

RECONSTRUCTION OF THE OBSERVABLE UNIVERSE FROM THE INTEGRATED SACHS-WOLFE EFFECT

JULIANNE CHUNG AND YIRAN WANG

ABSTRACT. The integrated Sachs-Wolfe (ISW) effect is a property of the Cosmic Microwave Background (CMB), in which photons from the CMB are gravitationally redshifted, causing the anisotropies in the CMB. An intriguing question is whether one can infer the gravitational perturbations from the ISW effect observed near the Earth. In this work, we address the question using a tomographic reconstruction approach, similar to X-ray CT reconstruction in medical imaging. We develop the mathematical analysis for the stable inversion of the X-ray transform in the cosmological setting. In addition, we provide a numerical study of reconstruction methods, thereby demonstrating the feasibility and potential of the tomography method.

CONTENTS

1. The physical background	1
2. The mathematical formulation and main results	4
3. Analysis of the backprojection	5
4. The microlocal inversion	11
5. Proof of Theorem 2.1	17
6. The numerical experiment	18
7. Conclusions	25
Acknowledgments	27
References	27

1. THE PHYSICAL BACKGROUND

This paper concerns the reconstruction of the initial status of the universe from the Cosmic Microwave Background (CMB). We start with the description of the physical problem. Consider the Friedman-Lemaître-Robertson-Walker (FLRW) model for the universe:

$$\mathcal{M} = (0, \infty) \times \mathbb{R}^3, \quad g_0 = -dt^2 + a^2(t)dx^2,$$

where $t \in (0, \infty), x \in \mathbb{R}^3$. The factor $a(t)$ is assumed to be positive and smooth in t . It represents the rate of expansion of the universe. We think of $(\mathcal{M}, g_0)$ as the background universe, which is isotropic and homogeneous. We note that the metric is conformal to the Minkowski metric. More precisely, after a change of variables from (t, x) to $(s, x), s > 0$ with $ds/dt = (a(t))^{-1}$, we can write the metric as $g_0 = a^2(s)(-ds^2 + dx^2)$.

Next, we consider the actual universe $(\mathcal{M}, g)$ as a perturbation of $(\mathcal{M}, g_0)$. The CMB perturbation theory is well-developed in the literature, see for instance [6, 8, 23]. Here, we

consider a universe governed by a scalar field φ . The stress energy tensor is

$$T^\mu{}_\nu = \nabla^\mu \varphi \nabla_\nu \varphi - \left[\frac{1}{2} \sum_{\alpha=0}^3 \nabla^\alpha \varphi \nabla_\alpha \varphi - V(\varphi) \right] \delta^\mu{}_\nu, \quad \mu, \nu = 0, 1, 2, 3,$$

see [23, equation (6.2)]. Here, V is the potential function for the scalar field φ . The field itself satisfies the Klein-Gordon equation $\square \varphi + \partial_\varphi V(\varphi) = 0$. According to Einstein's relativity theory, g satisfies the Einstein equation with the source T . We work with the linearized theory and assume that $\varphi = \varphi_0 + \delta\varphi$ where φ_0 is the scalar field which drives the background model and $\delta\varphi$ denotes the perturbation. Then we can split $T^\mu{}_\nu = {}^{(0)}T^\mu{}_\nu + \delta T^\mu{}_\nu$. From the linearized Einstein-Klein-Gordon equations, one deduces that the metric g up to the first order perturbation is of the form

$$(1.1) \quad g = a^2(s)[(1 + 2\Phi)ds^2 - (1 - 2\Phi)dx^2],$$

where Φ is a scalar functions on $\mathcal{M}$. In fact, $\Phi = \Psi$ satisfies the equation

$$(1.2) \quad \Phi'' + 2(H - \varphi_0''/\varphi_0')\Phi' - \Delta\Phi + 2(H' - H\varphi_0''/\varphi_0)\Phi = 0,$$

see [23, equation (6.48)]. Here, $H(s) = a'(s)/a(s)$ and $'$ denotes the derivative in s . Note that (1.2) is a damped wave equation. We remark that for a universe dominated by perfect fluids, there is a similar model where Φ satisfies the Bardeen's equation which is a wave equation with sound speed $c \in (0, 1)$, see for example [23, 28]. We consider the scalar field model because we are interested in finding gravitational waves which travel at the speed of light. The scalar field model can be regarded as a simplified problem, in addition to its own interest.

Now we consider the photon distribution in such a universe. It is known that trajectories of photons can be represented by light-like (or null) geodesics for the Lorentzian metric g on $\mathcal{M}$. Let $\mathcal{M}_0 = \{s_0\} \times \mathbb{R}^3$ be the surface of last scattering. This is the moment after which photons stopped interacting and started to travel freely in $\mathcal{M}$. Let $\mathcal{M}_1 = \{s_1\} \times \mathbb{R}^3$, $s_1 > s_0$ be the surface where we make observations of the photons or the CMB. We consider the photon energies observed at $\mathcal{M}_0, \mathcal{M}_1$,

$$E_0 = g_0(\dot{\gamma}(\tau_0), \partial_s) \quad \text{and} \quad E_1 = g_0(\dot{\gamma}(\tau_1), \partial_s).$$

Here, $\gamma(\tau)$, $\tau \in \mathbb{R}$ is a light-like geodesic with $\gamma(\tau_0) \in \mathcal{M}_0, \gamma(\tau_1) \in \mathcal{M}_1$. Also, the flow of the vector field ∂_s represents the observer. The redshift z is defined by

$$1 + z = E_1/E_0.$$

In [25], Sachs and Wolfe derived that to the first order linearization, the redshift is represented by an integral of the metric perturbations, see [25, equation (39)] and also the derivation in [19]. This is known as the integrated Sachs-Wolfe (ISW) effect. In our setup of scalar perturbations, the ISW is given by (up to some scalar factors),

$$(1.3) \quad \int \partial_s \Phi(\gamma(\tau)) d\tau$$

where $\gamma(\tau)$ is a light-like geodesic from $\mathcal{M}_0$ to $\mathcal{M}_1$. See also [8, Section 2.5] and [6].

It is clear from (1.3) that the ISW contains information of the gravitational perturbations. An outstanding question is whether one can extract the information. Since the discovery of the CMB by Penzias and Wilson in 1963, the measurements of the CMB have been significantly improved through multiple satellite projects such as COBE (Cosmic Background Explorer), WMAP (Wilkinson Microwave Anisotropy Probe), and the Planck Surveyor. Even

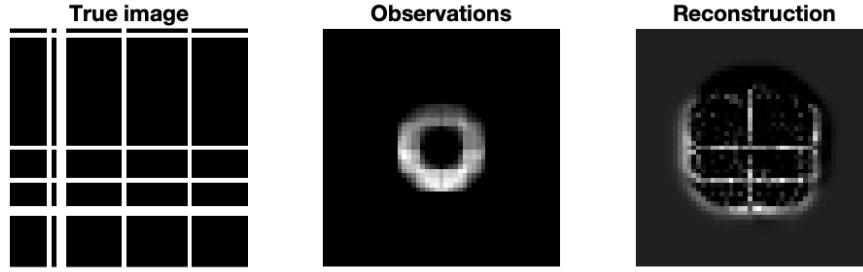


FIGURE 1. Cosmological X-ray tomography for a 2D Universe. The left figure is the true image which represents a collection of cosmic strings or gravitational plane waves at time $t = 0$. The middle figure shows the simulated ISW effect observed at $t = T$ by several detectors near the center of the image. One can see the “anisotropies” but the line structures are hardly discernible. The right figure shows the reconstruction using the tomography method in this paper. In the region given by Theorem 2.1, the line structures are clearly visible.

though ISW cannot be directly read off from the CMB data, there are many studies on extracting ISW from the CMB by combining it with other survey data. See [26, 7, 20, 22] for example. There is a huge literature on recovering cosmological parameters from the CMB and ISW, see [6, 8] and [15, 17] for example. Most of the existing work is based on the analysis of the CMB power spectrum and statistical inference methods. In this work, we treat the problem from a tomography point of view, which is very similar to X-ray computed tomography in medical imaging. One advantage of the tomography approach is the capability to reconstruct detailed object structures. To illustrate this point, we showcase the tomography reconstruction for a 2D universe in Figure 1 from one numerical simulation in Section 6.

The transform (1.3) (for tensors) appeared in [11] for the mathematical study of a cosmology problem. On page 186 of [11], Guillemin provisionally called the problem “cosmological X-ray tomography”. Recently, the relevance of the transform and the CMB inverse problem started to attract attention, see [31] for a survey. In particular, using microlocal methods, the authors of [19] demonstrated that time-like singularities (in the sense of wave front set) of Φ can be recovered from the ISW. See also [29] for related results on light-like singularities. Unfortunately, not all information can be stably recovered by using only the transform, which is an issue already pointed out in [11]. We refer to [5] for a recent numerical demonstration of the issue. A critical step was made in [28] that by taking into account the physical model, in particular the evolution equation (1.2), Φ can be stably recovered with the observation of (1.3) on a Cauchy surface, see also [30]. The result suggests that the tomography approach is promising, but to make it applicable, one must address the problem in a practical setting where the CMB is observed only *near* the Earth. The goal and novelty of this work is to address this problem by solving a cosmological X-ray tomography problem with partial data. We also conduct numerical studies to demonstrate the feasibility of the tomography method for cosmological applications.

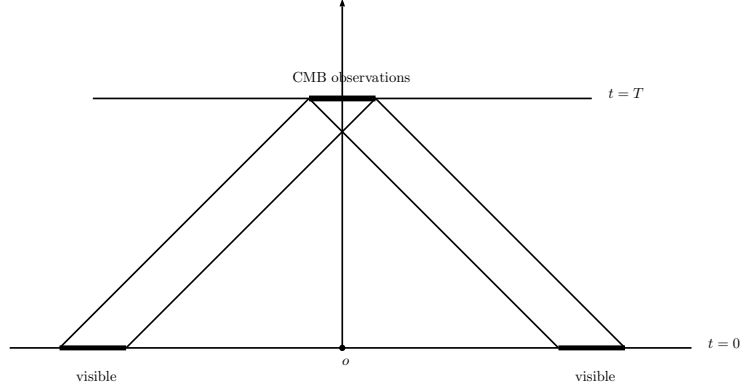


FIGURE 2. The setup of the partial data problem. The marked set on $t = 0$ is the visible set in Theorem 2.1.

2. THE MATHEMATICAL FORMULATION AND MAIN RESULTS

For $T > 0$, let $\mathcal{M} = [0, T] \times \mathbb{R}^3$ and $(t, x), t \in [0, T], x \in \mathbb{R}^3$ be the local coordinates. Let $g = -dt^2 + \sum_{i=1}^3 dx_i^2$ be the Minkowski metric on $\mathcal{M}$. We study the problem for a 3D universe for physical relevance, but remark that our method applies to dimensions $n \geq 2$. For $t \in \mathbb{R}$, we denote $\mathcal{M}_t = \{t\} \times \mathbb{R}^3$. Let $\mathcal{B} = \{x \in \mathbb{R}^3 : |x| < 1\}$ be the unit ball on $\mathbb{R}^3$. We consider null geodesics on $(\mathcal{M}, g)$ from $\mathcal{M}_0$ that meet $\mathcal{B}$ on $\mathcal{M}_T$. See Figure 2. This setup corresponds to making CMB observations near the Earth. It is convenient to parametrize the null geodesics using $y \in \mathcal{B}$ and $v \in \mathbb{S}^2$ as

$$(2.1) \quad \gamma_{y,v}(t) = (t, y + tv - Tv), \quad t \in \mathbb{R}.$$

For $t \in [0, T]$, $\gamma_{y,v}(t)$ is contained in $\mathcal{M}$. Note that the set of such null geodesics can be identified with the set $\mathcal{C} = \mathcal{B} \times \mathbb{S}^2$. For a scalar function f in $\mathcal{M}$, we consider the light ray transform of f defined as

$$(2.2) \quad Lf(y, v) = \int_0^T f(t, y + tv - Tv) dt,$$

which is a function on $\mathcal{C}$.

We study L acting on solutions to the Cauchy problem of wave equations on $\mathcal{M}$. Denote $\square = \partial_t^2 + \Delta$ where Δ is the positive Laplacian on $\mathbb{R}^3$, namely $\Delta = \sum_{i=1}^3 D_{x_i}^2, D_{x_i} = -i \frac{\partial}{\partial x_i}, i = \sqrt{-1}$. Let

$$(2.3) \quad P(t, x, \partial_t, D) = \square + a_0(t, x) \partial_t + \sum_{j=1}^3 a_j(t, x) D_j + b(t, x),$$

where $a_j, b \in C^\infty$. We consider the Cauchy problem

$$(2.4) \quad \begin{aligned} P(t, x, \partial_t, D)u(t, x) &= 0, \quad t > 0, x \in \mathbb{R}^3 \\ u(0, x) &= f_1(x), \quad \partial_t u(0, x) = f_2(x). \end{aligned}$$

It is well-known that under proper regularity assumptions on f_1, f_2 , there is a unique solution u . We study the inverse problem of recovering information of f_1, f_2 from Lu .

For the full data problem namely $\mathcal{B} = \mathbb{R}^3$, it was shown in [28] (see also [30]) that one can stably recover f_1, f_2 supported in a fixed compact set. The partial data problem we are considering is different and it is not always possible to recover f_1, f_2 . Consider the setup in

Figure 2. Let $\mathcal{B}_{1+T} = \{x \in \mathbb{R}^3 : |x| < 1+T\}$ on $\mathcal{M}_0$. If f_1, f_2 are supported outside of $\mathcal{B}_{1+T}$, by finite speed of propagation of (2.4), we see that $\text{supp } u$ does not meet any light rays that intersect $\mathcal{B}$ at $t = T$. So $Lu = 0$. Thus the best one can hope for is to recover f_1, f_2 in $\mathcal{B}_{1+T}$.

Our result concerns the information that can be recovered stably in the microlocal sense. For $b > a > 0$, we set

$$(2.5) \quad \mathcal{R}_{a,b} = \{x \in \mathbb{R}^3 : a < |x| < b\}.$$

We use the convention that if $a \leq 0$ then $\mathcal{R}_{a,b} = \mathcal{B}_b$. We define $\mathcal{W} = \mathcal{W}^+ \cup \mathcal{W}^-$ where

$$(2.6) \quad \mathcal{W}^\pm = \{(x, \xi) \in T^*\mathbb{R}^3 \setminus 0 : x \mp T\xi/|\xi| \in \mathcal{B}\}.$$

Our main result is

Theorem 2.1. *Let $s \geq 0$ be an integer. Suppose $f_1 \in H^{s+1}(\mathbb{R}^3)$, $f_2 \in H^s(\mathbb{R}^3)$ are compactly supported in $\mathcal{R}_{T-1, T+1}$. Then f_1, f_2 are uniquely determined by Lu . In addition, if $f_2 = 0$ and the wave front set $WF(f_1) \subset \mathcal{W}$, we have the following stability estimates*

$$(2.7) \quad \|f_1\|_{H^{s+1}(\mathbb{R}^3)} \leq C \|Lu\|_{H^{s+2}(\mathcal{C})}$$

for some constant $C > 0$ uniform for f_1 supported in a fixed compact set of $\mathcal{R}_{T-1, T+1}$.

In Theorem 4.3, we will give a more precise but technical statement on recovering singularities of f_1, f_2 even if they are not supported in $\mathcal{R}_{T-1, T+1}$.

The rest of the paper is organized as follows. In Section 3, we study the microlocal properties of the back-projection. In Section 4, we construct the microlocal inversion for the back-projection. Then we prove Theorem 2.1 in Section 5. Finally, we conduct a numerical study in Section 6, and conclusions are provided in Section 7.

3. ANALYSIS OF THE BACKPROJECTION

We start with the expression of the solution for the Cauchy problem (2.4). It is well-known that one can find approximate solutions as oscillatory integrals, see for example [27, Section 1, Chapter VI]. Let $E_j(t), j = 1, 2$ be the fundamental solutions such that $P(t, x, \partial_t, D)E_j(t) = 0$ for $t \in (0, T)$ and

$$\begin{aligned} E_1(0) &= \text{Id}, \quad \partial_t E_1(0) = 0, \\ E_2(0) &= 0, \quad \partial_t E_2(0) = \text{Id}, \end{aligned}$$

where Id denotes the identity operator. Then we can write the solution of (2.4) as $u(t, x) = E_1(t)f_1 + E_2(t)f_2$. In fact, $E_j(t), j = 1, 2$ are of the form

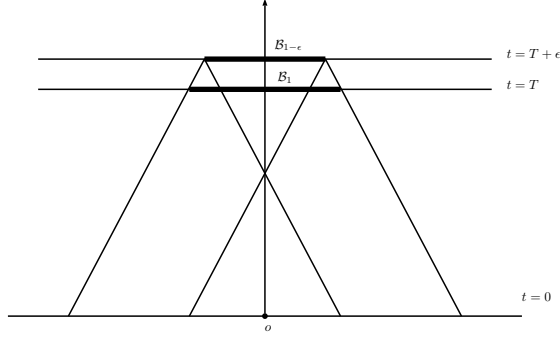
$$(3.1) \quad E_j(t)u(x) = E_j^+(t)u(x) + E_j^-(t)u(x) + R_j(t)u(x)$$

where

$$(3.2) \quad \begin{aligned} E_j^+(t)u(x) &= (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i((x-y) \cdot \xi + t|\xi|)} a_j^+(t, x, \xi) u(y) dy d\xi \\ E_j^-(t)u(x) &= (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i((x-y) \cdot \xi - t|\xi|)} a_j^-(t, x, \xi) u(y) dy d\xi \end{aligned}$$

and $R_j(t), j = 1, 2$ are regularizing, see [27, Chapter VI, (1.37)]. Here, the phase function is derived from the principal part of $P(t, x, \partial_t, D)$. The amplitudes $a_j^\pm$ have asymptotic summations of the form

$$a_j^\pm(t, x, \xi) \sim \sum_{k=0}^{\infty} a_{jk}^\pm(t, x, \xi),$$

FIGURE 3. Illustration of the modified light ray transform L_ϵ .

where each $a_{jk}^\pm$ is homogeneous of degree $-j-k$. They are determined via transport equations. Here, we recall the construction of the leading order terms for $k = 0$. They satisfy

$$(3.3) \quad \partial_t a_{j0}^\pm \mp \sum_{l=1}^3 \frac{\xi_l}{|\xi|} \partial_l a_{j0}^\pm \pm \frac{1}{2} (a_0(t, x) + \sum_{l=1}^3 a_l(t, x) \frac{\xi_l}{|\xi|}) a_{j0}^\pm = 0,$$

see [27, Chapter VI, (1.49)]. The initial condition at $t = 0$ is given by (see [27, Chapter VI, (1.53)])

$$(\imath|\xi|)^{j'-1} a_{j0}^+(0, x, \xi) + (-\imath|\xi|)^{j'-1} a_{j0}^-(0, x, \xi) = \delta_{jj'},$$

where $j, j' = 1, 2$ and $\delta_{jj'}$ is the Kronecker delta function. We find that

$$a_{10}^\pm(0, x, \xi) = \frac{1}{2}, \quad a_{20}^\pm(0, x, \xi) = \pm \frac{1}{2\imath} |\xi|^{-1}.$$

Then we can solve (3.3) with these initial conditions to get $a_{j0}^\pm(t, x, \xi)$. In particular, we see that $a_{j0}^\pm$ are non-vanishing.

To summarize, for $t > 0$, $E_j^\pm(t) \in I^{-j+1}(\mathbb{R}^3 \times \mathbb{R}^3; (C^\pm)')$ are elliptic Fourier integral operators in Hörmander's notation (see Definition 25.2.1 of [14]), where the canonical relations are

$$(3.4) \quad C_t^\pm = \{(x, \eta; y, \xi) \in T^*\mathbb{R}^3 \setminus 0 \times T^*\mathbb{R}^3 \setminus 0 : x = y - t(\pm\xi/|\xi|), \eta = \xi\}.$$

Finally, we write the approximate solution of (2.4) as

$$(3.5) \quad u(t, x) = E_1^+(t)f_1 + E_1^-(t)f_1 + E_2^+(t)f_2 + E_2^-(t)f_2 + R_1(t)f_1 + R_2(t)f_2$$

where $t \geq 0$.

We analyze Lu with u in (3.5). It suffices to consider the operator $(LE_j^\pm)$, $j = 1, 2$ defined by $(LE_j^\pm)f(x, v) = L(E_j^\pm(t)f)(x, v)$ for a scalar function f on $\mathbb{R}^3$. In order to “invert” the operators to get f , we use the idea of backprojection in X-ray tomography. In particular, we follow the approach of [28] to take the integral of the v variable on $\mathbb{S}^2$. For $h \in C^\infty(\mathbb{R}^3 \times \mathbb{S}^2)$, we define

$$Ih(x) = \int_{\mathbb{S}^2} h(x, v) dv.$$

The main purpose of this section is to analyze the microlocal structure of $ILLE_j^\pm$ and show that it can be decomposed into a sum of FIOs. The main results of this section are Lemma 3.1 and 3.2.

To avoid some technicalities, we will use a subset of the light rays. For $\epsilon > 0$ small, we consider the set $\mathcal{B}_{1-\epsilon} = \{y \in \mathbb{R}^3 : |y| < 1 - \epsilon\}$. Then we use the light rays that meet $\{T + \epsilon\} \times \mathcal{B}_{1-\epsilon}$, see Figure 3. We let φ_ϵ be a smooth cut-off function on $\mathbb{R}^3$ such that $\text{supp } \varphi_\epsilon = \overline{\mathcal{B}_{1-\epsilon}}$, $\varphi_\epsilon > 0$ on $\mathcal{B}_{1-\epsilon}$ and $\varphi_\epsilon = 1$ on $\overline{\mathcal{B}_{1-2\epsilon}}$. Then we let $L_\epsilon = \varphi_\epsilon L$ so that

$$(3.6) \quad L_\epsilon f(y, v) = \int_0^T \varphi_\epsilon(y) f(t, y + tv - (T + \epsilon)v) dt, \quad y \in \mathcal{B}_1, v \in \mathbb{S}^2.$$

Note that when $t = T$, $y + tv - (T + \epsilon)v = y - \epsilon v \in \mathcal{B}$ for $y \in \mathcal{B}_{1-\epsilon}$. The role of φ_ϵ is to avoid the cut-off singularities at the boundary of $\mathcal{B}_1$.

We now analyze $IL_\epsilon E_1^+$. In particular, we will find the integral kernel of the operator in the sense of distributions. Using (3.2), we have

$$IL_\epsilon E_1^+ f(x) = (2\pi)^{-3} \int_{\mathbb{S}^2} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(t-T-\epsilon)v \cdot \xi} e^{i t |\xi|} a_1^+(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt dv.$$

Note that for $t \in (0, T)$, $t - T - \epsilon < 0$. For $|\xi|$ large, the phase function

$$(t - T - \epsilon)v \cdot \xi = (t - T - \epsilon)|\xi|v \cdot (\xi/|\xi|)$$

has non-degenerate critical points at $v = \pm \xi/|\xi|$, see e.g., [21, (1.14)]. We can perform the stationary phase method for the integration in v to get $IL_\epsilon E_1^+ = B_{1,1}^+ + B_{1,2}^+ + R_1^+$ where R_1^+ is regularizing and

$$(3.7) \quad \begin{aligned} B_{1,1}^+ f(x) &= (2\pi)^{-3} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} b_{1,1}^+(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt, \\ B_{1,2}^+ f(x) &= (2\pi)^{-3} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} e^{i2|\xi|t} b_{1,2}^+(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt. \end{aligned}$$

Here, $b_{1,k}^+$, $k = 1, 2$ are smooth functions supported away from $\xi = 0$. They have asymptotic summations of the form

$$b_{1,k}^+(t, x, \xi) \sim \sum_{l=1}^{\infty} |\xi|^{-l} b_{1,k;l}^+(t, x, \xi/|\xi|)$$

for $|\xi|$ large. The leading order terms are

$$(3.8) \quad \begin{aligned} \sigma_{-1}(b_{1,1}^+)(t, x, \xi) &= C_1 \frac{1}{(t - T - \epsilon)|\xi|} a_{10}^+(t, x, \xi/|\xi|) = \frac{C_1}{2(t - T - \epsilon)|\xi|}, \\ \sigma_{-1}(b_{1,2}^+)(t, x, \xi) &= C_2 \frac{1}{(t - T - \epsilon)|\xi|} a_{10}^+(t, x, -\xi/|\xi|) = \frac{C_2}{2(t - T - \epsilon)|\xi|}, \end{aligned}$$

where we used $\sigma_{-1}(\cdot)$ to denote the leading order term or the principal symbol of order -1 , and C_1, C_2 come from the constants in the stationary phase argument, see [21, (1.14)]. More explicitly,

$$C_1 = e^{-\pi i/2} (2\pi)^2 = -4\pi^2 i, \quad C_2 = e^{\pi i/2} (2\pi)^2 = 4\pi^2 i.$$

Now we can write $B_{1,1}^+$ as

$$(3.9) \quad B_{1,1}^+ f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{1,1}^+(x, \xi) f(y) dy d\xi$$

where

$$\tilde{b}_{1,1}^+(x, \xi) = \int_0^T b_{1,1}^+(t, x, \xi) \varphi_\epsilon(x) dt$$

is a symbol of order -1 . The leading order term is given by

$$(3.10) \quad \sigma_{-1}(\tilde{b}_{1,1}^+)(x, \xi) = |\xi|^{-1} \varphi_\epsilon(x) \int_0^T \frac{-4\pi^2 i}{2(t-T-\epsilon)} dt = (-2\pi^2 i \ln \frac{\epsilon}{T+\epsilon}) |\xi|^{-1} \varphi_\epsilon(x)$$

for $|\xi|$ large. We observe that $B_{1,1}^+$ is an Fourier integral operator of order -1 with canonical relation $C_{T+\epsilon}$. Hereafter, for $a \in \mathbb{R}$, we define

$$C_a = \{(x, \xi; y, \eta) : \xi = \eta, x = y - a\xi/|\xi|, x, y \in \mathbb{R}^3, \xi, \eta \in \mathbb{R}^3 \setminus \{0\}\}.$$

In Hörmander's notation, we can write $B_{1,1}^+ \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon})$.

For $B_{1,2}^+$ in (3.7), we first compute

$$\begin{aligned} \tilde{b}_{1,2}^+(x, \xi) &= \int_0^T e^{i2|\xi|t} b_{1,2}^+(t, x, \xi) \varphi_\epsilon(x) dt = \frac{1}{i2|\xi|} b_{1,2}^+(T, x, \xi) \varphi_\epsilon(x) e^{i2|\xi|T} \\ &\quad - \frac{1}{i2|\xi|} b_{1,2}^+(0, x, \xi) \varphi_\epsilon(x) - \int_0^T \frac{1}{i2|\xi|} \partial_t b_{1,2}^+(t, x, \xi) \varphi_\epsilon(x) e^{i2|\xi|t} dt \end{aligned}$$

via integration by parts. This procedure can be continued to get

$$\tilde{b}_{1,2}^+(x, \xi) = \tilde{b}_{1,2,1}^+(x, \xi) + e^{i2|\xi|T} \tilde{b}_{1,2,2}^+(x, \xi)$$

where $\tilde{b}_{1,2,j}^+, j = 1, 2$ are symbols of order -2 . In fact, we can find the leading order terms as

$$\sigma_{-2}(\tilde{b}_{1,2,1}^+)(x, \xi) = \frac{\pi^2}{(T+\epsilon)} |\xi|^{-2} \varphi_\epsilon(x), \quad \sigma_{-2}(\tilde{b}_{1,2,2}^+)(x, \xi) = -\frac{\pi^2}{\epsilon} |\xi|^{-2} \varphi_\epsilon(x).$$

Therefore, we have

$$\begin{aligned} B_{1,2}^+ f(x) &= (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{1,2,1}^+(x, \xi) f(y) dy d\xi \\ &\quad + (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T-\epsilon)|\xi|} \tilde{b}_{1,2,2}^+(x, \xi) f(y) dy d\xi = B_{1,2,1}^+ f(x) + B_{1,2,2}^+ f(x). \end{aligned}$$

We observe that the operators $B_{1,2,k}^+, k = 1, 2$ are FIOs of order -2 with canonical relations $C_{-T-\epsilon}$ and $C_{T-\epsilon}$. This completes the analysis for $IL_\epsilon E_1^+$. In particular, we decomposed the operator into a sum of FIOs.

The analysis for $IL_\epsilon E_2^+$ is completely identical. The only difference is the order and the symbol. We will not show the details. To summarize, we proved

Lemma 3.1. *For $l = 1, 2$, we can write*

$$IL_\epsilon E_l^+ = B_{l,1}^+ + B_{l,2,1}^+ + B_{l,2,2}^+ + R_l^+,$$

where R_l^+ are regularizing operators and the rest are Fourier integral operators

$$\begin{aligned} B_{l,1}^+ &\in I^{-1-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ B_{l,2,1}^+ &\in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \quad B_{l,2,2}^+ \in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon}). \end{aligned}$$

Moreover,

(1) $B_{l,1}^+, l = 1, 2$ can be written as an oscillatory integrals

$$B_{l,1}^+ f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{l,1}^+(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,1}^+$ are symbols of order $-1 - (l - 1)$. The principal symbols of $B_{l,1}^+$ in this representation are given by

$$\begin{aligned}\sigma_{-1}(B_{1,1}^+)(x, y, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T + \epsilon}) \varphi_\epsilon(x) |\xi|^{-1}, \\ \sigma_{-2}(B_{2,1}^+)(x, y, \xi) &= (-2\pi^2 \ln \frac{\epsilon}{T + \epsilon}) \varphi_\epsilon(x) |\xi|^{-2}\end{aligned}$$

for $|\xi|$ large.

(2) $B_{l,2,1}^+, l = 1, 2$ can be written as an oscillatory integrals

$$B_{l,2,1}^+ f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{l,2,1}^+(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,1}^+$ are symbols of order $-2 - (l - 1)$. The principal symbols of $B_{l,2,1}^+$ in this representation are given by

$$\sigma_{-2}(B_{1,2,1}^+)(x, y, \xi) = \frac{\pi^2}{(T + \epsilon)} \varphi_\epsilon(x) |\xi|^{-2}, \quad \sigma_{-3}(B_{2,2,1}^+)(x, y, \xi) = \frac{\pi^2}{i(T + \epsilon)} \varphi_\epsilon(x) |\xi|^{-3}$$

for $|\xi|$ large.

(3) $B_{l,2,2}^+, l = 1, 2$ can be written as an oscillatory integrals

$$B_{l,2,2}^+ f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T-\epsilon)|\xi|} \tilde{b}_{l,2,2}^+(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,2}^+$ are symbols of order $-2 - (l - 1)$. The principal symbols of $B_{l,2,2}^+$ in this representation are given by

$$\sigma_{-2}(B_{1,2,2}^+)(x, y, \xi) = -\frac{\pi^2}{\epsilon} \varphi_\epsilon(x) |\xi|^{-2}, \quad \sigma_{-3}(B_{2,2,2}^+)(x, y, \xi) = -\frac{\pi^2}{\epsilon i} \varphi_\epsilon(x) |\xi|^{-3}$$

for $|\xi|$ large.

We can obtain a similar decomposition for $IL_\epsilon E_j^-$ following the same type of calculation. Because the canonical relations are different, we repeat some of the calculations. Consider $IL_\epsilon E_1^-$. We compute that

$$IL_\epsilon E_1^- f(x) = (2\pi)^{-3} \int_{\mathbb{S}^2} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(t-T-\epsilon)v \cdot \xi} e^{-it|\xi|} a_1^-(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt dv.$$

Again, for $|\xi|$ large, we perform the stationary phase method for the integration in v to get $IL_\epsilon E_1^- = B_{1,1}^- + B_{1,2}^- + R_1^-$ where R_1^- is regularizing and

$$\begin{aligned}(3.11) \quad B_{1,1}^- f(x) &= (2\pi)^{-3} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} b_{1,1}^-(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt, \\ B_{1,2}^- f(x) &= (2\pi)^{-3} \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} e^{-i2|\xi|t} b_{1,2}^-(t, x, \xi) \varphi_\epsilon(x) f(y) dy d\xi dt.\end{aligned}$$

Similarly, $b_{1,k}^-, k = 1, 2$ are smooth functions supported away from $\xi = 0$. They have asymptotic summations of the form

$$b_{1,k}^-(t, x, \xi) \sim \sum_{l=1}^{\infty} |\xi|^{-l} b_{1,k,l}^-(t, x, \xi/|\xi|)$$

for $|\xi|$ large. The leading order terms are

$$(3.12) \quad \begin{aligned} \sigma_{-1}(b_{1,1})(t, x, \xi) &= C_1 \frac{1}{(t - T - \epsilon)|\xi|} a_{10}^-(t, x, \xi/|\xi|) = \frac{C_1}{2(t - T - \epsilon)|\xi|}, \\ \sigma_{-1}(b_{1,2})(t, x, \xi) &= C_2 \frac{1}{(t - T - \epsilon)|\xi|} a_{10}^-(t, x, -\xi/|\xi|) = \frac{C_2}{2(t - T - \epsilon)|\xi|}. \end{aligned}$$

We integrate in t in $B_{1,1}^-$ to get

$$B_{1,1}^- f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{1,1}^-(x, \xi) f(y) dy d\xi,$$

where

$$\tilde{b}_{1,1}^-(x, \xi) = \int_0^T |\xi|^{-1} b_{1,1}^-(t, x, \xi) \varphi_\epsilon(x) dt$$

is a symbol of order -1 . For $B_{1,2}^-$, we obtain via integration by parts that

$$\begin{aligned} \tilde{b}_{1,2}^-(x, \xi) &= \int_0^T e^{-i2|\xi|t} |\xi|^{-1} b_{1,2}^-(t, x, \xi) \varphi_\epsilon(x) dt \\ &= \frac{1}{-i2|\xi|^2} b_{1,2}^-(T, x, \xi) \varphi_\epsilon(x) e^{-i2|\xi|T} - \frac{1}{-i2|\xi|^2} b_{1,2}^-(0, x, \xi) \varphi_\epsilon(x) \\ &\quad - \int_0^T \frac{1}{-i2|\xi|^2} \partial_t b_{1,2}^-(t, x, \xi) \varphi_\epsilon(x) e^{-i2|\xi|t} dt. \end{aligned}$$

The procedure can be continued to yield

$$\tilde{b}_{1,2}^-(x, \xi) = \tilde{b}_{1,2,1}^-(x, \xi) + e^{-i2|\xi|T} \tilde{b}_{1,2,2}^-(x, \xi),$$

where $\tilde{b}_{1,2,j}^-, j = 1, 2$ are symbols of order -2 . Therefore, we have

$$\begin{aligned} B_{1,2}^- f(x) &= (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T-\epsilon)|\xi|} \tilde{b}_{1,2,1}^-(x, \xi) f(y) dy d\xi \\ &\quad + (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{1,2,2}^-(x, \xi) f(y) dy d\xi = B_{1,2,1}^- f(x) + B_{1,2,2}^- f(x). \end{aligned}$$

To summarize, we proved

Lemma 3.2. *For $l = 1, 2$, we can write*

$$IL_\epsilon E_l^- = B_{l,1}^- + B_{l,2,1}^- + B_{l,2,2}^- + R_l^-,$$

where R_l^- are regularizing operators and the rest are elliptic Fourier integral operators

$$\begin{aligned} B_{l,1}^- &\in I^{-1-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \\ B_{l,2,1}^- &\in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon}), \quad B_{l,2,2}^- \in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}). \end{aligned}$$

Moreover,

(1) $B_{l,1}^-, l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,1}^- f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{l,1}^-(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,1}^-$ are symbols of order $-1 - (l - 1)$. The principal symbols of $B_{l,1}^-$ in this representation are given by

$$\begin{aligned}\sigma_{-1}(B_{1,1}^-)(x, y, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T + \epsilon}) |\xi|^{-1} \varphi_\epsilon(x), \\ \sigma_{-2}(B_{2,1}^-)(x, y, \xi) &= (2\pi^2 \ln \frac{\epsilon}{T + \epsilon}) |\xi|^{-2} \varphi_\epsilon(x),\end{aligned}$$

for $|\xi|$ large.

(2) $B_{l,2,1}^-$, $l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,2,1}^- f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(-T+\epsilon)|\xi|} \tilde{b}_{l,2,1}^-(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,1}^-$ are symbols of order $-2 - (l - 1)$. The principal symbols of $B_{l,2,1}^-$ in this representation are given by

$$\sigma_{-2}(B_{1,2,1}^-)(x, y, \xi) = -\frac{\pi^2}{(T + \epsilon)} \varphi_\epsilon(x) |\xi|^{-2}, \quad \sigma_{-3}(B_{2,2,1}^-)(x, y, \xi) = -\frac{\pi^2 i}{(T + \epsilon)} \varphi_\epsilon(x) |\xi|^{-3},$$

for $|\xi|$ large.

(3) $B_{l,2,2}^-$, $l = 1, 2$ can be written as an oscillatory integrals

$$B_{l,2,2}^- f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{l,2,2}^-(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,2}^-$ are symbols of order $-2 - (l - 1)$. The principal symbols of $B_{l,2,2}^-$ in this representation are given by

$$\sigma_{-2}(B_{1,2,2}^-)(x, y, \xi) = \frac{\pi^2}{\epsilon} \varphi_\epsilon(x) |\xi|^{-2}, \quad \sigma_{-3}(B_{2,2,2}^-)(x, y, \xi) = \frac{\pi^2 i}{\epsilon} \varphi_\epsilon(x) |\xi|^{-3},$$

for $|\xi|$ large.

4. THE MICROLOCAL INVERSION

For $\epsilon > 0$ small, we consider $IL_\epsilon u$ with u in (3.5). We apply Lemma 3.1 and 3.2 and group the terms as follows,

$$(4.1) \quad \begin{aligned}IL_\epsilon u &= \tilde{B}_1^+ f_1 + \tilde{B}_2^+ f_2 + B_{1,2,2}^+ f_1 + B_{2,2,2}^+ f_2 \\ &+ \tilde{B}_1^- f_1 + \tilde{B}_2^- f_2 + B_{1,2,1}^- f_1 + B_{2,2,1}^- f_2 + R_1 f_1 + R_2 f_2,\end{aligned}$$

where $R_1 = R_1^+ + R_1^-$, $R_2 = R_2^+ + R_2^-$ are regularizing operators and

$$(4.2) \quad \begin{aligned}\tilde{B}_1^+ &= B_{1,1}^+ + B_{1,2,2}^- \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ \tilde{B}_2^+ &= B_{2,1}^+ + B_{2,2,2}^- \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ \tilde{B}_1^- &= B_{1,1}^- + B_{1,2,1}^+ \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \\ \tilde{B}_2^- &= B_{2,1}^- + B_{2,2,1}^+ \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}).\end{aligned}$$

Also, $B_{1,2,2}^+ \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$, $B_{2,2,2}^+ \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$ and $B_{1,2,1}^- \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$, $B_{2,2,1}^- \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$. Note that the orders of these FIOs acting on f_1, f_2 are lower than those corresponding terms in (4.2), even though the canonical relations are different. The operators appearing in (4.1) are FIOs of graph type. Hence in principle, we can use FIO calculus (see for instance [14, Section 25.3]) to solve for f_1, f_2 . However, equation

(4.1) alone is not enough to solve for both f_1 and f_2 . We will obtain another equation by changing the back-projection operator I , similar to the approach in [28].

Let φ be a function on $\mathbb{S}^2$. Then for $h \in C^\infty(\mathbb{R}^3 \times \mathbb{S}^2)$, we define

$$I^\varphi h(x) = \int_{\mathbb{S}^2} \varphi(v) h(x, v) dv.$$

We consider the back-projection $I^\varphi L_\epsilon u$. Ideally, we would like to take an odd function φ such that $\varphi(-v) = -\varphi(v)$, $v \in \mathbb{S}^2$ but then φ would vanish somewhere on $\mathbb{S}^2$ so we proceed as follows. Let $x = (x_1, x_2, x_3)$ be the coordinate for $\mathbb{R}^3$. For $\delta > 0$, let $\mathcal{U}_k = \{v : v = (x_1, x_2, x_3), \|x\| = 1, |x_k| > \delta/2\}, k = 1, 2, 3$. For δ sufficiently small, $\mathcal{U}_k, k = 1, 2, 3$ form an open covering of $\mathbb{S}^2$. Let $\chi_k(v), k = 1, 2, 3$ be a partition of unity subordinated to this covering and $\chi_k(v) = 1$ on $\mathcal{V}_k = \{v : v = (x_1, x_2, x_3), \|x\| = 1, |x_k| > \delta\}, k = 1, 2, 3$. Here, by possibly taking δ smaller, we can assume that $\mathcal{V}_k$ also forms an open covering of $\mathbb{S}^2$. For $v \in \mathbb{S}^2$, we let

$$\varphi_k(v) = \chi_k(x) x_k + 2, \quad k = 1, 2, 3.$$

Then $\varphi_k(v) \neq 0$ and $\varphi_k(-v) - \varphi_k(v) \neq 0$ for $v \in \mathcal{U}_k$. We now analyze $I^{\varphi_k} L_\epsilon u$ as in Section 3. By repeating the calculations in Lemma 3.1 and 3.2, we have the following Lemmas 4.1 and 4.2. The proofs are omitted.

Lemma 4.1. *For $k = 1, 2, 3$, and $l = 1, 2$, we can write*

$$I^{\varphi_k} L_\epsilon E_l^+ = B_{l,1}^{\varphi_k,+} + B_{l,2,1}^{\varphi_k,+} + B_{l,2,2}^{\varphi_k,+} + R_l^{\varphi_k,+},$$

where $R_l^{\varphi_k,+}$ are regularizing operators and the rest are Fourier integral operators

$$\begin{aligned} B_{l,1}^{\varphi_k,+} &\in I^{-1-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ B_{l,2,1}^{\varphi_k,+} &\in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \quad B_{l,2,2}^{\varphi_k,+} \in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon}). \end{aligned}$$

Moreover,

- (1) $B_{l,1}^{\varphi_k,+}, l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,1}^{\varphi_k,+} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{l,1}^{\varphi_k,+}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,1}^{\varphi_k,+}$ are symbols of order $-1 - (l-1)$. The principal symbols of $B_{l,1}^{\varphi_k,+}$ in this representation are given by

$$\begin{aligned} \sigma_{-1}(B_{1,1}^{\varphi_k,+})(x, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T+\epsilon}) \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-1}, \\ \sigma_{-2}(B_{2,1}^{\varphi_k,+})(x, \xi) &= (-2\pi^2 \ln \frac{\epsilon}{T+\epsilon}) \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-2}, \end{aligned}$$

for $|\xi|$ large.

- (2) $B_{l,2,1}^{\varphi_k,+}, l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,2,1}^{\varphi_k,+} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{l,2,1}^{\varphi_k,+}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,1}^{\varphi_k,+}$ are symbols of order $-2 - (l-1)$. The principal symbols of $B_{l,2,1}^{\varphi_k,+}$ in this representation are given by

$$\begin{aligned}\sigma_{-2}(B_{1,2,1}^{\varphi_k,+})(x, y, \xi) &= -\frac{\pi^2}{(T+\epsilon)}\varphi_\epsilon(x)\varphi_k(\xi/|\xi|)|\xi|^{-2}, \\ \sigma_{-3}(B_{2,2,1}^{\varphi_k,+})(x, y, \xi) &= -\frac{\pi^2}{i(T+\epsilon)}\varphi_\epsilon(x)\varphi_k(\xi/|\xi|)|\xi|^{-3}\end{aligned}$$

for $|\xi|$ large.

(3) $B_{l,2,2}^{\varphi_k,+}$, $l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,2,2}^{\varphi_k,+} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y)\cdot\xi} e^{i(T-\epsilon)|\xi|} \tilde{b}_{l,2,1}^{\varphi_k,+}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,2}^{\varphi_k,+}$ are symbols of order $-2 - (l-1)$. The principal symbols of $B_{l,2,2}^{\varphi_k,+}$ in this representation are given by

$$\begin{aligned}\sigma_{-2}(B_{1,2,2}^{\varphi_k,+})(x, y, \xi) &= -\frac{\pi^2}{\epsilon}\varphi_\epsilon(x)\varphi_k(\xi/|\xi|)|\xi|^{-2}, \\ \sigma_{-3}(B_{2,2,2}^{\varphi_k,+})(x, y, \xi) &= -\frac{\pi^2}{i\epsilon}\varphi_\epsilon(x)\varphi_k(\xi/|\xi|)|\xi|^{-3}\end{aligned}$$

for $|\xi|$ large.

Lemma 4.2. For $k = 1, 2, 3$, and $l = 1, 2$, we can write

$$I^{\varphi_k} L_\epsilon E_l^- = B_{l,1}^{\varphi_k,-} + B_{l,2,1}^{\varphi_k,-} + B_{l,2,2}^{\varphi_k,-} + R_l^{\varphi_k,-},$$

where $R_l^{\varphi_k,-}$ are regularizing operators and the rest are elliptic Fourier integral operators

$$\begin{aligned}B_{l,1}^{\varphi_k,-} &\in I^{-1-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \\ B_{l,2,1}^{\varphi_k,-} &\in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon}), \quad B_{l,2,2}^{\varphi_k,-} \in I^{-2-(l-1)}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}).\end{aligned}$$

Moreover,

(1) $B_{l,1}^{\varphi_k,-}$, $l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,1}^{\varphi_k,-} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y)\cdot\xi} e^{-i(T+\epsilon)|\xi|} \tilde{b}_{l,1}^{\varphi_k,-}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,1}^{\varphi_k,-}$ are symbols of order $-1 - (l-1)$. The principal symbols of $B_{l,1}^{\varphi_k,-}$ in this representation are given by

$$\begin{aligned}\sigma_{-1}(B_{1,1}^{\varphi_k,-})(x, y, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T+\epsilon}) |\xi|^{-1} \varphi_k(\xi/|\xi|) \varphi_\epsilon(x), \\ \sigma_{-2}(B_{2,1}^{\varphi_k,-})(x, y, \xi) &= (2\pi^2 \ln \frac{\epsilon}{T+\epsilon}) |\xi|^{-2} \varphi_k(\xi/|\xi|) \varphi_\epsilon(x),\end{aligned}$$

for $|\xi|$ large.

(2) $B_{l,2,1}^{\varphi_k,-}$, $l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,2,1}^{\varphi_k,-} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y)\cdot\xi} e^{i(-T+\epsilon)|\xi|} \tilde{b}_{l,2,1}^{\varphi_k,-}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,1}^{\varphi_k,-}$ are symbols of order $-2 - (l-1)$. The principal symbols of $B_{l,2,1}^{\varphi_k,-}$ in this representation are given by

$$\begin{aligned}\sigma_{-2}(B_{1,2,1}^{\varphi_k,-})(x, y, \xi) &= -\frac{\pi^2}{(T+\epsilon)} \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-2}, \\ \sigma_{-3}(B_{2,2,1}^{\varphi_k,-})(x, y, \xi) &= -\frac{\pi^2 \iota}{(T+\epsilon)} \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-3},\end{aligned}$$

for $|\xi|$ large.

(3) $B_{l,2,2}^{\varphi_k,-}$, $l = 1, 2$ can be written as oscillatory integrals,

$$B_{l,2,2}^{\varphi_k,-} f(x) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(x-y) \cdot \xi} e^{i(T+\epsilon)|\xi|} \tilde{b}_{l,2,1}^{\varphi_k,-}(x, \xi) f(y) dy d\xi,$$

where $\tilde{b}_{l,2,2}^{\varphi_k,-}$ are symbols of order $-2 - (l-1)$. The principal symbols of $B_{l,2,2}^{\varphi_k,-}$ in this representation are given by

$$\begin{aligned}\sigma_{-2}(B_{1,2,2}^{\varphi_k,-})(x, y, \xi) &= \frac{\pi^2}{\epsilon} \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-2}, \\ \sigma_{-3}(B_{2,2,2}^{\varphi_k,-})(x, y, \xi) &= \frac{\pi^2 \iota}{\epsilon} \varphi_\epsilon(x) \varphi_k(-\xi/|\xi|) |\xi|^{-3},\end{aligned}$$

for $|\xi|$ large.

Using Lemma 4.1 and 4.2, we can write for $k = 1, 2, 3$ that

$$\begin{aligned}(4.3) \quad I^\varphi L_\epsilon u &= \tilde{B}_1^{\varphi_k,+} f_1 + \tilde{B}_2^{\varphi_k,+} f_2 + B_{1,2,2}^{\varphi_k,+} f_1 + B_{2,2,2}^{\varphi_k,+} f_2 \\ &+ \tilde{B}_1^{\varphi_k,-} f_1 + \tilde{B}_2^{\varphi_k,-} f_2 + B_{1,2,1}^{\varphi_k,-} f_1 + B_{2,2,1}^{\varphi_k,-} f_2 + R_1^{\varphi_k} f_1 + R_2^{\varphi_k} f_2,\end{aligned}$$

where $R_1^{\varphi_k} = R_1^{\varphi_k,+} + R_1^{\varphi_k,-}$, $R_2^{\varphi_k} = R_2^{\varphi_k,+} + R_2^{\varphi_k,-}$ are regularizing operators and

$$\begin{aligned}(4.4) \quad \tilde{B}_1^{\varphi_k,+} &= B_{1,1}^{\varphi_k,+} + B_{1,2,2}^{\varphi_k,-} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ \tilde{B}_2^{\varphi_k,+} &= B_{2,1}^{\varphi_k,+} + B_{2,2,2}^{\varphi_k,-} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon}), \\ \tilde{B}_1^{\varphi_k,-} &= B_{1,1}^{\varphi_k,-} + B_{1,2,1}^{\varphi_k,+} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}), \\ \tilde{B}_2^{\varphi_k,-} &= B_{2,1}^{\varphi_k,-} + B_{2,2,1}^{\varphi_k,+} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon}).\end{aligned}$$

Also, $B_{1,2,2}^{\varphi_k,+} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$, $B_{2,2,2}^{\varphi_k,+} \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$ and $B_{1,2,1}^{\varphi_k,-} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$, $B_{2,2,1}^{\varphi_k,-} \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$. We note again that the orders of these FIOs acting on f_1, f_2 are lower than those corresponding terms in (4.4), even though the canonical relations are different. We sum the equations (4.3) in k to get

$$\begin{aligned}(4.5) \quad I^\varphi L_\epsilon u &= \tilde{B}_1^{\varphi,+} f_1 + \tilde{B}_2^{\varphi,+} f_2 + B_{1,2,2}^{\varphi,+} f_1 + B_{2,2,2}^{\varphi,+} f_2 \\ &+ \tilde{B}_1^{\varphi,-} f_1 + \tilde{B}_2^{\varphi,-} f_2 + B_{1,2,1}^{\varphi,-} f_1 + B_{2,2,1}^{\varphi,-} f_2 + R_1^\varphi f_1 + R_2^\varphi f_2,\end{aligned}$$

where we let $\varphi = \sum_{k=1}^3 \varphi_k$, and the terms are described as follows:

- (1) $I^\varphi L_\epsilon u = \sum_{k=1}^3 I^{\varphi_k} L_\epsilon u$.
- (2) $\tilde{B}_1^{\varphi,+} = \sum_{k=1}^3 \tilde{B}_1^{\varphi_k,+} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon})$, $\tilde{B}_2^{\varphi,+} = \sum_{k=1}^3 \tilde{B}_2^{\varphi_k,+} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon})$. Moreover, these operators can be expressed as in Lemma 4.1 and 4.2.

The principal symbols of the operators are given by

$$\begin{aligned}\sigma_{-1}(\tilde{B}_1^{\varphi,+})(x, y, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T+\epsilon}) \varphi_\epsilon(x) \varphi(-\xi/|\xi|) |\xi|^{-1}, \\ \sigma_{-2}(\tilde{B}_2^{\varphi,+})(x, y, \xi) &= (-2\pi^2 \ln \frac{\epsilon}{T+\epsilon}) \varphi_\epsilon(x) \varphi(-\xi/|\xi|) |\xi|^{-2},\end{aligned}$$

for $|\xi|$ large.

- (3) $\tilde{B}_1^{\varphi,-} = \sum_{k=1}^3 \tilde{B}_1^{\varphi_k,-} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon})$, $\tilde{B}_2^{\varphi,-} = \sum_{k=1}^3 \tilde{B}_2^{\varphi_k,-} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T-\epsilon})$. Moreover, these operators can be expressed as in Lemmas 4.1 and 4.2. The principal symbols of the operators are given by

$$\begin{aligned}\sigma_{-1}(\tilde{B}_1^{\varphi,-})(x, y, \xi) &= (-2\pi^2 i \ln \frac{\epsilon}{T+\epsilon}) |\xi|^{-1} \varphi(\xi/|\xi|) \varphi_\epsilon(x), \\ \sigma_{-2}(\tilde{B}_2^{\varphi,-})(x, y, \xi) &= (2\pi^2 \ln \frac{\epsilon}{T+\epsilon}) |\xi|^{-2} \varphi(\xi/|\xi|) \varphi_\epsilon(x),\end{aligned}$$

for $|\xi|$ large.

- (4) $B_{1,2,2}^{\varphi,+} = \sum_{k=1}^3 B_{1,2,2}^{\varphi_k,+} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$, $B_{2,2,2}^{\varphi,+} = \sum_{k=1}^3 B_{2,2,2}^{\varphi_k,+} \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T-\epsilon})$, $B_{1,2,1}^{\varphi,-} = \sum_{k=1}^3 B_{1,2,1}^{\varphi_k,-} \in I^{-2}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$, $B_{2,2,1}^{\varphi,-} = \sum_{k=1}^3 B_{2,2,1}^{\varphi_k,-} \in I^{-3}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-T+\epsilon})$.
- (5) $R_l^\varphi = \sum_{k=1}^3 R_l^{\varphi_k}$, $l = 1, 2$ are regularizing operators.

Now we are ready to solve for f_1, f_2 from (4.1) and (4.5). Consider (4.1). First, because $\tilde{B}_1^+ \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{T+\epsilon})$ and the symbol is non-vanishing for $(x, \xi; y, \eta) \in C_{T+\epsilon}$ with $x \in \text{supp } \varphi_\epsilon = \overline{B}_{1-\epsilon}$, we can find a parametrix $Q^+ \in I^1(\mathbb{R}^3 \times \mathbb{R}^3; C_{-T-\epsilon})$ such that $Q^+ \circ \tilde{B}_1^+ = K^+$ modulo a regularizing operator where K^+ is a pseudo-differential operator on $\mathbb{R}^3$ of order 0 with symbol $\sigma_0(K^+)(x, \xi) = \varphi_{2\epsilon}(x - (T + \epsilon)\xi/|\xi|)$. Here, we note that $\varphi_{2\epsilon}\varphi_\epsilon = \varphi_{2\epsilon}$ by the definition of these cut-off functions. Applying Q^+ to (4.1), we get

$$(4.6) \quad Q^+ I L_\epsilon u = K^+ f_1 + Q^+ \tilde{B}_1^- f_1 + Q^+ \tilde{B}_2^+ f_2 + Q^+ \tilde{B}_2^- f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2),$$

where we have denoted by $\mathcal{I}^m(f)$ a collection of FIOs acting on f that are of order at most m and with canonical relations of the type C_a with some $a \in \mathbb{R}$. The specific forms of these operators will not be important for the analysis below. Similarly, we can find a parametrix $Q^{\varphi,+} \in I^1(\mathbb{R}^3 \times \mathbb{R}^3; C_{-T-\epsilon})$ such that $Q^{\varphi,+} \circ \tilde{B}_{\varphi,1}^+ = K^+$ modulo a regularizing operator. Applying $Q^{\varphi,+}$ to (4.5), we get

$$(4.7) \quad Q^{\varphi,+} I^\varphi L_\epsilon u = K^+ f_1 + Q^{\varphi,+} \tilde{B}_1^{\varphi,-} f_1 + Q^{\varphi,+} \tilde{B}_2^{\varphi,+} f_2 + Q^{\varphi,+} \tilde{B}_2^{\varphi,-} f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2).$$

From (4.6) and (4.7), we get

$$(4.8) \quad \begin{aligned}Q^+ I L_\epsilon u - Q^{\varphi,+} I^\varphi L_\epsilon u &= (Q^+ \tilde{B}_1^- - Q^{\varphi,+} \tilde{B}_1^{\varphi,-}) f_1 \\ &+ (Q^+ \tilde{B}_2^+ - Q^{\varphi,+} \tilde{B}_2^{\varphi,+}) f_2 + (Q^+ \tilde{B}_2^- - Q^{\varphi,+} \tilde{B}_2^{\varphi,-}) f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2).\end{aligned}$$

We note that $Q^+ \tilde{B}_1^- - Q^{\varphi,+} \tilde{B}_1^{\varphi,-} \in I^0(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-2T-2\epsilon})$, and it can be written as an oscillatory integral

$$(4.9) \quad (Q^+ \tilde{B}_1^- - Q^{\varphi,+} \tilde{B}_1^{\varphi,-}) f(z) = (2\pi)^{-3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{i(z-y) \cdot \xi} e^{i(2T+2\epsilon)|\xi|} a(y, \xi) f(y) dy d\xi,$$

where the principal symbol (the leading order term of a) is

$$\sigma_0(Q^+ \tilde{B}_1^- - Q^{\varphi,+} \tilde{B}_1^{\varphi,-})(y, \xi) = C(1 - \varphi(\xi/|\xi|)/\varphi(-\xi/|\xi|)) \varphi_{2\epsilon}(y + (T + \epsilon)\xi/|\xi|)$$

where C is a non-zero constant coming from the Maslov factor. Next, $Q^+ \tilde{B}_2^+ - Q^{\varphi,+} \tilde{B}_2^{\varphi,+} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_0)$ but the principal symbol vanishes so the term actually belongs to $\mathcal{J}^{-2}(f_2)$. Finally, $Q^+ \tilde{B}_2^- - Q^{\varphi,+} \tilde{B}_2^{\varphi,-} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{-2T-2\epsilon})$, and a similar calculation as (4.9) shows that the principal symbol is

$$\sigma_{-1}(Q^+ \tilde{B}_2^- - Q^{\varphi,+} \tilde{B}_2^{\varphi,-})(y, \xi) = C\iota(1 - \varphi(\xi/|\xi|)/\varphi(-\xi/|\xi|))|\xi|^{-1}\varphi_{2\epsilon}(y + (T + \epsilon)\xi/|\xi|)$$

with the same Maslov factor. Now we let $W^+ \in I^0(\mathbb{R}^3 \times \mathbb{R}^3; C'_{2T+2\epsilon})$ be a parametrix of $Q^+ \tilde{B}_2^+ - Q^{\varphi,+} \tilde{B}_2^{\varphi,+}$ and we get from (4.8) that

$$(4.10) \quad W^+ Q^+ IL_\epsilon u - W^+ Q^{\varphi,+} I^\varphi L_\epsilon u = K_1^+ f_1 + U^+ f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2),$$

where U^+ is a pseudo-differential operator of order -1 with principal symbol $\sigma_{-1}(U^+)(x, \xi) = \iota|\xi|^{-1}\varphi_{4\epsilon}(x + (T + \epsilon)\xi/|\xi|)$ and K_1^+ is a pseudo-differential operator of order 0 with principal symbol $\sigma_0(K_1^+)(x, \xi) = \varphi_{4\epsilon}(x + (T + \epsilon)\xi/|\xi|)$.

Next we repeat the analysis for the $-$ sign. In (4.1), we can find a parametrix $Q^- \in I^2(\mathbb{R}^3 \times \mathbb{R}^3; C_{T+\epsilon})$ such that $Q^- \circ \tilde{B}_2^- = K^-$ modulo a regularizing operator where K^- is a pseudo-differential operator on $\mathbb{R}^3$ of order 0 with symbol $\sigma_0(K^-)(x, \xi) = \varphi_{2\epsilon}(x + (T + \epsilon)\xi/|\xi|)$. Applying Q^- to (4.1), we get

$$(4.11) \quad Q^- IL_\epsilon u = K^- f_1 + Q^- \tilde{B}_1^+ f_1 + Q^- \tilde{B}_2^+ f_2 + Q^- \tilde{B}_2^- f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2).$$

Next, we find a parametrix $Q^{\varphi,-} \in I^1(\mathbb{R}^3 \times \mathbb{R}^3; C_{T+\epsilon})$ such that $Q^{\varphi,-} \circ \tilde{B}_{\varphi,1}^- = K^-$ modulo a regularizing operator. Applying $Q^{\varphi,-}$ to (4.5), we get

$$(4.12) \quad Q^{\varphi,-} I^\varphi L_\epsilon u = K^- f_1 + Q^{\varphi,-} \tilde{B}_1^{\varphi,+} f_1 + Q^{\varphi,-} \tilde{B}_2^{\varphi,+} f_2 + Q^{\varphi,-} \tilde{B}_2^{\varphi,-} f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2).$$

From (4.11) and (4.12), we get

$$(4.13) \quad \begin{aligned} Q^- IL_\epsilon u - Q^{\varphi,-} I^\varphi L_\epsilon u &= (Q^- \tilde{B}_1^+ - Q^{\varphi,-} \tilde{B}_1^{\varphi,+}) f_1 \\ &+ (Q^- \tilde{B}_2^+ - Q^{\varphi,-} \tilde{B}_2^{\varphi,+}) f_2 + (Q^- \tilde{B}_2^- - Q^{\varphi,-} \tilde{B}_2^{\varphi,-}) f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2). \end{aligned}$$

We note that $Q^- \tilde{B}_1^+ - Q^{\varphi,-} \tilde{B}_1^{\varphi,+} \in I^0(\mathbb{R}^3 \times \mathbb{R}^3; C'_{2T+2\epsilon})$. Writing the operator as in (4.9), we find that the principal symbol is

$$\sigma_0(Q^- \tilde{B}_1^+ - Q^{\varphi,-} \tilde{B}_1^{\varphi,+})(y, \xi) = C(1 - \varphi(\xi/|\xi|)/\varphi(-\xi/|\xi|))\varphi_{2\epsilon}(y - (T + \epsilon)\xi/|\xi|)$$

where C is a non-zero constant. Also, $Q^- \tilde{B}_2^- - Q^{\varphi,-} \tilde{B}_2^{\varphi,-} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_0)$ but the principal symbol is 0 so the term belongs to $\mathcal{J}^{-2}(f_2)$. Finally, $Q^- \tilde{B}_2^+ - Q^{\varphi,-} \tilde{B}_2^{\varphi,+} \in I^{-1}(\mathbb{R}^3 \times \mathbb{R}^3; C'_{2T+2\epsilon})$ and the principal symbol is

$$\sigma_{-1}(Q^- \tilde{B}_2^+ - Q^{\varphi,-} \tilde{B}_2^{\varphi,+})(y, \xi) = -C\iota(1 - \varphi(\xi/|\xi|)/\varphi(-\xi/|\xi|))|\xi|^{-1}\varphi_{2\epsilon}(y - (T + \epsilon)\xi/|\xi|).$$

Now we let $W^- \in I^0(\mathbb{R}^3 \times \mathbb{R}^3; C'_{2T+2\epsilon})$ be a parametrix of $Q^- \tilde{B}_2^+ - Q^{\varphi,-} \tilde{B}_2^{\varphi,+}$ and we get from (4.13) that

$$(4.14) \quad W^- Q^- IL_\epsilon u - W^- Q^{\varphi,-} I^\varphi L_\epsilon u = K_1^- f_1 + U^- f_2 + \mathcal{J}^{-1}(f_1) + \mathcal{J}^{-2}(f_2),$$

where U^- is a pseudo-differential operator of order -1 with principal symbol $\sigma_{-1}(U^-)(x, \xi) = -\iota|\xi|^{-1}\varphi_{4\epsilon}(x - (T + \epsilon)\xi/|\xi|)$ and K_1^- is a pseudo-differential operator of order 0 with principal symbol $\sigma_0(K_1^-)(x, \xi) = \varphi_{4\epsilon}(x - (T + \epsilon)\xi/|\xi|)$.

Finally, we can summarize the analysis for determining singularities in f_1, f_2 . Below, WF_s denotes the H^s wavefront set for $s \in \mathbb{R}$.

Theorem 4.3. *Let $H^+ = W^+Q^+I - W^+Q^{\varphi,+}I^{\varphi}$ and $H^- = W^-Q^-I - W^-Q^{\varphi,-}I^{\varphi}$. Suppose $f_1 \in H^{s'}(\mathbb{R}^3)$, $f_2 \in H^{s'-1}(\mathbb{R}^3)$ for $s' > s - 1$, $s \geq 0$. Then*

(1) *for any $(x, \xi) \in T^*\mathbb{R}^3 \setminus 0$ such that $x + T\xi/|\xi| \in \mathcal{B}_{1-4\epsilon}$, we have*

$$(x, \xi) \in WF_s(K_1^+ f_1 + U^+ f_2) \text{ if and only if } (x, \xi) \in WF_s(H^+ L_{\epsilon} u).$$

(2) *for any $(x, \xi) \in T^*\mathbb{R}^3 \setminus 0$ such that $x - T\xi/|\xi| \in \mathcal{B}_{1-4\epsilon}$, we have*

$$(x, \xi) \in WF_s(K_1^- f_1 + U^- f_2) \text{ if and only if } (x, \xi) \in WF_s(H^- L_{\epsilon} u).$$

Proof. The proof follows from a wavefront analysis from (4.10) and (4.14) on the elliptic region of $K_1^{\pm}$. Note that $\mathcal{I}^{-1}(f_1)$ and $\mathcal{I}^{-2}(f_2)$ contains FIOs of graph type with higher orders. Hence, they do not contribute to H^s wavefront sets. $\square$

We remark that $K_1^{\pm} f_1 + U^{\pm} f_2$ are essentially the re-parametrized Cauchy data for the wave equation (2.4). It might be possible to remove the assumption $f_1 \in H^{s'}(\mathbb{R}^3)$, $f_2 \in H^{s'-1}(\mathbb{R}^3)$ if one examines the detailed structures of the operators in $\mathcal{I}^{-1}(f_1)$ and $\mathcal{I}^{-2}(f_2)$. We do not pursue it here. However, we point out that for some cases, it is easy to see that the assumption can be removed, for example $f_1 = 0$ or $f_2 = 0$.

5. PROOF OF THEOREM 2.1

We start with the Sobolev estimates for recovering f_1 assuming $f_2 = 0$. We recall the mapping properties of FIOs of graph type on Sobolev spaces, see [14, Section 25.3]. In particular, for $A \in I^{\mu}(\mathbb{R}^3 \times \mathbb{R}^3; C'_a)$, where $a \in \mathbb{R}$, $\mu \in \mathbb{R}$, we have for $s \in \mathbb{R}$ that

$$\|Af\|_{H^s(\mathbb{R}^3)} \leq C\|f\|_{H^{s+\mu}(\mathbb{R}^3)}$$

for some $C > 0$. Hereafter, we use C for a generic constant which could depend on ϵ but is uniform in f_1, f_2 . Assume that f_1 is compactly supported in $\mathcal{R}_{T-1, T+1}$ and $WF(f_1) \subset \mathcal{W}$. Then there is $\epsilon > 0$ small such that (4.10) and (4.14) hold. We use $U^{\pm} \in \Psi^{-1}(\mathbb{R}^3)$ to get that

$$\|f_1\|_{H^s} \leq C(\|IL_{\epsilon}u\|_{H^{s+1}} + \|I^{\varphi}L_{\epsilon}u\|_{H^{s+1}}) + C\|f_1\|_{H^{s-1}}.$$

We observe that for $s + 2 \geq 0$ an integer,

$$\|IL_{\epsilon}u\|_{H^{s+2}} \leq C\|Lu\|_{H^{s+2}}, \quad \|I^{\varphi}L_{\epsilon}u\|_{H^{s+2}} \leq C\|Lu\|_{H^{s+2}},$$

which can be seen by directly estimating the integral in I and I^{φ} . Thus

$$(5.1) \quad \|f_1\|_{H^{s+1}} \leq C\|Lu\|_{H^{s+2}} + C\|f_1\|_{H^s}.$$

Next, we remove the last term. We will need the following uniqueness result for the light ray transform with partial data.

Lemma 5.1. *Let $u \in H^s(\mathcal{M})$, $s \geq 0$ and $Lu = 0$. Then $u = 0$ on the set*

$$(5.2) \quad \{(t, x) \in \mathcal{M} : \min(0, T - 1 - t) < |x| < T + 1 - t\}.$$

Proof. Let $\varphi(x)$ be the characteristic function of $\mathcal{B}$ in $\mathbb{R}^3$. Let $\chi(t)$ be the characteristic function of $[0, T]$ in $\mathbb{R}$. We write

$$Lu(y, v) = \int_{\mathbb{R}} \chi(t) \varphi(y) u(t, y + tv - Tv) dt,$$

where $y \in \mathbb{R}^3, v \in \mathbb{S}^2$. We follow the proof of the Fourier Slice Theorem to compute that

$$(5.3) \quad \begin{aligned} 0 &= \int_{\mathbb{R}^3} e^{-iy \cdot \xi} Lu(y, v) dy = \int_{\mathbb{R}^3} e^{-iy \cdot \xi} \chi(t) \varphi(y) u(t, y + tv - Tv) dt dy \\ &= \int_{\mathbb{R}^3} e^{-i(x-tv+Tv) \cdot \xi} \chi(t) \varphi(x - tv + Tv) u(t, x) dt dx = e^{-iTv \cdot \xi} \hat{A}(-v \cdot \xi, \xi), \end{aligned}$$

where $\hat{A}(\tau, \xi)$ is the Fourier transform of $A(t, x) = \chi(t) \varphi(x - tv + Tv) u(t, x)$. Note that A is compactly supported in $\mathbb{R}^{3+1}$ for each fixed v so $\hat{A}$ is analytic for fixed v . We see from (5.3) that $\hat{A}$ vanishes on some open set with non-empty interior. Thus $\hat{A}$ is identically zero. So we get $\chi(t) \varphi(x - tv + Tv) u(t, x) = 0$ for any $v \in \mathbb{S}^2$ so u vanishes on the set (5.2). $\square$

Proof of Theorem 2.1. The unique determination is contained in Lemma 5.1. Next, we consider the stability. Let $f_1 \in H^{s+1}(\mathbb{R}^3)$, $s \geq 0$ be supported in $\mathcal{R}_0$ which is compact in $\mathcal{R}_{T-1, T+1}$ and $\text{WF}(f_1) \subset \mathcal{W}$. Then there exists $\epsilon > 0$ such that (5.1) holds with C depending on ϵ . We prove (2.7), that is,

$$\|f_1\|_{H^{s+1}} \leq C \|Lu\|_{H^{s+2}}.$$

By contradiction, we assume that for $n = 1, 2, \dots$, there are $f_1^n \in H^{s+1}(\mathbb{R}^3)$ supported in $\mathcal{R}_0$ and $\text{WF}(f_1^n) \subset \mathcal{W}$ such that $\|f_1^n\|_{H^{s+1}} = 1$, and

$$(5.4) \quad \|f_1^n\|_{H^{s+1}} \geq n \|Lu^n\|_{H^{s+2}},$$

where u^n is the solution of (2.4) with Cauchy data f_1^n and $f_2^n = 0$. First, we note that on the fixed compact set $\mathcal{R}_0$, $H^{s+1}(\mathcal{R}_0)$ is compactly embedded in $H^s(\mathcal{R}_0)$. Thus by passing to a subsequence, we can assume that f_1^n converges in H^s , and it follows from (5.1) that f_1^n converges to some f_1 in H^{s+1} . For the Cauchy problem (2.4), we have the standard energy estimate

$$\|u\|_{H^{s+1}} \leq C \|f_1\|_{H^{s+1}}.$$

Therefore, u^n converges to u in H^{s+1} . We get from (5.4) that $0 = \|Lu\|_{H^{s+2}} \geq \|Lu\|_{H^{s+1}}$. By Lemma 5.1, we deduce that $f_1 = 0$. But this contradicts to $\|f_1\|_{H^{s+1}} = 1$. Thus, we proved (2.7), and this completes the proof of Theorem 2.1. $\square$

6. THE NUMERICAL EXPERIMENT

6.1. The numerical setup. We conduct numerical experiments for a 2D universe. In Section 2, we explained the local nature of the problem. So it suffices to consider the numerical simulation on a compact region. We take $\mathcal{M} = [0, 2] \times [-7, 7]^2$ and consider the Cauchy problem of the standard wave equations

$$(6.1) \quad \begin{aligned} \square u(t, x) &= 0, \quad t > 0, x \in [-7, 7]^2 \\ u(0, x) &= f(x), \quad \partial_t u(0, x) = 0, \end{aligned}$$

where $\square = \partial_t^2 + \Delta$ with $\Delta = -\sum_{i=1}^2 \partial_{x_i}^2$. Instead of using techniques such as Perfectly Matched Layer, we will impose zero boundary conditions on the lateral boundary $[0, 2] \times \partial[-7, 7]^2$. This means that waves are reflected from the boundary. In the numerical experiments, we will make sure that either these reflected waves are not observed (in the partial data reconstruction) or we restrict f to be supported in a sufficiently small region (for the full data reconstruction) so that the waves do not meet the boundary for $t \in [0, 2]$. Hence, they do not affect the reconstruction problem.

Let $\mathcal{U}$ be an open set of $\{2\} \times [-7, 7]^2$. We consider light-like geodesics from $\{0\} \times [-7, 7]^2$ that meet $\mathcal{U}$ and the light ray transform

$$(6.2) \quad Lu(y, v) = \int_0^T u(t, y + tv - Tv) dt, \quad y \in \mathcal{U}, v \in \mathbb{S}^1.$$

This will be the simulated CMB data. The inverse problem is to recover f from Lu . In principle, we treat this as a PDE constrained optimization problem. However, for (6.1), it is relatively straightforward to find the (discretized) solution operator. We let $u = Sf$ and $w = Lu$ be the measurement. Then we can solve the more direct optimization problem,

$$\min_f \|LSf - w\| + \alpha \|f\|$$

with suitable norms and regularization terms. Here, $\alpha \geq 0$ is a regularization parameter.

We start with the discretization of S . We will solve (6.1) by the classical finite difference method on a rectangular grid. We refer to [18] for the details. The mesh in time consists of T time points $t_0 = 0 < t_1 < \dots < t_T$, with constant spacing Δt , and mesh points in each space direction are given as

$$x_1 < x_2 < \dots < x_n \quad \text{and} \quad y_1 < y_2 < \dots < y_n$$

with equally spaced spatial grids Δx and Δy . For any time point t_τ , we denote $u_{i,j}^\tau$ as u at mesh point (x_i, y_j, t_τ) and let $\mathbf{u}^\tau \in \mathbb{R}^{n^2}$ be the discretized and vectorized image containing values $u_{i,j}^\tau$. That is, at $\tau = 0$ (i.e., the initial time point t_0), the unknown initial condition is given by $\mathbf{u}^0 = \mathbf{f}$. For large 2D meshes, vectorization is essential for efficient computation; thus, we describe the solution operator in terms of matrix multiplications. Let $\mathbf{T}_x$ and $\mathbf{T}_y$ represent Poisson matrices obtained using finite differences, i.e.,

$$\mathbf{T}_x = \mathbf{T}_y = \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Then, given the discretized initial condition $\mathbf{f}$, the solution of the PDE at the first time point can be computed as

$$\mathbf{u}^1 = \mathbf{T}\mathbf{u}^0 \quad \text{where} \quad \mathbf{T} = \mathbf{I} + \frac{\sigma^2}{2}(\mathbf{I} \otimes \mathbf{T}_x) + \frac{\gamma^2}{2}(\mathbf{T}_y \otimes \mathbf{I})$$

with $\sigma = \gamma = \Delta t / \Delta x$. Subsequent time points can be computed as,

$$(6.3) \quad \mathbf{u}^\tau = 2\mathbf{T}\mathbf{u}^{\tau-1} - \mathbf{u}^{\tau-2} \quad \text{for} \quad \tau = 2, \dots, T.$$

One can represent the entire vector of spatio-temporal voxels $\mathbf{u} \in \mathbb{R}^{n^2 T}$ as

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \vdots \\ \mathbf{u}^T \end{bmatrix} = \mathbf{S}\mathbf{u}^0 = \mathbf{S}\mathbf{f},$$

for some matrix $\mathbf{S} \in \mathbb{R}^{n^2 T \times n^2}$.

For the discretization of the light ray transform L , we use the method described in [5]. Roughly speaking, we think of each grid point on $\{0\} \times [-7, 7]^2$ as a point source, and each grid point on $\{2\} \times [-7, 7]^2$ as a receiver. Then we find the light rays from each source point

to the (nearest) receivers which yields the discretization of the light rays. Assume there are N source grid locations. Then for each source location s_i for $i = 1, 2, \dots, N$, there are m_i observations that are collected in vector,

$$(6.4) \quad \mathbf{b}_i = \mathbf{H}_i \mathbf{u} + \mathbf{e}_i$$

where $\mathbf{b}_i \in \mathbb{R}^{m_i}$ and $\mathbf{H}_i \in \mathbb{R}^{m_i \times n^2 T}$ with noise contamination $\mathbf{e}_i \in \mathbb{R}^{m_i}$. The number of observations m_i corresponds to the number of observable detector grid locations such that light-rays from source point s_i have a non-zero interaction with the space-time mesh grid. Thus, if we let

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_N \end{bmatrix} \quad \text{and} \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_N \end{bmatrix},$$

we have an inverse problem where noisy data is given by

$$(6.5) \quad \mathbf{b} = \mathbf{H}\mathbf{u} + \mathbf{e} = \underbrace{\mathbf{H}\mathbf{S}}_{\mathbf{A}} \mathbf{f} + \mathbf{e},$$

where $\mathbf{A} \in \mathbb{R}^{m \times n^2}$ with $m = \sum_{i=1}^N m_i$, and the goal of the reconstruction problem is to obtain an approximation of the initial time point $\mathbf{f}$, given $\mathbf{b}$ and $\mathbf{A}$. We assume that $\mathbf{e}$ is i.i.d. standard Gaussian noise.

6.2. The spectral analysis. Before we consider the reconstruction problem, we investigate some of the spectral properties of $\mathbf{A}$. For the discretization of the world sheet, the x, y -values are linearly spaced on $[-7, 7] \times [-7, 7]$ and discretized using 51×51 points. Thus, $\mathbf{f} \in \mathbb{R}^{2601}$. We assume that we have $T = 40$ time slices where the planes are fixed to be the midpoints of the subintervals of $[0, 2]$ so that the first and last do not coincide with the source and detector planes. The PDE solution matrix $\mathbf{S}$ that takes the initial condition $\mathbf{f}$ to the full space-time representation is $104,040 \times 2,601$. The cone beam light ray transform matrix $\mathbf{H}$ is defined by the locations of the sources and detectors. The sources and detectors are equally spaced on a 51×51 grid on $[-7, 7] \times [-7, 7]$. For the full data reconstruction problem (where all detectors receive measurements), the resulting matrix $\mathbf{H}$ is $103,384 \times 104,040$.

We also consider partial data reconstruction problems where the locations of the detectors remain fixed, but the active detectors that collect measurements are only those located in a square region around the origin. For different sizes of detector grids, we provide in Table 1 the corresponding total number of observed measurements. As expected, as the number of active detectors decreases, the total number of observations also decreases and the condition number of the forward model matrix $\mathbf{A}$ increases rapidly. The number of unknowns $n^2 = 2,601$ remains fixed.

detector grid	m	$\kappa(\mathbf{A})$
51×51	103,384	60.1800
21×21	21,168	558.0165
7×7	2,352	∞
3×3	432	∞

TABLE 1. For different sized detector grids, we provide the total number of observations m , which corresponds to the number of rows of $\mathbf{A}$. The number of columns of $\mathbf{A}$ remains fixed at 2,601. κ denotes the numerically computed condition number.

For each of the matrices $\mathbf{A}$ corresponding to different numbers of detectors, we provide the decay of the singular values in Figure 4. We observe that with more detectors (e.g., a larger detector grid), there are very few small singular values and so the matrix is well-conditioned. As the number of observations decreases, the problem becomes more ill-posed, evident in the small (and numerically zero) singular values.

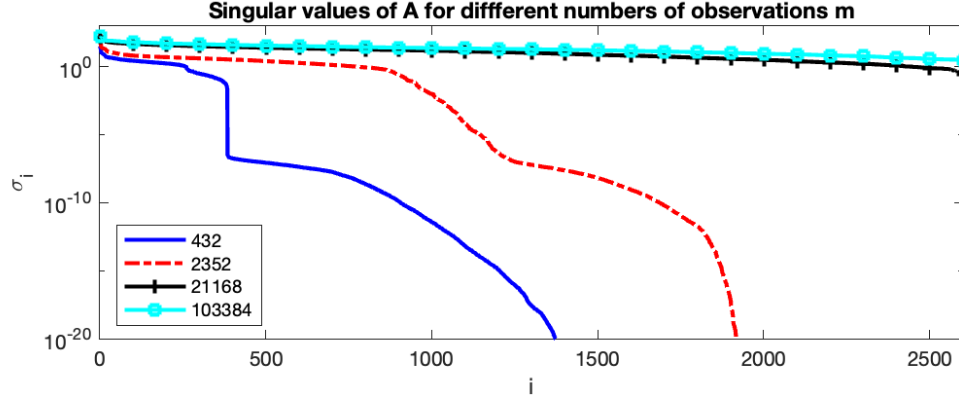


FIGURE 4. Singular values σ_i for matrices $\mathbf{A}$ corresponding to different sized detector grids. In the legend, m corresponds to the number of observed measurements.

For the 7×7 grid, we provide in Figure 5 the Picard plot [13], which contains the singular values of $\mathbf{A}$, σ_i , the SVD coefficients $|\mathbf{u}_i^\top \mathbf{b}|$ and the solution coefficients $\frac{|\mathbf{u}_i^\top \mathbf{b}|}{\sigma_i}$ where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the left and right singular vectors of $\mathbf{A}$ respectively. The observation vector $\mathbf{b}$ was constructed using the true image in the bottom row of Figure 6 with 10^{-4} noise level. In addition to the decay of singular values, we observe that the SVD coefficients decay on average faster than the singular values and they level off at the noise level. We also observe that once the SVD coefficients level off at the noise level, the solution coefficients increase significantly. Thus, the inverse solution, $\sum_i \frac{\mathbf{u}_i^\top \mathbf{b}}{\sigma_i} \mathbf{v}_i$, will likely be dominated by noise. Regularization will be essential to obtain reasonable reconstructions, especially for the partial data problem. Next we provide numerical results and comparisons for the full data and partial data reconstruction problems.

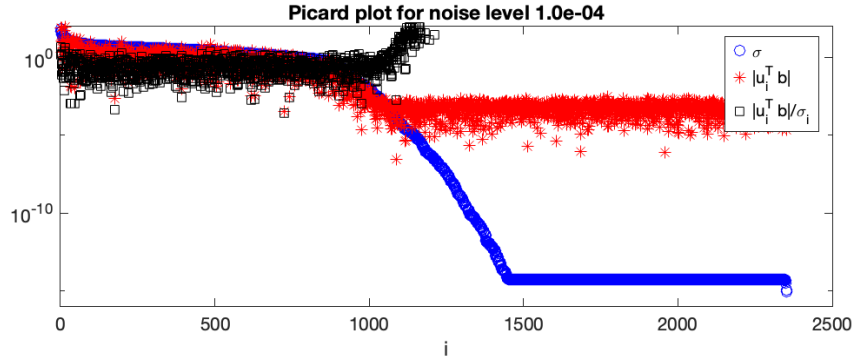


FIGURE 5. Picard plot for the 7×7 detector grid demonstrating the impact of noise on the SVD and solution coefficients.

6.3. Full data reconstruction problem. For the full data problem set up, we have 51^2 detectors and the corresponding matrix $\mathbf{A}$ has a condition number of 60.18. This is relatively small, so we anticipate that even with noise in the data, solving the least-squares (LS) problem should yield good solutions. So for the discrete inverse problem (6.5), we consider solving

$$(6.6) \quad \min_{\mathbf{f}} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|^2,$$

where unless specified otherwise, $\|\cdot\|$ will denote the Euclidean norm.

We use two different true images, denoted $\mathbf{f}_{\text{true}}$, to represent the true initial condition, see the first column of Figure 6. Each image is of size 51×51 . The first one contains random dots that simulate gravitational wave point sources. The second image contains randomly placed vertical and horizontal lines that simulate sources of gravitational plane waves or cosmic strings. Note that these images have relatively simple singularities. For the full-data reconstruction, we need to avoid the reflected waves from the lateral boundary and ensure that we collect all rays from the cone beam light transform that intersect the object. By the finite speed of propagation, it suffices to restrict the non-zero pixels of the true image to $[-3, 3] \times [-3, 3]$. Thus, for the full data problem, we use the true image in the left column of Figure 6, where every pixel outside of the yellow box is set to zero. Then, for each image, we simulated observations as in (6.5) where we added Gaussian white noise that was scaled so that the noise level was $\frac{\|\mathbf{e}\|}{\|\mathbf{A}\mathbf{f}_{\text{true}}\|} = 0.02$.

For this example, we can compute the LS reconstruction using standard techniques (e.g., via the QR factorization [3]), and the reconstruction $\hat{\mathbf{f}}$ is provided in the right column of Figure 6, with the relative reconstruction error, computed as $\frac{\|\hat{\mathbf{f}} - \mathbf{f}_{\text{true}}\|}{\|\mathbf{f}_{\text{true}}\|}$, provided in the title. We also compute reconstructions obtained using an iterative LS solver, namely LSQR [24]. We observe that after 100 iterations of LSQR, we get comparably good reconstructions. However, we do note that for larger noise levels, the LS solution may become contaminated with noise even in the full data case, so regularization may still be required. For iterative methods such as LSQR, early termination of the iterations is a common approach to avoid potential noise amplification in the solution [13]. These and other regularization techniques are considered next for the partial data problem.

6.4. Partial data reconstruction problem. Next we consider the partial data reconstruction problem, for which we demonstrate two main points. First, for the partial data problem, regularization is critical for obtaining a good reconstruction. Second, different regularization choices can be used to incorporate prior knowledge about the solution. We remark that contrary to our previous CMB reconstruction study [5], the unknowns here represent a 2D image (i.e., the initial condition), rather than a 3D spatial-temporal image. That is, by incorporating the PDE solution operator, we obtain a CMB reconstruction problem that is better conditioned and requires estimating significantly fewer unknown parameters. Moreover, only spatial regularization is needed here, rather than expensive spatiotemporal priors, which presents a wide array of options for priors.

We consider various forms of regularization. For large-scale problems, iterative methods can be used to compute approximate solutions to (6.6), where at the k th iteration, a solution is sought in a k -dimensional Krylov subspace, $\mathcal{K}_k$. That is, the k^{th} iterate is given by

$$(6.7) \quad \mathbf{f}^{(k)} = \arg \min_{\mathbf{f} \in \mathcal{R}(\mathbf{V}_k)} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|^2$$

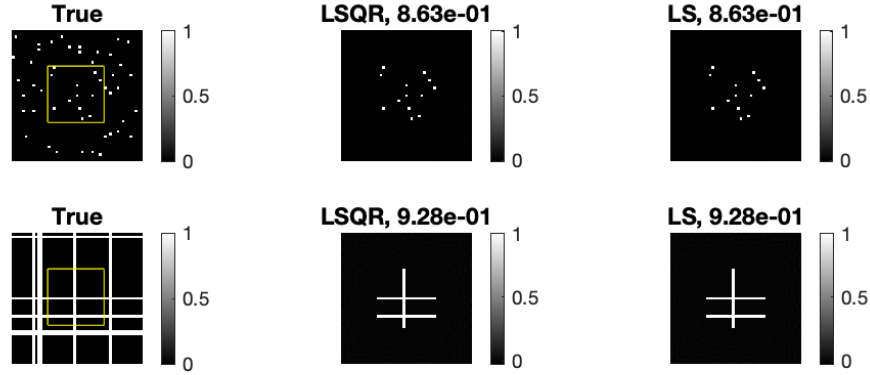


FIGURE 6. We provide two example true images, where the box identifies the region $[-3, 3] \times [-3, 3]$ used for the full data problem. We provide the reconstructions for the full data problem using iterative method LSQR (after 100 iterations) and the LS solution. Relative reconstruction error norms are provided in the titles.

where $\mathcal{R}(\cdot)$ denotes the range of a given matrix and $\mathcal{R}(\mathbf{V}_k) = \mathcal{K}_k(\mathbf{A}^\top \mathbf{A}, \mathbf{A}^\top \mathbf{b})$ when the starting guess, $\mathbf{x}^{(0)}$, is the zero vector. It is well known that for ill-posed inverse problems, iterative methods such as LSQR exhibit semiconvergence behavior whereby early iterations reconstruction good solution approximations, but latter iterations are dominated by inverted noise [12]. An early termination of the iterative method can produce a regularized solution, where the stopping iteration serves as the regularization parameter.

We also consider variational regularization approaches where the goal is to solve a penalized problem of the form,

$$(6.8) \quad \min_{\mathbf{f}} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|^2 + \lambda R(\mathbf{f})$$

where R is a regularization term and $\lambda > 0$ is a regularization parameter that balances the trade-off between the data fit and the regularization term. For example, the ℓ_1 regularization problem,

$$(6.9) \quad \min_{\mathbf{f}} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|^2 + \lambda \|\mathbf{f}\|_1,$$

is known to promote sparsity in solutions by reducing the sensitivity to outliers. Various algorithms can be used to solve (6.9), and we use the Fast-ISTA (FISTA) method [2] with backtracking as implemented in the software package IR-Tools [9].

As an alternative strategy, we consider edge preserving regularization, where an edge preserving prior is used to impose sparsity on derivative images [4, 10]. Following [1], we use intrinsic Gaussian Markov random field (IGMRF) priors where the increment variance is allowed to be larger in some locations. More precisely, recall that the unknown initial state is defined on a two dimensional mesh, so let f_{ij} denote the pixel value at mesh point (x_i, y_j) for $i, j = 1, \dots, n$. Then, we assume the independent increment model,

$$\begin{aligned} \Delta_s f_{ij} &\sim \mathcal{N}(0, (w_{ij}^s \delta)^{-1}), & i = 1, \dots, n-1, & j = 1, \dots, n \\ \Delta_t f_{ij} &\sim \mathcal{N}(0, (w_{ij}^t \delta)^{-1}), & i = 1, \dots, n, & j = 1, \dots, n-1 \end{aligned}$$

where the horizontal and vertical increments are given by $\Delta_s f_{ij} = f_{i+1,j} - f_{ij}$ and $\Delta_t f_{ij} = f_{i,j+1} - f_{ij}$ respectively and $\mathcal{N}$ denotes a normal distribution. If we take $w_{ij} = w_{ij}^s = w_{ij}^t$ and let $\mathbf{A} = \text{diag}(\text{vec}(\{w_{ij}\}))$, we get a prior of the form,

$$(6.10) \quad p(\mathbf{f} \mid \delta) \propto \delta^{(n^2-2)/2} \exp \left(-\frac{\delta}{2} \mathbf{f}^\top (\mathbf{D}_s^\top \mathbf{A} \mathbf{D}_s + \mathbf{D}_t^\top \mathbf{A} \mathbf{D}_t) \mathbf{f} \right)$$

where $\mathbf{D}_s = \mathbf{I} \otimes \mathbf{D}$ and $\mathbf{D}_t = \mathbf{D} \otimes \mathbf{I}$ with

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \cdots & -1 & 1 \\ 1 & 0 & \cdots & 0 & -1 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

and $\delta > 0$. If the goal is to preserve edges, we can assume that we have an estimate of $\mathbf{f}$ and take $\mathbf{A}(\mathbf{f}) = \text{diag}(\mathbf{1} / \sqrt{(\mathbf{D}_s \mathbf{f})^2 + (\mathbf{D}_t \mathbf{f})^2 + \beta \mathbf{1}})$ in (6.10) where vector operations are computed element wise and $0 < \beta \ll 1$. We use an iterative process to estimate $\mathbf{f}$ and use the estimates to update the prior. See Algorithm 1 for a description of the edge preserving algorithm. Note that at each iteration, we must solve a Tikhonov regularization problem, and for this, we use the conjugate gradient (CG) method with a fixed regularization parameter λ_k .

Algorithm 1 Edge preserving algorithm using IGMRF

Require: $\mathbf{A}$, $\mathbf{b}$, K (maximum outer iterations)

Set $\mathbf{L}_1 = \mathbf{D}_s^\top \mathbf{D}_s + \mathbf{D}_t^\top \mathbf{D}_t$

Compute $\mathbf{f}^{(1)} = (\mathbf{A}^\top \mathbf{A} + \lambda_1 \mathbf{L}_1)^{-1} \mathbf{A}^\top \mathbf{b}$ using CG

for $k = 2, \dots, K$ **do**

Set $\mathbf{L}_k = \mathbf{D}_s^\top \mathbf{A}(\mathbf{f}^{(k-1)}) \mathbf{D}_s + \mathbf{D}_t^\top \mathbf{A}(\mathbf{f}^{(k-1)}) \mathbf{D}_t$

Compute $\mathbf{f}^{(k)} = (\mathbf{A}^\top \mathbf{A} + \lambda_k \mathbf{L}_k)^{-1} \mathbf{A}^\top \mathbf{b}$ using CG

end for

Next we present various numerical results comparing different reconstruction algorithms. For the numerical experiments presented here, observations have added noise such that the noise level is 0.02. We provide the LS solution, the LSQR reconstruction that corresponds to the smallest reconstruction error norm, the FISTA reconstruction after 500 iterations with $\lambda = 6.6 \times 10^{-5}$, and the edge preserving reconstruction with $K = 5, 100$ maximum CG iterations, $\lambda_k = e^k$ where k is the outer iteration counter, and $\beta = 0.001$. Finally, we remark that in the following experiments, the reflected waves from the lateral boundaries are not observed. Thus they do not affect the reconstructions.

First, for the tomography reconstruction problem with a 7×7 detector grid, we provide reconstructions for the random dots example and for the random lines example in Figure 7. In Figure 8, we provide the relative residual norm per iteration and the relative reconstruction error norms per iteration. We observe that without regularization, the least-squares reconstruction is very poor because the reconstruction is dominated by error and noise. The LSQR reconstruction is significantly better and corresponds to the LSQR iterate with the smallest relative error norm. This is marked by a blue star in the right figures of Figure 8. Although the error curves are not available in practice, the results show that if a good stopping iteration or stopping criteria is known, then the LSQR reconstruction is good. However, without an

appropriate stopping criteria, the reconstruction can become dominated with noise. From the plot of the relative errors for LSQR, we observe semiconvergence whereby the errors decrease in early iterations but increase at later iterations.

Both FISTA and the edge-preserving algorithm provide good reconstructions. Notice that since IGMRF has an inner and outer iteration, the x-axis in Figure 8 for IGMRF corresponds to the total number of CG iterations. Moreover, the stars correspond to the reconstructions in Figure 7. We note that due to the limited view from the detectors, only unknowns in the central region of the image are able to be reconstructed accurately. Regions along the borders are not visible and hence do not get reconstructed.

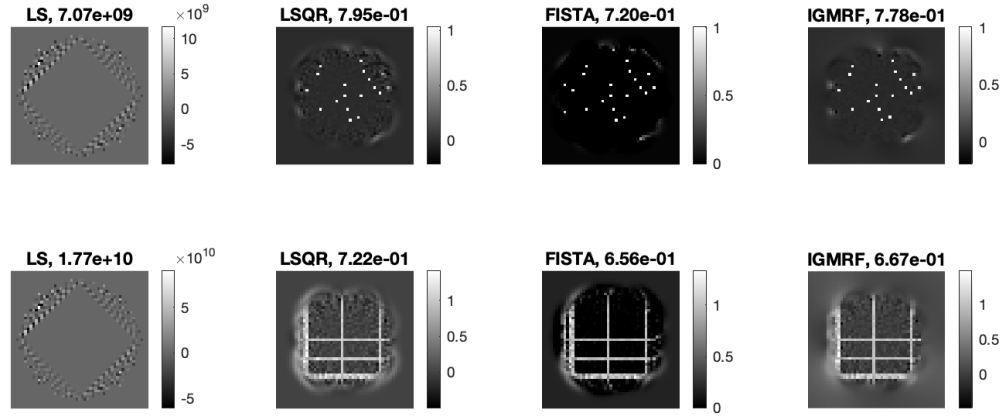


FIGURE 7. Partial data reconstructions for the 7×7 detector grid for the random dots example (top) and the random lines example (bottom). Relative reconstruction error norms are provided in the titles.

For a smaller 3×3 detector grid, we compare reconstructions for LSQR, FISTA, and IGMRF in Figure 9. We do not provide results for the LS solution because due to the ill-posed nature of the problem, the LS solution is useless (i.e., dominated by noise as illustrated in the leftmost column of Figure 7). In both examples, we observe that due to the smaller grid of detectors, the window or region of reconstruction is smaller. In particular, there is a very small circular region in the center of the image that cannot be reconstructed as well as the ring region. This agrees with our theoretical analysis. However, the reconstruction in the ring region is very well due to the fact that the singularities of the image are relatively simple and we included regularization. Also, we observe that the size of the ring region seems to be larger than that predicted in Theorem 2.1. This is not an issue because our initial data is not supported in the ring region as required by Theorem 2.1. Nevertheless, the numerical results seem to indicate that some partial information can be reconstructed in a larger region that is not covered by our main results.

7. CONCLUSIONS

In this work, we have shown that it is possible to stably recover some information of the initial status of the observable universe using only the CMB observed near the Earth. Mathematically, we show that the inversion of the cosmological X-ray transform, which is a severely ill-posed inverse problem, can be stabilized by incorporating the physical model. In

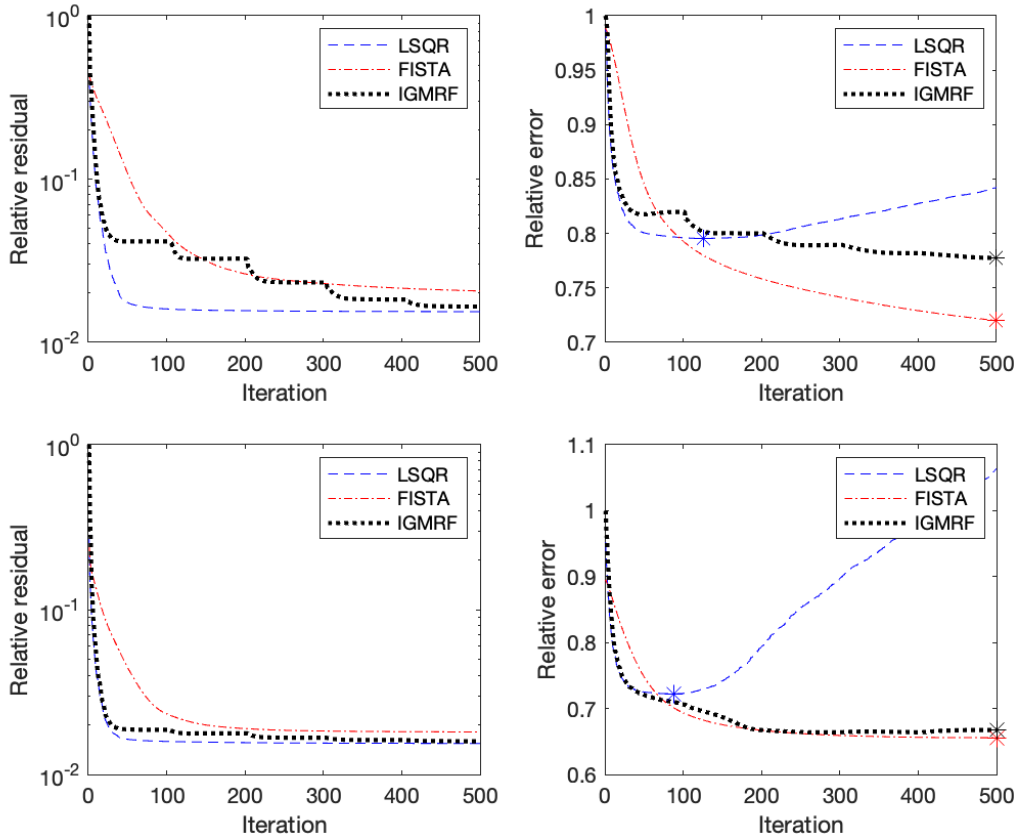


FIGURE 8. Relative residual norms (left) and relative reconstruction error norms (right) per iteration for the partial data reconstruction problems. These results correspond to the 7×7 detector grid for the random dots example (top) and random lines example (bottom).

addition to the analysis, we provide a numerical study for inverting the ray transform. We investigate the spectral properties of the discretized problem which reveal the importance of incorporating regularization for the partial data problem. Also, we compare various reconstruction methods to demonstrate the feasibility of the tomography method for cosmological applications.

Our study opens up many possibilities. One outstanding problem in cosmology is the detection of primordial gravitational waves generated in the beginning of Universe, see for example [16]. Theoretical studies have shown that these waves will leave signatures (e.g., polarizations) in the CMB, although such signatures have not been identified yet. The tomography approach we studied in this paper provides another way to find the primordial waves indirectly. For this purpose, one needs to study the inversion of the cosmological X-ray transform acting on metric tensors. This is a more challenging problem due to the presence of a non-trivial null space for the transform, see [19, 31]. Another interesting problem is to develop accurate and efficient numerical methods for simulations on a larger scale, especially for the realistic 3D universe. In particular, the tomography method can be adapted to geometric backgrounds (instead of the flat Minkowski background) and handle CMB data collected along trajectories of the satellites.

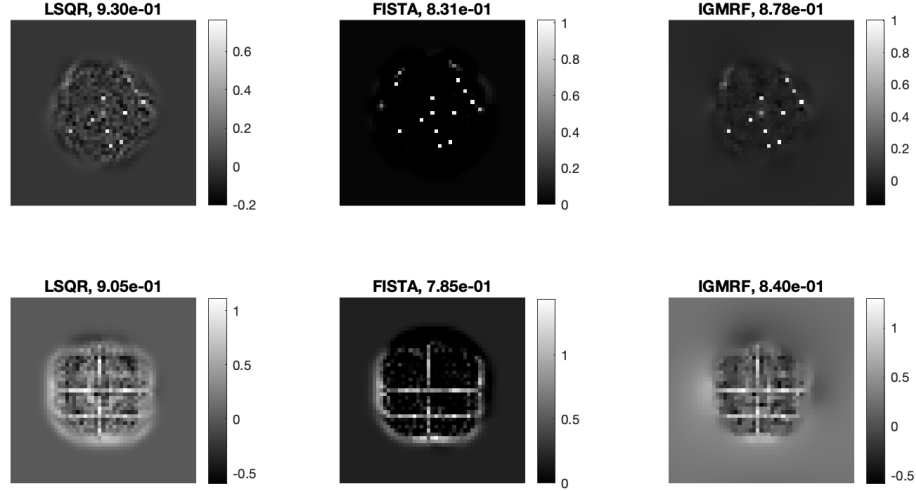


FIGURE 9. Reconstructions for LSQR, FISTA, and IGMRF for the partial data problem with a 3×3 detector grid. The top row contains reconstructions for the random dots example and the bottom row contains reconstructions for the random lines example. Relative reconstruction error norms are provided in the titles.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (NSF) under grants DMS-2411197 (J. Chung) and DMS-2205266 (Y. Wang).

REFERENCES

- [1] J. M. Bardsley. *Computational Uncertainty Quantification for Inverse Problems*. SIAM, 2018.
- [2] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [3] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, 2024.
- [4] V. Caselles, A. Chambolle, and M. M. Novaga. Total variation in imaging. *Handbook of mathematical methods in imaging*, 1(2):1455–1499, 2015.
- [5] J. Chung, L. Onisk, and Y. Wang. Iterative reconstruction methods for cosmological x -ray tomography. *SIAM Journal on Imaging Sciences*, to appear, 2025.
- [6] S. Dodelson. *Modern Cosmology*. Academic Press, Amsterdam (Netherlands), 2003.
- [7] F. X. Dupé, A. Rassat, J. L. Starck, and M. J. Fadili. Measuring the integrated Sachs-Wolfe effect. *Astronomy and Astrophysics*, 534:A51, 2011.
- [8] R. Durrer. *The Cosmic Microwave Background*. Cambridge University Press, Cambridge, UK, 2008.
- [9] S. Gazzola, P. C. Hansen, and J. G. Nagy. IR Tools: a MATLAB package of iterative regularization methods and large-scale test problems. *Numerical Algorithms*, 81(3):773–811, 2019.
- [10] G. González, V. Kolehmainen, and A. Seppänen. Isotropic and anisotropic total variation regularization in electrical impedance tomography. *Computers & Mathematics with Applications*, 74(3):564–576, 2017.
- [11] V. Guillemin. *Cosmology in $(2+1)$ -dimensions, cyclic models, and deformations of M_2* , 1. Number 121. Princeton University Press, 1989.
- [12] M. Hanke and P. C. Hansen. Regularization methods for large-scale problems. *Surv. Math. Ind*, 3(4):253–315, 1993.
- [13] P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, 2010.

- [14] L. Hörmander. *The analysis of linear partial differential operators IV: Fourier integral operators*. Springer, 2009.
- [15] J. A. Kable, G. Benevento, N. Frusciante, A. De Felice, and S. Tsujikawa. Probing modified gravity with integrated Sachs-Wolfe CMB and galaxy cross-correlations. *Journal of Cosmology and Astroparticle Physics*, 2022(09):002, 2022.
- [16] M. Krauss, S. Dodelson, and S. Meyer. Primordial gravitational waves and cosmology. *Science*, 328(5981):989–992, 2010.
- [17] A. Krolewski and S. Ferraro. The integrated Sachs Wolfe effect: unWISE and Planck constraints on dynamical dark energy. *Journal of Cosmology and Astroparticle Physics*, 2022(04):033, 2022.
- [18] H. P. Langtangen and S. Linge. *Finite difference computing with PDEs: a modern software approach*. Springer Nature, 2017.
- [19] M. Lassas, L. Oksanen, P. Stefanov, and G. Uhlmann. On the inverse problem of finding cosmic strings and other topological defects. *Communications in Mathematical Physics*, 357:569–595, 2018.
- [20] A. Manzotti and S. Dodelson. Mapping the integrated Sachs-Wolfe effect. *Physical Review D*, 90(12):123009, 2014.
- [21] R. B. Melrose. *Geometric scattering theory*, volume 1. Cambridge University Press, 1995.
- [22] J. Muir and D. Huterer. Reconstructing the integrated Sachs-Wolfe map with galaxy surveys. *Physical Review D*, 94:043503, 2016.
- [23] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger. Theory of cosmological perturbations. *Physics Reports*, 215(5-6):203–333, 1992.
- [24] C. C. Paige and M. A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software (TOMS)*, 8(1):43–71, 1982.
- [25] R. K. Sachs and A. M. Wolfe. Perturbations of a cosmological model and angular variations of the microwave background. *The Astrophysical Journal*, 1967.
- [26] A. Shajib and E. Wright. Measurement of the integrated Sachs–Wolfe effect using the allwise data release. *The Astrophysical Journal*, 827(2):116, 2016.
- [27] F. Trèves. *Introduction to pseudodifferential and Fourier integral operators Volume 2: Fourier integral operators*, volume 2. Springer Science & Business Media, 1980.
- [28] A. Vasy and Y. Wang. On the light ray transform of wave equation solutions. *Communications in Mathematical Physics*, 384:503–532, 2021.
- [29] Y. Wang. Microlocal analysis of the light ray transform on globally hyperbolic Lorentzian manifolds. *arXiv preprint arXiv:2104.08576*, 2021.
- [30] Y. Wang. Some integral geometry problems for wave equations. *Inverse Problems*, 38(8):084001, 2022.
- [31] Y. Wang. Inverse problems in cosmological x-ray tomography. *Microlocal Analysis and Inverse Problems in Tomography and Geometry, Radon Series on Computational and Applied Mathematics*, 30:139–165, 2024.

JULIANNE CHUNG

DEPARTMENT OF MATHEMATICS, EMORY UNIVERSITY
 Email address: `jmchung@emory.edu`

YIRAN WANG

DEPARTMENT OF MATHEMATICS, EMORY UNIVERSITY
 Email address: `yiran.wang@emory.edu`