ &\quad + \|u_k^n - u_k\|_{C([0,T],H^s)} \\ &\lesssim \|u_k^n - u^n\|_{C([0,T],H^s)} + c_{\geq n} + c_{\geq n}^k. \end{aligned}$$

The convergence  $u_{0k} \rightarrow u_0$  in  $H^s$  allows us to choose a sequence of frequency envelopes  $c^k \rightarrow c$  in  $\ell^2$ . We can thus take  $n$  uniformly in  $k$  such that the second and third term become arbitrarily small, and then let  $k \rightarrow \infty$ .  $\square$

## 7. SHARPNESS OF DERIVATIVE LOSS

Next, we connect the Maxwell system with a wave equation to infer the sharpness of derivative loss in Theorem 1.4 for permittivity coefficients  $\varepsilon^{ij} \in C^s$  with  $1 \leq s \leq 2$ . For this purpose, we recall the counterexamples in Smith–Tataru [32]. Smith and Tataru constructed time-independent  $C^s$ -metrics  $g$  for  $0 \leq s \leq 2$  and solutions  $u$  to the corresponding wave equations which exhaust an derivative loss analogous to that proved in Theorem 1.4. They treated the second-order hyperbolic operator

$$Q(t, x, \partial_t, \partial_x) = \partial_t^2 - \partial_i g^{ij}(t, x) \partial_j$$

on  $[0, 1] \times \mathbb{R}^n$  with  $g^{ij} \in C^s$  for  $n \geq 2$  and  $0 \leq s \leq 2$ , and the Strichartz estimates

$$\|u\|_{L^p(0,1;L^q)} \lesssim \|u\|_{L^\infty(0,1;H^{\rho'})} + \|\partial_t u\|_{L^\infty(0,1;H^{\rho'-1})} + \|Qu\|_{L_t^1(0,1;H^{\rho'-1})},$$

where  $(\rho, p, q, n)$  is a Strichartz pair and

$$\rho' = \rho + \frac{\sigma}{p}, \quad \sigma = \frac{2-s}{2+s}, \quad \delta = \frac{2}{2+s} = \frac{1+\sigma}{2}.$$

In [32] it is proved that  $\rho'$  cannot be lowered. For  $2/3 \leq s \leq 2$ , the authors at first look at the equation

$$(83) \quad Q_\lambda(y, \partial_t, \partial_x, \partial_y) = \partial_t^2 - g_\lambda(y) \partial_x^2 - \Delta_y,$$

with  $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1}$  and  $g_\lambda(y) = 1 + \lambda^{2\sigma} |y|^2$ . Later  $g_\lambda$  is replaced by  $\tilde{g}_\lambda(y) = 1 + \lambda^{2\sigma-2\delta} a(\lambda^\delta |y|)$  for a smooth  $a \geq 0$  supported in  $[0, 2)$  with  $a(r) = r^2$  for  $r \in [0, 1]$ . The  $C^s$  norm of  $\tilde{g}_\lambda$  is uniformly bounded in  $\lambda$ .

We consider smooth solutions to Maxwell equations in two spatial dimensions

$$(84) \quad \begin{cases} \partial_t D &= \nabla_\perp H, & \partial_1 D_1 + \partial_2 D_2 &= \rho_e, \\ \partial_t H &= -\nabla \times E = \partial_2 E_1 - \partial_1 E_2, \end{cases}$$

with the rough permittivity  $\varepsilon_\lambda$  given by

$$\varepsilon_\lambda^{-1}(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 + \lambda^{2\sigma} y^2 \end{pmatrix}$$

with  $\lambda \geq 2$ . Corresponding to the Strichartz estimate above, we look at

$$(85) \quad \|(D, H)\|_{L^p L^q} \lesssim \|(D, H)\|_{L^\infty H^\gamma} + \|P(D, H)\|_{L^1 H^{\gamma-\sigma}} + \|\rho_e\|_{L^\infty H^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}}$$

with  $P$  as in (6), omitting the time interval  $(0, 1)$ . We want to show that this estimate can only hold for all  $u$  if  $\gamma \geq \rho + \frac{\sigma}{p}$ , cf. Theorem 1.4.

The field  $H$  from (84) solves the wave equation

$$\partial_t^2 H = (1 + \lambda^{2\sigma} y^2) \partial_1^2 H + \partial_2^2 H =: \Delta_{\varepsilon_\lambda^{-1}} H.$$

As in [32] we choose

$$(86) \quad H(t, x, y) = H^\lambda(t, x, y) = \frac{1}{(\log \lambda)^2} \int \beta((\log \lambda)^{-2} r) u_r^\lambda(t, x, y) dr$$

for the function

$$(87) \quad u_r^\lambda(t, x, y) = e^{ir\lambda(t-x) - \frac{i}{2}\lambda^\sigma t - \frac{i}{2}r\lambda^{2\sigma} y^2}$$

with  $0 \neq \beta \in C_c^\infty(\mathbb{R})$  satisfying  $\text{supp } \beta \subseteq [1, 2]$  and  $\beta \geq 0$ . As noted in [32], the function  $H^\lambda$  is essentially supported in  $K_\lambda^t = \{|t-x| \leq \lambda^{-1}(\log \lambda)^{-2}, |y| \leq \lambda^{-\delta}(\log \lambda)^{-1}\}$ .

To solve (84), we set

$$(88) \quad D^\lambda(t, x, y) = \int_0^t \nabla_\perp H^\lambda(s, x, y) ds + C^\lambda(x, y) + \varepsilon_\lambda \nabla \psi^\lambda(x, y)$$

with  $C^\lambda$  and  $\psi^\lambda$  to be determined.

Below we choose  $C^\lambda$  with support in

$$K_\lambda^0 = \{|x| \leq \lambda^{-1}(\log \lambda)^{-2}, |y| \leq \lambda^{-\delta}(\log \lambda)^{-1}\}.$$

In view of the second line of (84), we compute

$$\partial_t H^\lambda + \nabla \times (\varepsilon_\lambda^{-1} D^\lambda) = \partial_t H^\lambda(t, x, y) - \int_0^t \Delta_{\varepsilon_\lambda^{-1}} H^\lambda(s, x, y) ds + \nabla \times (\varepsilon_\lambda^{-1} C^\lambda).$$

Inserting the equation

$$\partial_t^2 H^\lambda - \Delta_{\varepsilon_\lambda^{-1}} H^\lambda = -\frac{1}{4} \lambda^{2\sigma} H^\lambda$$

from [32], we obtain

$$\partial_t H^\lambda(t, x, y) - \int_0^t \Delta_{\varepsilon_\lambda^{-1}} H^\lambda(s, x, y) ds = \partial_t H^\lambda(0, x, y) - \frac{\lambda^{2\sigma}}{4} \int_0^t H^\lambda(s, x, y) ds.$$

Because of (86) and (87), we can calculate the time integral. Compared to (89) it gives a lower-order term on the right-hand side (85) as  $\lambda \rightarrow \infty$ . (Here we use that  $2\sigma \leq 1$  since  $s \geq 2/3$ .) Hence, we seek  $\nabla \times (\varepsilon_\lambda^{-1} C^\lambda) \approx \partial_t H^\lambda(0, x, y)$ . We set

$$C_1^\lambda = 0, \quad C_2^\lambda(x, y) = \varphi(x, y) g_\lambda(y)^{-1} \int_{-\infty}^x \partial_t H^\lambda(0, x', y) dx',$$

with  $\varphi$  localized to  $B(0, 1/2)$ , which contains the essential support of  $\partial_t H^\lambda(0, x, y)$  for  $\lambda \gg 1$ . Due to (86) and (87), one can check that also the term

$$r(x, y) = \nabla \times (\varepsilon_\lambda^{-1} C^\lambda) - \partial_t H^\lambda(0)$$

gives a lower-order term. Using (5) and (6) in [32], we have

$$(89) \quad \begin{aligned} \|H^\lambda\|_{L^p L^q} &\gtrsim \lambda^{-(1+\delta)/q} (\log \lambda)^{-3/q}, \\ \|H^\lambda(t)\|_{H^\gamma(\mathbb{R}^2)} &\lesssim \lambda^{\gamma-(1+\delta)/2} (\log \lambda)^{2\gamma-3/2}. \end{aligned}$$

If we can also prove that

$$(90) \quad \|\rho_e^\lambda\|_{H^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} \lesssim \|H^\lambda\|_{H^\gamma} \quad \text{and} \quad \|D^\lambda\|_{H^\gamma} \lesssim \|H^\lambda\|_{H^\gamma},$$

letting  $\lambda \rightarrow \infty$  we see that the validity of (85) implies  $\gamma \geq \rho + \frac{\sigma}{p}$ . This remains true if we modify  $1 + \lambda^{2\sigma}|y|^2$  to  $\tilde{g}_\lambda(y)$  in  $\varepsilon_\lambda^{-1}$ , since  $H^\lambda$  decays fast enough outside  $K_\lambda$ . (See p. 202 in [32].)

To make the charge as small as possible, in view of (88) we solve approximately

$$(91) \quad \partial_1 C_1^\lambda + \partial_2 C_2^\lambda = -(\partial_1^2 + \partial_2(1 + \lambda^{2\sigma} y^2)^{-1} \partial_2) \psi^\lambda = -\Delta_{\varepsilon_\lambda^{-1}} \psi^\lambda.$$

For this purpose, we solve (91) in  $B(0, 1)$  with Dirichlet boundary conditions. Let  $\psi_0^\lambda$  denote the solution in  $B(0, 1)$ . Let  $\chi : [0, \infty) \rightarrow [0, 1]$  be a smooth, monotone decreasing function with  $\chi(r) = 1$  for  $r \in [0, \frac{1}{2}]$  and  $\chi(r) = 0$  for  $r \geq \frac{3}{4}$ . We set  $\psi^\lambda(x, y) = \chi(|(x, y)|) \psi_0^\lambda(x, y)$ .

We can now show (90). We treat  $D^\lambda$  by estimating the contributions of the terms in (88) separately: From (87) follows that the contribution of  $\int_0^t \nabla_\perp H^\lambda ds$  is acceptable because the anti-derivative in  $t$  cancels the derivative loss. The same argument bounds the contribution of  $C^\lambda$ , so that  $\|C^\lambda\|_{H^\gamma} \lesssim \|H^\lambda\|_{H^\gamma}$ . By elliptic regularity, we thus obtain the bound  $\|\psi^\lambda\|_{H^{\gamma+1}} \lesssim \|C^\lambda\|_{H^\gamma} \lesssim \|H^\lambda\|_{H^\gamma}$  since the coefficients in  $\Delta_{\varepsilon_\lambda^{-1}}$  are uniformly bounded in  $C^s$ . Here we need the additional assumption  $1 \leq s \leq 2$ , which is not necessary in [32].

Lastly, we compute for the charge

$$\rho_e^\lambda = \nabla \cdot D^\lambda = \nabla \cdot C^\lambda + \Delta_{\varepsilon_\lambda^{-1}} \psi^\lambda.$$

Hence, we find  $\rho_e^\lambda = 0$  in  $B(0, 1/2)$  and  $B(0, 3/4)^c$ . In  $B(0, 3/4) \setminus B(0, 1/2)$ , we find

$$\rho_e^\lambda = 2(\varepsilon_\lambda \nabla \chi)(\nabla \psi_0^\lambda) + (\partial_2 g_\lambda^{-1})(\partial_2 \psi_0^\lambda) \chi + \text{l.o.t.}$$

Since  $\|\psi^\lambda\|_{H^{\gamma+1}} \lesssim \|H^\lambda\|_{H^\gamma}$  and  $g_\lambda^{-1} \in C^s$  with  $s \geq 1$ , only the second summand in the display is problematic. First we handle  $\gamma - \frac{1}{2} - \frac{\sigma}{p} \leq 0$ , in which case the estimate

$$\|\rho_e^\lambda\|_{H^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} \lesssim \|H^\lambda\|_{H^\gamma}$$

is straight-forward. For  $\gamma - \frac{1}{2} - \frac{\sigma}{p} > 0$ , observe that  $\|g_\lambda^{-1}\|_{C^\theta(K_\lambda^0)} \lesssim \lambda^{2\sigma-(2-\theta)\delta}$ . We refer to [32] for the true coefficient  $\tilde{g}_\lambda$  with  $\|\tilde{g}_\lambda^{-1}\|_{C^\theta(B(0,1))} \lesssim \|g_\lambda^{-1}\|_{C^\theta(K_\lambda^0)}$ . Suppose

that  $0 < \gamma < \rho + \frac{\sigma}{p}$ . Hence,  $\|\partial_2 g_\lambda^{-1}\|_{C^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} \lesssim \lambda^\alpha$  with  $\alpha = 2\sigma - (\frac{3}{2} - \gamma + \frac{\sigma}{p})\delta$ . Furthermore,

$$\begin{aligned} \|(\partial_2 g_\lambda^{-1})(\partial_2 \psi_0^\lambda)\chi\|_{H^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} &\lesssim \|\partial_2 g_\lambda^{-1}\|_{C^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} \|\partial_2 \psi_0^\lambda\|_{L^2} + \|\partial_2 \psi_0^\lambda\|_{H^{\gamma-\frac{1}{2}-\frac{\sigma}{p}}} \\ &\lesssim \lambda^\alpha \lambda^{-\frac{1+\delta}{2}} (\log \lambda)^{-\frac{3}{2}} + \|H_\lambda\|_{H^\gamma}, \end{aligned}$$

and it remains to check  $\alpha < \gamma$ , which follows from  $(1 - \delta)\gamma > \frac{5}{4}\sigma - \frac{3}{4}$ . The latter inequality holds as  $\gamma > 0$ ,  $\delta \leq \frac{2}{3}$ , and  $\sigma \leq \frac{1}{3}$  because of  $s \geq 1$ .  $\square$

**Acknowledgement.** Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

#### REFERENCES

- [1] Hajer Bahouri and Jean-Yves Chemin. Équations d’ondes quasilineaires et estimations de Strichartz. *Amer. J. Math.*, 121(6):1337–1377, 1999.
- [2] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2011.
- [3] Sylvie Benzoni-Gavage and Denis Serre. *Multidimensional hyperbolic partial differential equations*. The Clarendon Press, Oxford University Press, Oxford, 2007.
- [4] D. Boito, L. N. S. de Andrade, G. de Sousa, R. Gama, and C. Y. M. London. On Maxwell’s electrodynamics in two spatial dimensions. Preprint, arXiv:1809.07368, 2018.
- [5] Federico Cacciafesta and Anne-Sophie de Suzzoni. Weak dispersion for the Dirac equation on asymptotic flat and warped product spaces. *Discrete Contin. Dyn. Syst.*, 39(8):4359–4398, 2019.
- [6] Michael Christ and Alexander Kiselev. Maximal functions associated to filtrations. *J. Funct. Anal.*, 179(2):409–425, 2001.
- [7] Piero D’Ancona and Roland Schnaubelt. Global Strichartz estimates for an inhomogeneous Maxwell system. CRC 1173 Preprint 2020/28, Karlsruhe Institute of Technology, oct 2020.
- [8] Jean-Marc Delort. *F.B.I. transformation*, volume 1522 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1992.
- [9] Eric Dumas and Franck Sueur. Cauchy problem and quasi-stationary limit for the Maxwell-Landau-Lifschitz and Maxwell-Bloch equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 11(3):503–543, 2012.
- [10] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman lectures on physics. Vol. 2: Mainly electromagnetism and matter*. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1964.
- [11] Andrew Hassell, Terence Tao, and Jared Wunsch. Sharp Strichartz estimates on nontrapping asymptotically conic manifolds. *Amer. J. Math.*, 128(4):963–1024, 2006.
- [12] Lars Hörmander. *The analysis of linear partial differential operators. I*. Springer-Verlag, Berlin, 2003. Reprint of the second (1990) edition.
- [13] Lars Hörmander. *The analysis of linear partial differential operators. III*. Springer, Berlin, 2007. Reprint of the 1994 edition.
- [14] Mihaela Ifrim and Daniel Tataru. Local well-posedness for quasilinear problems: a primer. Preprint, arXiv:2008.05684, 2020.
- [15] L. V. Kapitanskiĭ. Some generalizations of the Strichartz-Brenner inequality. *Algebra i Analiz*, 1(3):127–159, 1989.
- [16] L. V. Kapitanskiĭ. Estimates for norms in Besov and Lizorkin-Triebel spaces for solutions of second-order linear hyperbolic equations. *J. Soviet Math.*, 56:2348–2389, 1991. Appeared 1989 in Russian.
- [17] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.
- [18] S. Klainerman. A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations. *Internat. Math. Res. Notices*, (5):221–274, 2001.
- [19] Sergiu Klainerman and Igor Rodnianski. Rough solutions of the Einstein-vacuum equations. *Ann. of Math. (2)*, 161(3):1143–1193, 2005.



- [20] L. D. Landau and E. M. Lifschitz. *Lehrbuch der theoretischen Physik ("Landau-Lifschitz")*. Band VIII. Akademie-Verlag, Berlin, fifth edition, 1990.
- [21] Otto Liess. Decay estimates for the solutions of the system of crystal optics. *Asymptotic Anal.*, 4(1):61–95, 1991.
- [22] A. Majda. *Compressible fluid flow and systems of conservation laws in several space variables*, volume 53 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1984.
- [23] Rainer Mandel and Robert Schippa. Time-harmonic solutions for maxwell's equations in anisotropic media and bochner-riesz estimates with negative index for non-elliptic surfaces. Preprint, arXiv:2103.17176, 2021.
- [24] Jeremy Marzuola, Jason Metcalfe, Daniel Tataru, and Mihai Tohaneanu. Strichartz estimates on Schwarzschild black hole backgrounds. *Comm. Math. Phys.*, 293(1):37–83, 2010.
- [25] Kirk McDonald. Electrodynamics in 1 and 2 spatial dimensions. Unpublished note, see <https://www.physics.princeton.edu/~mcdonald/examples/2dem.pdf>, 2019.
- [26] Jason Metcalfe, Daniel Tataru, and Mihai Tohaneanu. Pointwise decay for the Maxwell field on black hole space-times. *Adv. Math.*, 316:53–93, 2017.
- [27] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. *J. Amer. Math. Soc.*, 6(1):65–130, 1993.
- [28] Jerome Moloney and Alan Newell. *Nonlinear optics*. Westview Press. Boulder, CO, 2004.
- [29] Andreas Seeger, Christopher D. Sogge, and Elias M. Stein. Regularity properties of Fourier integral operators. *Ann. of Math. (2)*, 134(2):231–251, 1991.
- [30] Hart F. Smith. A parametrix construction for wave equations with  $C^{1,1}$  coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.
- [31] Hart F. Smith and Christopher D. Sogge. On Strichartz and eigenfunction estimates for low regularity metrics. *Math. Res. Lett.*, 1(6):729–737, 1994.
- [32] Hart F. Smith and Daniel Tataru. Sharp counterexamples for Strichartz estimates for low regularity metrics. *Math. Res. Lett.*, 9(2-3):199–204, 2002.
- [33] Christopher D. Sogge. *Fourier integrals in classical analysis*, volume 210 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 2017.
- [34] Martin Spitz. *Local wellposedness of nonlinear Maxwell equations*. PhD thesis, Karlsruhe Institute of Technology (KIT), 2017.
- [35] Martin Spitz. Local wellposedness of nonlinear Maxwell equations with perfectly conducting boundary conditions. *J. Differential Equations*, 266(8):5012–5063, 2019.
- [36] Terence Tao. Global regularity of wave maps. II. Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.
- [37] Daniel Tataru. Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation. *Amer. J. Math.*, 122(2):349–376, 2000.
- [38] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.
- [39] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III. *J. Amer. Math. Soc.*, 15(2):419–442, 2002.
- [40] Michael E. Taylor. *Pseudodifferential operators and nonlinear PDE*, volume 100 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1991.

*E-mail address:* robert.schippa@kit.edu

*E-mail address:* schnaubelt@kit.edu

DEPARTMENT OF MATHEMATICS, KARLSRUHE INSTITUTE OF TECHNOLOGY, ENGLERSTRASSE 2,  
76131 KARLSRUHE, GERMANY