# LOCAL CHARACTERIZATION OF NOISE IN ITERATIVE RECONSTRUCTION OF THE GENERALIZED RADON TRANSFORM

#### ALEXANDER KATSEVICH1

ABSTRACT. We study noise in iterative reconstruction from discrete noisy data of a generalized Radon transform in the plane. Our approach builds on Local Reconstruction Analysis (LRA), a framework for analyzing reconstructions at the native scale. We establish that the rescaled reconstruction error converges in distribution to a zero-mean Gaussian random field with explicitly computable covariance, providing a complete local characterization of noise in iterative reconstruction. Numerical experiments show strong agreement with the theoretical predictions. Combined with earlier deterministic results, our findings complete the analysis of iterative reconstruction at the native scale with respect to the two most fundamental limitations: the discreteness of the data and the presence of noise.

# 1. Introduction

1.1. Two strands of Local Reconstruction Analysis (LRA). Main results. Let  $\mathcal{R}$  denote the generalized Radon transform (GRT), which integrates a function f defined on an open set  $\mathcal{U} \subset \mathbb{R}^n$  (the image domain) over a family of surfaces parameterized by points in a set  $\mathcal{V} \subset \mathbb{R}^n$  (the data domain). The GRT data of f, denoted  $\hat{f}(y) = (\mathcal{R}f)(y)$  for  $y \in \mathcal{V}$ , are sampled on a rectangular grid  $\{y_j\}_{j\in\mathbb{Z}^n} \subset \mathcal{V}$  with step size proportional to a small parameter  $\epsilon > 0$ . If the data are noisy, the measurements take the form

(1.1) 
$$g_j := (\mathcal{R}f)(y_j) + v_j, \ j \in J := \{ j \in \mathbb{Z}^n : y_j \in \mathcal{V} \},$$

where  $\{v_i\}_{i\in J}$  are independent random variables.

We begin by considering reconstructions of the form  $f_0 = \mathcal{R}^*\mathcal{B}\hat{f}$ , where  $\mathcal{R}^*$  is the adjoint transform and  $\mathcal{B}$  is a pseudo-differential operator ( $\Psi$ DO). Let  $\varphi$  be a compactly supported, sufficiently smooth interpolation kernel, and let  $g_{\epsilon}$  denote the interpolant of the discrete data  $\{g_j\}_{j\in J}$  (see (3.7)). The reconstruction from discrete data is then  $f_{\epsilon} = \mathcal{R}^*\mathcal{B}g_{\epsilon}$ . This is of filtered backprojection (FBP) type: interpolation and the application of  $\mathcal{B}$  constitute the filtering step, while  $\mathcal{R}^*$  performs backprojection.

LRA, introduced by the author in [18], provides a systematic framework for studying local properties of such reconstructions. The central idea is to derive an explicit asymptotic formula for  $f_{\epsilon}$  in an  $\epsilon$ -neighborhood of a point

<sup>2020</sup> Mathematics Subject Classification. 44A12, 60G60, 65R32.

Key words and phrases. Generalized Radon transform, discrete data, random noise, reconstruction error, Gaussian random field, iterative reconstruction.

<sup>&</sup>lt;sup>1</sup>Department of Mathematics, University of Central Florida, Orlando, FL 32816 (Alexander.Katsevich@ucf.edu).

 $x_0 \in \mathcal{U}$ . This local approximation serves as a surrogate reconstruction at the native  $\sim \epsilon$  scale: once available, it can be readily analyzed from any perspective – edge smoothing, random-error covariance, aliasing, or feature detection – without further algorithm-specific derivations.

Two main strands of LRA have been developed:

- In the deterministic setting, the data are noise-free  $(v_j \equiv 0)$ , and one studies effects such as edge smoothing and aliasing artifacts in  $f_{\epsilon}$  [19, 20, 22].
- In the *stochastic setting*, by linearity, one may assume  $f \equiv 0$  and analyze reconstructions from noise  $(g_j = v_j)$ . This yields a detailed description of the local reconstruction error caused solely by random noise [2, 21].

Since exact inversion formulas exist for only a few GRTs, iterative reconstruction is often necessary. In a recent work by the author [23], LRA was extended to iterative reconstruction for discrete GRT data in the plane. The reconstruction  $f_{\epsilon}$  is obtained as the solution of the quadratic minimization problem

$$(1.2) \ f_{\epsilon} = \operatorname{argmin}_{f \in H_0^1(\mathcal{U}_b)} \Psi(f), \quad \Psi(f) := \|\mathcal{R}f - g_{\epsilon}\|_{L^2(\mathcal{V})}^2 + \kappa \epsilon^3 \|\partial_x f\|_{L^2(\mathcal{U})}^2,$$

where  $\kappa > 0$  is a fixed regularization parameter,  $\mathcal{U}_b \in \mathcal{U}$  is a subdomain, and  $H_0^1(\mathcal{U}_b)$  is the closure of  $C_0^{\infty}(\mathcal{U}_b)$  in the Sobolev  $H^1(\mathbb{R}^2)$  norm. The factor  $\epsilon^3$  in the Tikhonov term ensures that the reconstruction resolution remains at the native scale  $\epsilon$ .

The analysis in [23] concerns the deterministic setting  $(v_j \equiv 0)$ . The main result is an accurate formula describing how sharp edges in f are smoothed in  $f_{\epsilon}$ . The present paper complements that work by addressing the stochastic setting. Since  $\Psi$  is quadratic, the reconstruction depends linearly on the data. Thus it suffices to study the case  $f \equiv 0$  and  $g_j \equiv v_j$ . In this case the reconstructed image is denoted  $N_{\epsilon}^{\text{rec}}(x)$ . In the spirit of LRA, we focus on an  $O(\epsilon)$  neighborhood of a generic point  $x_0 \in \mathcal{U}_b \subset \mathbb{R}^2$ .

Let C(D) denote the space of continuous functions on a bounded Lipschitz domain  $D \subset \mathbb{R}^2$ . The main result of this paper is that, under suitable assumptions on the first four moments of  $\{v_j\}_{j\in J}$  and on  $x_0$ , the following limit exists:

(1.3) 
$$N^{\text{rec}}(\check{x}; x_0) = \lim_{\epsilon \to 0} N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}), \quad \check{x} \in \bar{D}.$$

Here,  $N^{\rm rec}$  and  $N^{\rm rec}_{\epsilon}$  are viewed as  $C(\bar{D})$ -valued random variables (random fields), and the limit is understood in the sense of probability distributions. We also prove that  $N^{\rm rec}(\check{x};x_0)$ ,  $\check{x}\in\bar{D}$ , is a zero-mean Gaussian random field (GRF) with explicitly computable covariance. Finally, we present numerical experiments demonstrating strong agreement between theoretical predictions and simulated reconstructions.

As in [23], our emphasis is on the most fundamental limitation to image quality: discrete noisy data. Although numerical implementation requires finite-dimensional approximation, its effect can be made negligible compared to data discretization and noise, so our analysis isolates the intrinsic contribution of noise.

1.2. **Related Work.** Classical analyses of noise in CT focus on analytic inversion schemes such as FBP. The book by Kak and Slaney [17, Sec. 5.2] derives explicit expressions for how additive measurement noise propagates through FBP, including the expectation and variance of the reconstructed image. A related treatment is given in [13, Sec. 16.2]. However, the resulting expressions are not available in closed form and are therefore difficult to use for detailed theoretical analysis.

Kernel-type estimators for tomographic problems with additive noise were studied in [27, 26], where minimax rates were obtained both pointwise and in  $L^2$ . Related results on maximal deviations appear in [6], while [8, 9] established pointwise asymptotic efficiency for density estimation from noise-free Radon data on random grids. A common feature of these works is the reliance on additional smoothing at a scale much larger than the sampling step, which entails loss of resolution and requires strong smoothness assumptions on f. None of them derive the distribution of the reconstruction error, whereas our analysis works at the native resolution and yields a complete local characterization of the error.

Bayesian inversion has also been developed for tomographic settings [28, 30, 35]. For Gaussian white noise, [28] showed asymptotic normality of the posterior and MAP estimator for smooth test functionals  $\int f\psi$ , i.e. reconstructions at macroscopic scales. By contrast, we quantify pointwise error uniformly over  $O(\epsilon)$  neighborhoods.

Another perspective is semiclassical analysis [31], which characterizes the empirical spatial mean and variance of reconstruction noise for a single experiment as  $\epsilon \to 0$ . Our approach differs by describing the entire limiting error distribution across repeated reconstructions.

Applied studies (e.g. [36, 10]) have examined image-domain noise with numerical and semi-empirical methods. These works provide valuable heuristics but do not yield theoretical PDFs of the reconstructed noise in small neighborhoods.

Finally, the existing work on stochastic LRA [2, 21] does characterize noise in the reconstructed image, but the analysis is confined to FBP-type reconstruction.

1.3. Significance. Organization of the paper. Empirical noise studies in tomography typically require many independent experiments: one generates noisy data, reconstructs, and then averages results to estimate statistics such as variance or correlation. For iterative methods, however, each reconstruction involves solving a large optimization problem, so repeating this process many times is computationally expensive.

The significance of this work is twofold. First, it eliminates the need for extensive empirical sampling. By extending LRA to the stochastic iterative setting, we obtain a rigorous description of the entire limiting distribution of the reconstruction error at the native scale. In effect, our analysis provides directly what empirical studies would reveal only after a prohibitive number of reconstructions. This advances the theoretical understanding of errors in iterative methods while also offering practical value: reliable predictions of reconstruction error without repeated costly simulations.

Second, this work broadens the scope of LRA itself. Until recently, LRA had been applied only to FBP-type algorithms, yet exact FBP inversion formulas exist for only a few transforms. By extending LRA to iterative reconstruction, we expand its applicability to the much wider class of GRTs for which only iterative methods are available. Taken together with the deterministic analysis in [23], the present results complete the study of reconstruction at the native scale with respect to the two most fundamental phenomena: data discreteness and noise. In turn, this opens the way to applying additional analyses to iterative reconstruction, such as the statistical microlocal analysis recently developed for FBP reconstruction [3].

The remainder of the paper is organized as follows. Section 2 introduces the basic notation. In Section 3 we describe the problem setup, state the assumptions, and present the iterative reconstruction algorithm. The main results, Theorems 4.2 and 4.4, are given in Section 4. Properties of the reconstructed image are examined in Section 5. Subsection 5.3 illustrates how our results apply to the classical Radon transform. The proofs of the main theorems are contained in Sections 6 and 7. Section 8 reports numerical experiments that confirm our theoretical predictions. Finally, the proofs of auxiliary results are collected in Appendices A–D.

# 2. Basic notation, function spaces, symbols, operators

Denote  $\mathbb{N} = \{1, 2, ...\}$  and  $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ . Let  $U \subset \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , be a domain, which is defined as a non-empty, connected, open set. For clarity, the zero vector in  $\mathbb{R}^n$ ,  $n \geq 2$ , is denoted  $0_n$ .

We identify tangent and cotangent spaces on Euclidean spaces with the underlying Euclidean spaces. Thus, for a function f(x) defined on a domain  $U \subset \mathbb{R}^n$  we have

$$\partial_x f(x) := f'_x(x) := (\partial_{x_1} f(x), \dots, \partial_{x_n} f(x))^T,$$

$$(2.1) \qquad d_x f = (\partial_{x_1} f(x), \dots, \partial_{x_n} f(x)),$$

$$|\partial_x f(x)| = |d_x f(x)| = \left[ (\partial_{x_1} f(x))^2 + \dots + (\partial_{x_n} f(x))^2 \right]^{1/2}.$$

We use  $f'_x$  or  $\partial_x f$  interchangeably depending on what is more convenient from the notational perspective.

The Sobolev space  $H^s(\mathbb{R}^n)$ ,  $s \in \mathbb{R}$ , is the space of all tempered distributions f for which

(2.2) 
$$||f||_{H^s(\mathbb{R}^n)}^2 = (2\pi)^{-n} \int_{\mathbb{R}^n} |\tilde{f}(\xi)|^2 (1+|\xi|^2)^s d\xi < \infty.$$

 $H_0^s(U)$  is the closure of  $C_0^\infty(U)$  in  $H^s(\mathbb{R}^n)$ .

for some constants  $\delta > 0$  and  $c_m, c_{m_1, m_2} > 0$ .

**Definition 2.1.** Given a domain  $U \subset \mathbb{R}^n$  and  $r \in \mathbb{R}$ ,  $S^r(U)$  denotes the vector-space of  $C^{\infty}(U \times (\mathbb{R}^n \setminus 0_n))$  functions,  $\tilde{B}(x,\xi)$ , having the following properties

$$(2.3) |\partial_x^m \tilde{B}(x,\xi)| \le c_m |\xi|^{-n+\delta} \ \forall m \in \mathbb{N}_0^n, x \in U, 0 < |\xi| \le 1; |\partial_x^{m_1} \partial_{\xi}^{m_2} \tilde{B}(x,\xi)| \le c_{m_1,m_2} |\xi|^{r-|m_2|}, \ \forall m_1 \in \mathbb{N}_0^n, m_2 \in \mathbb{N}_0^n, x \in U, |\xi| \ge 1;$$

The elements of  $S^r$  are called symbols of order r. We modify the conventional definition slightly to allow for symbols to be non-smooth at the origin. This makes our analysis more streamlined, but otherwise has no effects. The space of the corresponding  $\Psi DOs$ , given by

(2.4) 
$$(\operatorname{Op}(\tilde{B}(x,\xi))f)(y) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \tilde{B}(x,\xi)\tilde{f}(\xi)e^{-i\xi\cdot y} d\xi, \ f \in C_0^{\infty}(U),$$

is denoted  $L^r(U)$ . The notation  $\mathcal{U}_b \in \mathcal{U}$  means that the closure  $\overline{\mathcal{U}_b}$  is compact and  $\overline{\mathcal{U}_b} \subset \mathcal{U}$ .

For convenience, throughout the paper we use the following convention. If a constant c is used in an equation, the qualifier 'for some c>0' is assumed. If several c are used in a string of (in)equalities, then 'for some' applies to each of them, and the values of different c's may all be different.

Additional function spaces, notation, and conventions are introduced as needed.

- 3. Setting of the problem, assumptions, and reconstruction algorithm
- 3.1. **GRT and its properties.** Consider a GRT, which integrates over the curves  $S_y := \{x \in \mathcal{U} : p = \Phi(x, \alpha)\}, \ y := (\alpha, p) \in \mathcal{V} := \mathcal{I}_{\alpha} \times \mathcal{I}_p$ :

(3.1) 
$$(\mathcal{R}f)(y) = \int_{\mathcal{U}} f(x)W(x,y)\delta(p - \Phi(x,\alpha))dx$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathcal{U}} f(x)W(x,y)e^{i\nu(p-\Phi(x,\alpha))}dxd\nu, \ y \in \mathcal{V},$$

where  $W \in C^{\infty}(\mathcal{U} \times \mathcal{V})$ . An open set  $\mathcal{U} \subset \mathbb{R}^2$  is the image domain, and  $\mathcal{V} = \mathcal{I}_{\alpha} \times \mathcal{I}_p$  is the data domain. Here and everywhere below we use the convention that whenever y,  $\alpha$ , and p appear in the same equation or sentence, then  $y = (\alpha, p)$ .

**Assumption 3.1** (Properties of the GRT - I).

- (1)  $\mathcal{I}_{\alpha} \subset \mathbb{R}$  is a closed interval with the endpoints identified, so it can be viewed as a circle,
- (2)  $\mathcal{I}_p \subset \mathbb{R}$  is an open interval (or all of  $\mathbb{R}$ ), and
- (3)  $\Phi \in C^{\infty}(\mathcal{U} \times \mathcal{I}_{\alpha})$  and  $\Phi(\mathcal{U} \times \mathcal{I}_{\alpha}) \subset \mathcal{I}_{p}$ .

Both  $\mathcal{U}$  and  $\mathcal{V}$  are endowed with the usual Euclidean metric.

**Assumption 3.2** (Properties of the GRT - II).

- (1)  $d_x \Phi(x, \alpha)$  and  $d_x \Phi'_{\alpha}(x, \alpha)$  are linearly independent on  $\mathcal{U} \times \mathcal{I}_{\alpha}$ .
- (2) The Bolker condition: if  $\Phi(x_1, \alpha) = \Phi(x_2, \alpha)$  for some  $x_1, x_2 \in \mathcal{U}$ ,  $x_1 \neq x_2$ , and  $\alpha \in \mathcal{I}_{\alpha}$ , then  $\Phi'_{\alpha}(x_1, \alpha) \neq \Phi'_{\alpha}(x_2, \alpha)$ .
- (3) For each  $(x,\xi) \in T^*\mathcal{U}$ , there exist  $\alpha \in \mathcal{I}_{\alpha}$  and  $\nu \in \mathbb{R}$  such that  $\xi = \nu d_x \Phi(x,\alpha)$ .
- (4)  $W \in C^{\infty}(\mathcal{U} \times \mathcal{V})$  and W(x, y) > c on  $\mathcal{U} \times \mathcal{V}$ .
- (5) For any domain  $\mathcal{U}_h \subseteq \mathcal{U}$  there exists  $c = c(\mathcal{U}_h)$  such that

(3.2) 
$$||f||_{H^{-1/2}(\mathbb{R}^2)} \le c ||\mathcal{R}f||_{L^2(\mathcal{V})} \text{ for any } f \in H_0^{-1/2}(\mathcal{U}_b).$$

Assumption 3.2(1) implies that the curves  $S_y$  are smooth. See [23] for a more detailed discussion of these assumptions.

3.2. **GRT** data and the reconstruction algorithm. Realistic data can usually be represented as the sum of a useful signal (the GRT of some function,  $\hat{f} = \mathcal{R}f$ ) and noise. Hence we assume that our data are  $g(y_j) = \hat{f}(y_j) + v_j$ , where

$$(3.3) y_j = (\Delta \alpha j_1, \Delta p j_2) \in \mathcal{V}, \ \Delta p = \epsilon, \ \Delta \alpha = \mu \epsilon, \ j = (j_1, j_2) \in J,$$

 $v_j$  represents random noise, and  $J:=\{j\in\mathbb{Z}^2:y_j\in\mathcal{V}\}$ . Likewise,  $J_\alpha:=\{j_1\in\mathbb{Z}:\alpha_{j_1}\in\mathcal{I}_\alpha\}$  and  $J_p:=\{j_2\in\mathbb{Z}:p_{j_2}\in\mathcal{I}_p\}$ . For simplicity, the dependence of J,  $J_\alpha$ , and  $J_p$  on  $\epsilon$  is suppressed.

We assume that f is a sufficiently regular function supported in  $\mathcal{U}_b \subseteq \mathcal{U}$ . Since we study noise, the only property of f needed here is that  $\|\hat{f}\|_{L^{\infty}(\mathcal{V})} \leq C$  for some *known* constant C. For the complete set of assumptions on f required for image resolution analysis, see [23, Assumption 3.3].

Let  $\mathbb{E}v$  denote the expectation of a random variable v. We also use the shorthand notation  $\bar{v} = \mathbb{E}v$ .

# **Assumption 3.3** (Properties of noise $v_i$ ).

- (1)  $v_j$  are independent (but not necessarily identically distributed) random variables,
- (2) One has

(3.6)

(3.4) 
$$\mathbb{E}v_j = 0$$
,  $\mathbb{E}v_j^2 = \sigma^2(y_j)\Delta\alpha$ ,  $\mathbb{E}|v_j|^3 = O(\epsilon^{3/2})$ ,  $\mathbb{E}v_j^4 = O(\epsilon^{3/2})$ , where  $\sigma \in L^{\infty}(\mathcal{V})$ , and all the big- $O$  terms are uniform in  $j \in J$  as  $\epsilon \to 0$ .

Prior to reconstruction, we modify the data  $g(y_j)$  by applying a hard-thresholding procedure: we keep  $g(y_j)$  if  $|g(y_j)| \leq 2C$ , and set  $g(y_j) = 0$  if  $|g(y_j)| > 2C$ . This is equivalent to replacing the original noise  $v_j$  with a new noise  $\eta_j$ , defined as follows:

(3.5) 
$$\eta_{j} := \begin{cases} v_{j}, & \text{if } |g(y_{j})| \leq 2C \\ -\hat{f}(y_{j}), & \text{if } |g(y_{j})| > 2C, \end{cases} j \in J.$$

Thus, instead of  $g(y_j)$ , our data are  $f(y_j) + \eta_j$ .

The proof of the following result is in Appendix A.

**Lemma 3.4.** The random variables  $\eta_j$ ,  $j \in J$ , are independent and they satisfy

If 
$$v_j = 0$$
 for some  $j \in J$ , then  $\eta_j = 0$  for that  $j$ ,  

$$\mathbb{E}\eta_j = O(\epsilon^{3/2}), \ \mathbb{E}\eta_j^2 = \left(\sigma^2(y_j) + O(\epsilon^{1/2})\right)\Delta\alpha,$$

$$\mathbb{E}|\eta_j - \bar{\eta}_j|^3 = O(\epsilon^{3/2}),$$

where the big-O terms are uniform in  $j \in J$  as  $\epsilon \to 0$ .

The interpolated data, denoted  $g_{\epsilon}(y)$ , is computed by:

(3.7) 
$$g_{\epsilon}(y) := \sum_{y_{j} \in \mathcal{V}} \varphi_{\epsilon}(y - y_{j})(\hat{f}(y_{j}) + \eta_{j}), \ y \in \mathcal{V},$$
$$\varphi_{\epsilon}(y) := \varphi(y/\epsilon), \ \varphi(y) := \varphi_{\alpha}(\alpha/\mu)\varphi_{p}(p),$$

where  $\varphi$  is an interpolation kernel.

**Assumption 3.5** (Properties of the interpolation kernel,  $\varphi$ ). One has

- (1)  $\varphi_{\alpha}$  and  $\varphi_{p}$  are compactly supported,  $\varphi_{\alpha} \in L^{\infty}(\mathbb{R})$ , and  $\varphi_{p}^{(3)} \in L^{\infty}(\mathbb{R})$ .
- (2)  $\varphi_*$  is even:  $\varphi_*(u) = \varphi_*(-u), u \in \mathbb{R}, * = \alpha, p$ .

Given that the set  $\Phi(\mathcal{U}_b, \mathcal{I}_\alpha)$  is bounded and the modified data  $g_\epsilon$  are bounded due to the hard thresholding procedure, it follows that  $\|g_\epsilon\|_{L^2(\mathcal{V})} \le c$ ,  $0 < \epsilon \ll 1$ , with probability 1. The reconstruction is achieved by minimizing the quadratic functional

(3.8) 
$$f_{\epsilon} = \operatorname{argmin}_{f \in H_0^1(\mathcal{U}_b)} \Psi(f), \ \Psi(f) := \|\mathcal{R}f - g_{\epsilon}\|_{L^2(\mathcal{V})}^2 + \kappa \epsilon^3 \|\partial_x f\|_{L^2(\mathcal{U})}^2,$$

with some fixed regularization parameter  $\kappa > 0$ . [23, Lemma 3.5] yields the following result.

**Lemma 3.6.** Suppose Assumptions 3.1–3.5 are satisfied. With probability 1, the solution to (3.8) exists and is unique for each  $0 < \epsilon \ll 1$ .

The minimization problem in (3.8) is quadratic, hence the reconstructed image depends linearly on input data (see also (5.2) below). More precisely, we talk here about the modified data in (3.7). Therefore, in the rest of the paper we assume that  $\hat{f}(y_j) \equiv 0$ , and the data consist only of noise:  $g(y_j) = \eta_j$ ,  $j \in J$ . Thus,

(3.9) 
$$g_{\epsilon}(y) = \sum_{j \in J} \varphi_{\epsilon}(y - y_j) \eta_j, \ y \in \mathcal{V}.$$

The corresponding reconstructed image is denoted  $N_{\epsilon}^{\rm rec}$ .

We study  $N_{\epsilon}^{\text{rec}}(x)$  in an  $\epsilon$ -neighborhood of a point  $x_0$ , which is assumed to satisfy the following conditions:

**Assumption 3.7** (Properties of  $x_0$ ).

- (1) The equation  $\Phi''_{\alpha\alpha}(x_0,\alpha) = 0$ ,  $\alpha \in \mathcal{I}_{\alpha}$ , has finitely many solutions.
- (2)  $\sigma(y)$  is Lipschitz continuous in a neighborhood of the set

$$(3.10) \Gamma_{x_0} := \{ y \in \mathcal{V} : x_0 \in \mathcal{S}_y \}.$$

(3) There exists an open set  $I_{\alpha} \in \mathcal{I}_{\alpha}$  such that

(3.11) 
$$\sigma(\alpha, \Phi(x_0, \alpha)) \neq 0 \ \forall \alpha \in I_{\alpha}.$$

#### 4. Main results

Pick any  $L \in \mathbb{N}$  and choose L distinct points  $\check{x}_1, \dots, \check{x}_L \in \mathbb{R}^n$ . The corresponding reconstructed vector is

(4.1) 
$$\vec{N}_{\epsilon}^{\text{rec}} := (N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}_1), \dots, N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}_L)) \in \mathbb{R}^L.$$

For simplicity, the dependence of  $\vec{N}_{\epsilon}^{\text{rec}}$  on  $x_0$  and  $\check{x}_l$  is suppressed from notation. We remind the reader the notion of weak convergence of random variables.

**Definition 4.1** ([25, p. 185]). Suppose  $X_n$ ,  $1 \le n \le \infty$ , are random variables. Then  $X_n$  converges to  $X_\infty$  weakly, which is denoted  $X_n \Rightarrow X_\infty$ , as  $n \to \infty$  if the distribution of  $X_n$  converges to that of  $X_\infty$ .

**Theorem 4.2.** Let  $x_0 \in \mathcal{U}$  and  $\check{x}_l \in \mathbb{R}^n$ , l = 1, 2, ..., L, be fixed. Suppose Assumptions 3.1, 3.2, 3.3, 3.5, and 3.7 are satisfied. Suppose f is compactly supported and  $\|\mathcal{R}f\|_{L^{\infty}(\mathcal{V})} \leq C$  for some known C. The limit  $\vec{N}^{rec} := \lim_{\epsilon \to 0} \vec{N}^{rec}_{\epsilon}$  exists in the sense of weak convergence and is a Gaussian random vector.

The theorem is proven in Section 6. Before we state our next result, we remind the reader the definition of a random field (also known as a random function or topological space-valued random variable).

**Definition 4.3.** [25, p. 182] Let T be a topological space, endowed with its Borel field  $\mathcal{B}(T)$ . A T-valued random variable X on the probability space  $(\Omega, \mathcal{G}, \mathbb{P})$  is a measurable map  $X : \Omega \to T$ . In other words, for all  $E \in \mathcal{B}(T)$ ,  $\{\omega \in \Omega : X(\omega) \in E\} \in \mathcal{G}$ .

Let  $D \subset \mathbb{R}^n$  be a bounded Lipschitz domain. Define  $C := C(\bar{D}, \mathbb{R})$  to be the collection of all functions continuous up to the boundary  $f : \bar{D} \to \mathbb{R}$  metrized by

(4.2) 
$$d(f,g) = \max_{\check{x} \in D} |f(\check{x}) - g(\check{x})|, \ f, g \in C.$$

Recall that N(x),  $x \in \bar{D}$ , is a GRF if  $(N(x_1), \dots, N(x_L))$  is a Gaussian random vector for any  $L \geq 1$  and any collection of points  $x_1, \dots, x_L \in \bar{D}$  [4, Section 1.7]. As is known, a GRF is completely characterized by its mean  $\bar{G}(x)$  and covariance function  $\mathrm{Cov}(x,y) = \mathbb{E} \left[ (N(x) - \bar{N}(x))(N(y) - \bar{N}(y)) \right], x, y \in \bar{D}$  [4, Section 1.7]. Thus, Theorem 4.2 implies that  $N^{\mathrm{rec}}(\check{x}), \ \check{x} \in \bar{D}$ , is a zero mean GRF. For simplicity, we drop the dependence of  $N^{\mathrm{rec}}$  and its covariance function on  $x_0$  from notation.

In the next theorem, we show that  $N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}) \Rightarrow N^{\text{rec}}(\check{x}), \epsilon \to 0$ , as C-valued random variables ([25, p. 185]).

**Theorem 4.4.** Let D be a bounded domain with a Lipschitz boundary. Suppose the assumptions of Theorem 4.2 hold. Let the function G be defined as in (5.11). Then  $N_{\epsilon}^{rec}(x_0 + \epsilon \check{x}) \Rightarrow N^{rec}(\check{x})$ ,  $\check{x} \in \bar{D}$ ,  $\epsilon \to 0$ , as GRFs, i.e. in the sense of weak convergence of C-valued random variables. Furthermore,  $N^{rec}(\check{x})$ ,  $\check{x} \in \bar{D}$ , is a GRF with zero mean and covariance

$$(4.3) Cov(\check{x},\check{y}) = C(\check{x} - \check{y}),$$

$$C(\check{x}) := \int_{\mathcal{I}_{\alpha}} (G \star G) (x_0, \alpha, d_x \Phi(x_0, \alpha) \cdot \check{x}) \sigma^2(\alpha, \Phi(x_0, \alpha)) d\alpha,$$

$$(G \star G)(x_0, \alpha, p) := \int_{\mathbb{R}} G(x_0, \alpha, p + t) G(x_0, \alpha, t) dt,$$

and sample paths of  $N^{rec}(\check{x})$  are continuous with probability 1.

The theorem is proven in Section 7.

#### 5. Regularized solution

5.1. Equation for the solution to (3.8). Let  $\mathcal{R}^*$  be the adjoint of  $\mathcal{R}$ :  $L^2(\mathcal{U}) \to L^2(\mathcal{V})$ :

(5.1) 
$$(\mathcal{R}^*g)(x) = \int_{\mathcal{V}} g(y)W(x,y)\delta(p - \Phi(x,\alpha))\mathrm{d}y$$

$$= \int_{\mathcal{I}_{\alpha}} g(\alpha, \Phi(x,\alpha))W(x, (\alpha, \Phi(x,\alpha)))\mathrm{d}\alpha, \ x \in \mathcal{U}.$$

Assumption 3.2 ensures that  $\mathcal{R}^*\mathcal{R}$  is an elliptic  $\Psi DO$  [29, 16], [34, Section VIII.6.2] of order -1 in  $\mathcal{U}$ , i.e.  $\mathcal{R}^*\mathcal{R} \in L^{-1}(\mathcal{U})$ .

As is shown in [23], the solution to (3.8) is the unique solution to the following equation

$$(5.2) (\mathcal{R}^*\mathcal{R} - \kappa \epsilon^3 \Delta) N_{\epsilon}^{\text{rec}} = \mathcal{R}^* g_{\epsilon} \text{ in } H^{-1}(\mathcal{U}_b), \ N_{\epsilon}^{\text{rec}} \in H_0^1(\mathcal{U}_b).$$

Here we use that  $(H_0^1(\mathcal{U}_b))^* = H^{-1}(\mathcal{U}_b)$ ; see [32, Chapter 4, Proposition 5.1]. Recall that elements of  $H^s(\Omega)$  (for a domain  $\Omega$ ) can be viewed as equivalence classes of functions in  $H^s(\mathbb{R}^n)$ , where two functions are identified if they agree on  $\Omega$  [11, Section 1.6.1]. We also use that  $C^{\infty}(\overline{\Omega})$  is dense in  $H^{-s}(\Omega)$  for s < 0 [32, Exercise 7, Section 4.5].

Recall that we write  $N_{\epsilon}^{\text{rec}}$  instead of  $f_{\epsilon}^{\text{rec}}$  because  $g(y_j) = \eta_j, j \in J$ .

**Lemma 5.1.** Let  $\psi_{\epsilon}[h]$  be the solution to (3.8) (equivalently, (5.2)) with  $g_{\epsilon}$  replaced by  $h \in L^2(\mathcal{V})$ . One has

(5.3) 
$$\|\psi_{\epsilon}[h]\|_{H^{-1/2}(\mathcal{U})} \le c\|h\|_{L^{2}(\mathcal{V})}$$

for any  $0 < \epsilon \ll 1$  and  $h \in L^2(\mathcal{V})$ .

Proof. Since  $\Psi(\psi \equiv 0) = \|h\|_{L^2(\mathcal{V})}^2$  (cf. (3.8)), the solution  $\psi_{\epsilon}$  satisfies  $\|\mathcal{R}\psi_{\epsilon}\|_{L^2(\mathcal{V})} \leq 2\|h\|_{L^2(\mathcal{V})}$ ,  $0 < \epsilon \ll 1$ . By construction,  $\psi_{\epsilon} \in H_0^1(\mathcal{U}_b) \subset H_0^{-1/2}(\mathcal{U}_b)$ , where  $\mathcal{U}_b$  is bounded. Application of (3.2) completes the proof.

Let  $\tilde{Q}$  be the complete symbol of  $\mathcal{R}^*\mathcal{R}$ , i.e.  $\mathcal{R}^*\mathcal{R} = \operatorname{Op}(\tilde{Q}(x,\xi))$  (see [15, Theorem 3.4]). Let  $\mathcal{U}_0$  be a small neighborhood of  $x_0$ . Pick a function  $\chi \in C_0^{\infty}(\mathcal{U})$  such that  $\chi(x) \equiv 1$  in  $\mathcal{U}_0$ .

Let  $\mathcal{B}_{\epsilon} \in L^{-2}(\mathcal{U})$  be a parametrix for the operator on the left in (5.2). Additionally,  $\mathcal{B}_{\epsilon} \in L^{1}(\mathcal{U})$ , and its seminorms (i.e., the smallest constants  $c_{m}$ ,  $c_{m_{1},m_{2}}$  in (2.3)) are bounded with respect to  $0 < \epsilon \ll 1$ . Thus,

(5.4) 
$$N_{\epsilon}^{\text{rec}}(x) = \left(\operatorname{Op}\left(\tilde{B}_{\epsilon}(x,\xi)\right)\chi(z)\mathcal{R}^{*}g_{\epsilon}\right)(x) + \left(\operatorname{Op}\left(\tilde{B}_{\epsilon}(x,\xi)\right)(1-\chi(z))\mathcal{R}^{*}g_{\epsilon}\right)(x) + \int_{\mathcal{U}_{b}} T_{\epsilon}(x,x-z)N_{\epsilon}^{\text{rec}}(z)\mathrm{d}z, \ x \in \mathcal{U}_{0}.$$

Here T(x, w) is the smooth kernel of a regularizing  $\Psi$ DO  $\mathcal{T}_{\epsilon}$ , whose existence follows from the parametrix construction [33, Appendix to Section I.4].

5.2. Approximation of  $N_{\epsilon}^{\text{rec}}$ . It is noted in [23, (9.1) and Appendix C.2] that

(5.5) 
$$\tilde{Q}_0(x,\xi) := \lim_{\epsilon \to 0} |\xi/\epsilon| \tilde{Q}(x,\xi/\epsilon) = 2\pi \sum_k \frac{|\xi|}{|\nu|} \frac{W^2(x,y)}{|\Delta_{\Phi}(x,\alpha)|},$$
$$y = y_k(x,\xi), \nu = \nu_k(x,\xi), \ (x,\xi) \in \mathcal{U} \times (\mathbb{R}^2 \setminus 0_2).$$

Here

(5.6) 
$$\Delta_{\Phi}(x,y) = \det \begin{pmatrix} d_x \Phi(x,\alpha) \\ d_x \Phi'_{\alpha}(x,\alpha) \end{pmatrix},$$

and the functions  $y_k(x,\xi) \in \mathcal{V}$  and  $\nu_k(x,\xi) \in \mathbb{R}$  are local solutions to the equations:

(5.7) 
$$p = \Phi(x, \alpha), \ \xi = -\nu d_x \Phi(x, \alpha), \ (\alpha, p) \in \mathcal{V}.$$

The existence of at least one local solution follows from Assumption 3.2(3). The local smoothness of solutions follows from Assumption 3.2(1). The function  $\tilde{Q}_0(x,\xi)$  is positively homogeneous of degree zero in  $\xi$ .

Denote

(5.8) 
$$\tilde{B}_0(x,\hat{\xi}) := \frac{|\hat{\xi}|}{\tilde{Q}_0(x,\hat{\xi}) + \kappa |\hat{\xi}|^3}, \ x \in \mathcal{U}, \hat{\xi} \in \mathbb{R}^2 \setminus 0_2.$$

Denote further

(5.9) 
$$K_0(\check{t}; x, \alpha) := \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{B}_0(x, \lambda d_x \Phi(x, \alpha)) e^{-i\lambda \check{t}} d\lambda,$$

(5.10) 
$$\vartheta(\check{t}; x, \alpha) = \int_{\text{supp }\varphi_{\alpha}} \varphi_{p} (\check{t} + \mu \Phi'_{\alpha}(x, \alpha)\check{\alpha}) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha},$$

$$(5.11) \qquad G(x,\alpha,\check{q}) := \mu W(x,(\alpha,\Phi(x,\alpha))) \int_{\mathbb{R}} K_0(\check{q}-\check{t};x,\alpha)\vartheta(\check{t};x,\alpha)d\check{t}.$$

The following result is proven in Appendix B.

Lemma 5.2. Under the assumptions of Theorem 4.2 one has

(5.12) 
$$N_{\epsilon}^{rec}(x_0 + \epsilon \check{x}) = \sum_{j \in J} \left[ G(x_0, \alpha_{j_1}, r_{j_1} - j_2) + R_j \right] \eta_j,$$
$$r_{j_1} := \frac{\Phi(x_0, \alpha_{j_1})}{\epsilon} + d_x \Phi(x_0, \alpha_{j_1}) \cdot \check{x}, \ |R_j| \le c\epsilon \ln(1/\epsilon).$$

This lemma is a key ingredient in the proof of Theorem 4.2 in Section 6.

5.3. **Example.** As a brief initial illustration of Theorems 4.2 and 4.4, we consider the classical Radon transform in the plane, which integrates over lines  $S_y = \{x \in \mathbb{R}^2 : \vec{\alpha} \cdot x = p\}, \ \vec{\alpha} := (\cos \alpha, \sin \alpha)$ . In this case  $\mathcal{I}_{\alpha} = [0, 2\pi) \sim S^1, \ \mathcal{I}_p = \mathbb{R}, \ \Phi(x, \alpha) = \vec{\alpha} \cdot x, \ W(x, y) \equiv 1$ , and

(5.13) 
$$d_x \Phi(x, \alpha) = \vec{\alpha}, \ |d_x \Phi(x, \alpha)| = 1, \ \Phi'_{\alpha}(x, \alpha) = \vec{\alpha}^{\perp} \cdot x,$$
$$\Delta_{\Phi}(x, y) = \det \begin{pmatrix} d_x \Phi(x, \alpha) \\ d_x \Phi'_{\alpha}(x, \alpha) \end{pmatrix} = 1, \ \vec{\alpha}^{\perp} := (-\sin \alpha, \cos \alpha).$$

From (5.5) and (5.8):

(5.14) 
$$\tilde{Q}_0(x,\xi) = 4\pi, \ \tilde{B}_0(x,\hat{\xi}) = \frac{|\hat{\xi}|}{4\pi + \kappa |\xi|^3}.$$

Here we use that each singularity is visible in the data two times. Finally, from (5.9)-(5.11),

$$K_{0}(\check{t}; x, \alpha) = \frac{1}{\pi} \int_{0}^{\infty} \frac{\lambda}{4\pi + \kappa \lambda^{3}} \cos(\lambda \check{t}) d\lambda,$$

$$(5.15) \qquad \vartheta(\check{t}; x, \alpha) = \int_{\text{supp }\varphi_{\alpha}} \varphi_{p} (\check{t} + \mu(\alpha^{\perp} \cdot x)\check{\alpha}) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha},$$

$$G(x, \alpha, \check{q}) := \mu \int_{\mathbb{R}} K_{0}(\check{q} - \check{t}; x, \alpha) \vartheta(\check{t}; x, \alpha) d\check{t}.$$

See Section 8 for a more detailed and extensive example involving the GRT that integrates over circles.

# 6. Proof of Theorem 4.2

Recall that the random vector  $\vec{N}_{\epsilon}^{\text{rec}}$  is defined in (4.1). Pick any vector  $\vec{\theta} \in \mathbb{R}^L$ . By (5.12),

(6.1) 
$$\zeta_{\epsilon} := \vec{\theta} \cdot \vec{N}_{\epsilon}^{\text{rec}} = \sum_{l=1}^{L} \theta_{l} N_{\epsilon}^{\text{rec}}(x_{0} + \epsilon \check{x}_{l}) = \sum_{j \in J} \zeta_{\epsilon, j},$$

$$\zeta_{\epsilon, j} = \left[ \sum_{l=1}^{L} \theta_{l} \left[ G\left(\alpha_{j_{1}}, r_{j_{1}}^{(l)} - j_{2}\right) + R_{j}^{(l)} \right] \right] \eta_{j},$$

$$r_{j_{1}}^{(l)} := \frac{\Phi(x_{0}, \alpha_{j_{1}})}{\epsilon} + d_{x} \Phi(x_{0}, \alpha_{j_{1}}) \cdot \check{x}_{l}.$$

To prove that  $\vec{N}_{\epsilon}^{\text{rec}}$  converges in distribution to a Gaussian random vector, it suffices to show that for any  $\vec{\theta} \in \mathbb{R}^L \setminus 0_L$ ,  $\lim_{\epsilon \to 0} \zeta_{\epsilon}$  is a Gaussian random variable [5, Theorem 10.4.5]. The following result is proven in Appendix C.

**Lemma 6.1.** Suppose the assumptions of Theorem 4.2 are satisfied. One has

(6.2) 
$$\mathbb{E}N_{\epsilon}^{rec}(x_0 + \epsilon \check{x}) = O(\epsilon^{1/2}\ln(1/\epsilon)), \ \epsilon \to 0,$$

for any fixed  $\check{x}$ . Moreover, with  $\zeta_{\epsilon,j}$  defined in (6.1), one has

(6.3) 
$$\lim_{\epsilon \to 0} \frac{\sum_{j \in J} \mathbb{E} |\zeta_{\epsilon,j} - \bar{\zeta}_{\epsilon,j}|^3}{\left[\sum_{j \in J} Var(\zeta_{\epsilon,j})\right]^{3/2}} = 0.$$

Lemma 6.1 implies that the family of random variables,  $\zeta_{\epsilon,j} - \bar{\zeta}_{\epsilon,j}$ ,  $\epsilon > 0$ , satisfies the Lyapunov condition for triangular arrays [5, Definition 11.1.3]. It follows from [5, Corollary 11.1.4], the first line in (6.1), and (6.2) that  $\zeta = \lim_{\epsilon \to 0} \zeta_{\epsilon}$  is a zero mean Gaussian random variable. Hence, by [5, Theorem 10.4.5],  $\lim_{\epsilon \to 0} \vec{N}_{\epsilon}^{\text{rec}}$  is a zero mean Gaussian random vector, where the limit is in the sense of weak convergence. This completes the proof of Theorem 4.2.

# 7. Proof of Theorem 4.4

Our goal is to show that  $N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}) \Rightarrow N^{\text{rec}}(\check{x}), \ \check{x} \in D$ , as C-valued random variables as  $\epsilon \to 0$ . To do this, we use the following definition and theorem.

**Definition 7.1** ([25, p. 189]). Let  $\mathbb{P}_n$  be the distribution of a C-valued random variable  $X_n$ ,  $1 \leq n \leq \infty$ . The collection  $(X_n)$  is tight if for all  $\delta \in (0,1)$ , there exists a compact set  $\Gamma_{\delta} \in C$  such that  $\sup_{n} \mathbb{P}(X_n \notin \Gamma_{\delta}) \leq \delta$ .

By finite-dimensional distributions of a C-valued random variable  $X(\check{x})$ ,  $\check{x} \in D$ , we mean the collection of all the distributions of the  $\mathbb{R}^L$ -dimensional random variables  $\{X(\check{x}_1), X(\check{x}_2), \cdots, X(\check{x}_L)\}$  as we vary  $L \geq 1$  and the points  $\check{x}_1, \ldots, \check{x}_L \in D$ , see e.g. [25, p. 154].

**Theorem 7.2** ([25, Proposition 3.3.1, p. 197]). Suppose  $X_n$ ,  $1 \le n \le \infty$ , are C-valued random variables. Then  $X_n \Rightarrow X_\infty$  provided that:

- (1) Finite dimensional distributions of  $X_n$  converge to that of  $X_{\infty}$ .
- (2)  $(X_n)$  is a tight sequence.

The preceding definition and theorem are formulated for sequences of C-valued random variables indexed by n and the limit is taken as  $n \to \infty$ . With obvious modifications, these notions and results extend to families of random variables indexed by a continuous variable  $\epsilon > 0$  and when the limit is taken as  $\epsilon \to 0$ .

By Theorem 4.2, all the finite-dimensional distributions of  $N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x})$  converge to that of  $N^{\text{rec}}(\check{x})$ . Thus it remains to verify Property 2 of Theorem 7.2. The following result is proven in Appendix D.

**Lemma 7.3.**  $(N_{\epsilon}^{rec}(x_0 + \epsilon \check{x}))$  is a tight family.

By Theorem 7.2,  $N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}) \Rightarrow N^{\text{rec}}(\check{x})$  as C-valued random variables. Since C is a complete metric space, it follows that  $N^{\text{rec}}(\check{x})$  has continuous sample paths with probability 1.

By Lemma 6.1,  $N^{\text{rec}}(\check{x})$  is a zero mean GRF. To completely characterize this GRF, we calculate its covariance function

(7.1) 
$$\operatorname{Cov}(\check{x}, \check{y}) = \mathbb{E}(N^{\operatorname{rec}}(\check{x})N^{\operatorname{rec}}(\check{y})), \ \check{x}, \check{y} \in D.$$

In fact, essentially this has already been done in the proof of Lemma 6.1, see Section C.2. From (5.9)–(5.11) and (C.22) we obtain

(7.2) 
$$\operatorname{Cov}(\check{x}, \check{y}) = \mu^{2} \int_{\mathcal{I}_{\alpha}} W^{2}(x_{0}, \Phi(x_{0}, \alpha)) \sigma^{2}(\alpha, \Phi(x_{0}, \alpha)) \times \mathcal{F}_{\lambda \to r}(\left|\tilde{B}_{0}(x_{0}, \lambda d_{x} \Phi(x_{0}, \alpha))\tilde{\vartheta}(\lambda; x_{0}, \alpha)\right|^{2}) d\alpha, \ r = \alpha \cdot (\check{x} - \check{y}).$$

Here  $\vartheta$  is the 1D Fourier transform of  $\vartheta$  with respect to the first argument.

#### 8. Numerical experiments

The GRT integrates over circles  $S_y$ ,  $y = (\alpha, \rho)$ , with various radii  $\rho > 0$  and centers  $R\vec{\alpha}$ ,  $\alpha \in [0, 2\pi)$ , where R = 10. The integration weight is  $W \equiv 1$ . Thus,  $\rho = \Phi(x, \alpha) = |x - R\vec{\alpha}|$  and

$$d_{x}\Phi(x,\alpha) = \frac{x - R\vec{\alpha}}{|x - R\vec{\alpha}|} =: \vec{\Theta}, \ \Phi'_{\alpha}(x,\alpha) = R\vec{\Theta} \cdot \vec{\alpha}^{\perp},$$

$$(8.1) \qquad d_{x}\Phi'_{\alpha}(x,\alpha) = -\frac{R}{|x - R\vec{\alpha}|}\vec{\alpha}^{\perp} + c\vec{\Theta},$$

$$\Delta_{\Phi}(x,\alpha) = \det\left(\frac{\vec{\Theta}}{-\frac{R}{|x - R\vec{\alpha}|}}\vec{\alpha}^{\perp} + c\vec{\Theta}\right) = \frac{R\vec{\alpha} \cdot \vec{\Theta}}{|x - R\vec{\alpha}|} \neq 0, \ |x| < R.$$

Here  $\vec{\alpha}^{\perp} = (-\sin\alpha, \cos\alpha)$ , and  $\vec{\Theta}^{\perp}$  is defined similarly. The continuous data corresponds to  $\alpha \in \mathcal{I}_{\alpha} := [0, 2\pi) \sim S^1$  and  $\rho \in \mathcal{I}_p := (0, 2R)$ . The image domain is  $\mathcal{U} = \{|x| < R\}$ . The support of the object is contained in the square reconstruction region  $\mathcal{U}_b := (-R_{rec}, R_{rec})^2$ ,  $R_{rec} = 3.7$ . This corresponds to the set  $\mathcal{U}_b$  used in (3.8).

As is easily seen, for each  $(x,\xi) \in T^*\mathcal{U}_b \setminus 0_2$  there are two solutions,  $y_k(x,\xi)$ ,  $\nu_k(x,\xi)$ , k = 1, 2, to (5.7). To find  $y_{1,2}$ , we first solve  $|x+t\vec{\xi}| = R$ . This gives two values  $t_1, t_2, t_2 < 0 < t_1$ . Then  $\alpha_k$  is determined from  $x + t_k \xi = R\vec{\alpha}_k$ , and  $\rho_k = |x - R\vec{\alpha}_k|$ , k = 1, 2. By a simple geometric argument (cf. (8.1)),

(8.2) 
$$|R\vec{\alpha}_k \cdot \vec{\Theta}_k| = |R\vec{\alpha}_1 - R\vec{\alpha}_2|/2, \ k = 1, 2.$$

Further, substituting the above into (5.5) gives

(8.3) 
$$\tilde{Q}_{0}(x,\xi) = 2\pi \sum_{k=1}^{2} \frac{|\mathbf{d}_{x}\Phi(x,\alpha_{k})|}{|\Delta_{\Phi}(x,\alpha_{k})|} = 2\pi \left[ \frac{|x-R\vec{\alpha}_{1}|}{|R\vec{\alpha}_{1}\cdot\vec{\Theta}_{1}|} + \frac{|x-R\vec{\alpha}_{2}|}{|R\vec{\alpha}_{2}\cdot\vec{\Theta}_{2}|} \right]$$
$$= 2\pi \left[ \frac{|x-R\vec{\alpha}_{1}|}{|R\vec{\alpha}_{1}-R\vec{\alpha}_{2}|/2} + \frac{|x-R\vec{\alpha}_{2}|}{|R\vec{\alpha}_{1}-R\vec{\alpha}_{2}|/2} \right] = 4\pi.$$

Combining with (5.8)–(5.11) we obtain

(8.4) 
$$G(x,\alpha,q) = \frac{\mu}{2\pi} \int \frac{|\lambda|}{4\pi + \kappa |\lambda|^3} \tilde{\varphi}_{\rho}(\lambda) \tilde{\varphi}_{\alpha}(-\mu R \vec{\Theta} \cdot \vec{\alpha}^{\perp} \lambda) e^{-i\lambda q} d\lambda,$$

where  $\vec{\Theta}$  is the same as in (8.1).

It is shown in [23] that the circular GRT used in this section satisfies Assumptions 3.1 and 3.2. Since  $\Phi(x,\alpha)$  is analytic in  $\alpha$ , Assumption 3.7(1) holds for every  $x_0$  with  $|x_0| < R$ .

The discrete data are given at the points

$$\alpha_{j_1} = j_1 \Delta \alpha, \ \Delta \alpha = 2\pi/N_{\alpha}, \ 0 \le j_1 < N_{\alpha}, \ N_{\alpha} = 300,$$

$$(8.5) \ \rho_{j_2} = \rho_{\min} + j_2 \Delta \rho, \ \epsilon := \Delta \rho = (\rho_{\max} - \rho_{\min})/(N_{\rho} - 1), \ 0 \le j_2 < N_{\rho},$$

$$\rho_{\min} = R - R_{rec} \sqrt{2}, \ \rho_{\max} = R + R_{rec} \sqrt{2}, \ N_{\rho} = 451.$$

Before reconstruction, the GRT data  $g(y_j)$  are interpolated from the grid (8.5) to a denser grid using (3.7). This refined grid, where  $g_{\epsilon}$  is specified, has the same structure as (8.5) but with  $N'_{\alpha} > N_{\alpha}$  and  $N'_{\rho} > N_{\rho}$ . For this interpolation, we use the Keys interpolation kernel [24, 7]

(8.6) 
$$\varphi_{\alpha}(t) = \varphi(t/\mu), \ \varphi_{\rho}(t) = \varphi(t), \ \mu = \Delta \alpha/\Delta \rho, \\ \varphi(t) = 3B_3(t) - (B_2(t - 1/2) + B_2(t + 1/2)),$$

where  $B_n$  is the cardinal B-spline of degree n supported on [-(n+1)/2, (n+1)/2]. Therefore supp  $\varphi = [-2, 2]$ . The kernel is a piecewise-cubic polynomial with continuous  $\varphi, \varphi'$  and bounded  $\varphi''$ , so  $\varphi \in C_0^2(\mathbb{R})$ . As is well-known,

(8.7) 
$$(\mathcal{F}B_n)(\lambda) = \operatorname{sinc}(\lambda/2)^{n+1}, \ \operatorname{sinc}(t) := \sin t/t.$$

Therefore,

(8.8) 
$$\tilde{\varphi}(\lambda) = (\operatorname{sinc} t)^3 [3 \operatorname{sinc} t - 2 \operatorname{cos} t], \ t = \lambda/2.$$

Substituting (8.6) and (8.8) into (8.4) we find:

(8.9) 
$$G(x,\alpha,q) = \frac{\mu}{\pi} \int_0^\infty \frac{\lambda}{4\pi + \kappa \lambda^3} \tilde{\varphi}(\lambda) \tilde{\varphi}(-\mu a(x,\alpha)\lambda) \cos(\lambda q) d\lambda,$$

where  $a(x, \alpha)$  is the distance from the origin to the chord of the circle |x| = R passing through the points x and  $R\vec{\alpha}$ . Similarly, the covariance function is computed by (cf. (4.3), (7.2))

(8.10) 
$$C(\check{x}) = \frac{\mu^2}{\pi} \int_0^{2\pi} \int_0^{\infty} \left[ \frac{\lambda}{4\pi + \kappa \lambda^3} \tilde{\varphi}(\lambda) \tilde{\varphi}(-\mu a(x_0, \alpha) \lambda) \right]^2 \cos(\lambda \vec{\Theta} \cdot \check{x}) d\lambda \\ \times \sigma^2(\alpha, |x_0 - R\vec{\alpha}|) d\alpha, \ \vec{\Theta} = \frac{x_0 - R\vec{\alpha}}{|x_0 - R\vec{\alpha}|}.$$

In order to approximate the continuous GRT  $\mathcal{R}$  and its adjoint  $\mathcal{R}^*$ , the reconstruction is performed on a dense  $801 \times 801$  grid covering the same square region  $\mathcal{U}_b$ . The regularization parameter is set at  $\kappa = 0.5$ .

The functional (3.8) is minimized using gradient descent. The stopping criterion is met when either

(8.11) (i) 
$$||f_{k+1} - f_k||_{L^{\infty}(\mathcal{U}_b)} < 10^{-4} \text{ or (ii) } \Psi(f_k) \le 10^{-6} \Psi(f_0)$$

occurs three times (not necessarily consecutively). Here  $f_k$  denotes the image at the k-th iteration, and  $f_0(x) \equiv 0$ . Occurrences of (i) and (ii) are counted separately; mixed occurrences are not combined toward the total of three.

From the proof of Lemma 5.2 it follows that (5.12) holds for any choice of  $\mathcal{U}_b \in \mathcal{U}$ , provided  $x_0 \in \mathcal{U}_b$ . However, computing  $N_{\epsilon}^{\text{rec}}(x)$  by numerically solving (3.8) is sensitive to discretization errors. To mitigate this, we chose a smaller domain  $\mathcal{U}_b' = (1, 1.4) \times (0.5, 0.9)$  centered at the fixed point  $x_0 = (1.2, 0.7)$ . We set  $\eta_j = 1$  for a single index  $j_0 = (j_1, j_2) = (100, 229)$  and  $\eta_j = 0$  for all  $j \neq j_0$ . The index  $j_0$  was selected so that  $\rho_{j_2} \approx |x_0 - R\vec{\alpha}_{j_1}|$ , with  $\alpha_{j_1} = 2.1$  and  $\rho_{j_2} = 10.1$ . The data are then interpolated to a fine grid with  $N_\alpha' = 2400$  and  $N_\rho' = 3601$ . Finally, the reconstruction uses an  $801 \times 801$  grid covering  $\mathcal{U}_b'$  (rather than the original  $\mathcal{U}_b$ ). The large values of  $N_\alpha'$  and  $N_\rho'$ , together with a dense reconstruction grid on a small region, ensure high accuracy.

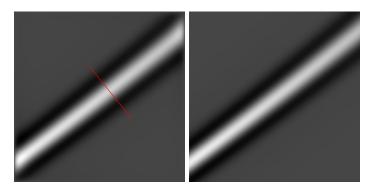


FIGURE 1. Left: iteratively reconstructed  $N_{\epsilon}^{\text{rec}}(x)$  in the square  $\mathcal{U}_b' = [1, 1.4] \times [0.5, 0.9]$ . Right:  $G(x_0, \alpha_{j_1}, (|x - R\vec{\alpha}_{j_1}| - \rho_{j_2})/\epsilon)$  for  $x \in \mathcal{U}_b'$ , computed using (8.9).

The results are shown in Figures 1, 2, and 3. The left panel of Figure 1 shows the density plot of the iteratively reconstructed  $N_{\epsilon}^{\rm rec}(x)$  for  $x \in \mathcal{U}_b'$ , while the right panel displays the function  $G(x_0, \alpha_{j_1}, (|x - R\vec{\alpha}_{j_1}| - \rho_{j_2})/\epsilon)$ ,  $x \in \mathcal{U}_b'$ , computed using (8.9), over the same domain. Both panels use a display window of [-0.02, 0.06]. Figure 2 shows the difference between these two images, with a tighter display window of  $[-5 \cdot 10^{-3}, 5 \cdot 10^{-3}]$ . Consistent with the lemma, the difference is negligible except near the boundary of  $\mathcal{U}_b'$ . To better illustrate the agreement, Figure 3 plots cross-sections of the three images along the red line shown in the left panel of Figure 1.



FIGURE 2. Difference between the two images shown in Figure 1, displayed with the window  $[-5 \cdot 10^{-3}, 5 \cdot 10^{-3}]$ .

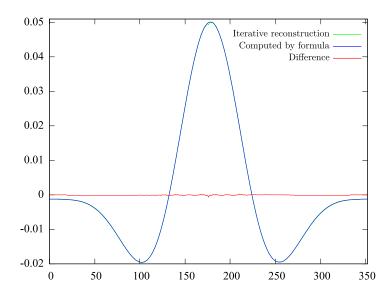


FIGURE 3. Cross-sections of the three images from Figures 1 and 2 along the red line indicated in the left panel of Figure 1.

To validate the predicted covariance function  $C(\check{x})$  in (4.3), we simulate a complete data set with noisy measurements of a test function f. We choose f to be the characteristic function of the disk centered at  $x_c = (1,1)$  with radius r = 2. The center of a local region of interest (ROI) is the

same point  $x_0 = (1.2, 0.7)$  as before. Then we simulate 1500 noisy data sets. For each data set, the noise values  $v_j$  are drawn from the uniform distribution on [-1, 1], which has a standard deviation of 1/3, and then scaled by  $(3\Delta\alpha)^{1/2}\sigma(y_j)$ , where

(8.12) 
$$\sigma(y) = (1 + 0.5\sin(2\alpha))(1 - 0.4\cos\rho)\chi_{\text{supp}(\mathcal{R}f)}(y).$$

Here  $\chi_{\text{supp}(\mathcal{R}f)}(y)$  is the characteristic function of the support of  $\mathcal{R}f$ . This choice ensures that Assumptions 3.3 and 3.7(2,3) are satisfied.

In the spirit of conventional X-ray CT, if the curve  $S_{y_j}$  does not intersect the disk, the corresponding noise value  $v_j$  is set to zero. Physically, in X-ray CT a beam passing only through air experiences negligible attenuation, so its associated noise level is much smaller than that of beams traversing the object and undergoing substantial attenuation. The discontinuity of  $\sigma(y)$ at the boundary of the projected disk in the data domain is not problematic, since this discontinuity remains at a positive distance from  $\Gamma_{x_0}$  (cf. Assumption 3.7(2)).

To reduce computation time, the data interpolation grid was reduced to  $N'_{\alpha} = 1200$ ,  $N'_{\rho} = 1801$ . The entire original  $\mathcal{U}_b = [-R_{rec}, R_{rec}]^2$  was reconstructed on an  $801 \times 801$  grid. We recorded the reconstructed values at the center  $x_0$  and at a nearby point  $x_1 = x_0 + \epsilon \check{x}$ , where  $\check{x} = (1.24, -1.77)$ . The sample variances at these points are 0.042 and 0.039, respectively, compared to the predicted value of 0.043. Figure 4 shows the sample and predicted histograms, which agree well. The sample covariance between  $N_{\epsilon}^{\text{rec}}(x_0)$  and  $N_{\epsilon}^{\text{rec}}(x_1)$  is 0.024, while the predicted value, computed using (4.3), is  $C(\check{x}) = 0.027$ .

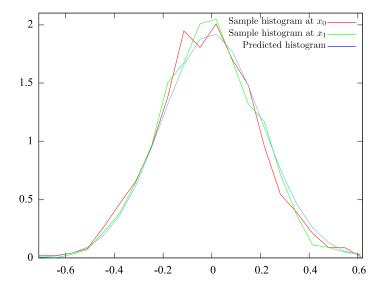


FIGURE 4. Sample histograms of the computed values  $N_{\epsilon}^{\rm rec}(x_0)$  (red) and  $N_{\epsilon}^{\rm rec}(x_1)$  (green), along with the predicted histogram (blue), which does not depend on the point.

# Appendix A. Proof of Lemma 3.4

Since each  $\eta_j$  is a function of only one of the independent variables  $v_j$ , the variables  $\eta_j$  are also independent. For simplicity, we drop the subscript j for the remainder of the proof and denote  $\hat{f} := \hat{f}(y_j)$ .

If v = 0, then  $|\hat{f}| \leq C$  implies that the corresponding g is not replaced with zero. In this case,  $\eta = v = 0$ .

Let  $\pi$  be the probability measure of  $\nu$ . By construction,

(A.1) 
$$\mathbb{E}|\eta - v| \le \int_{|x + \hat{f}| \ge 2C} |x + \hat{f}| d\pi(x) \le 2 \int_{|x| \ge C} |x| d\pi(x)$$

$$\le \frac{2}{C^2} \int_{|x| \ge C} |x|^3 d\pi(x) \le \frac{2}{C^2} \mathbb{E}|v|^3 = O(\epsilon^{3/2}).$$

Using the triangle inequality and that  $\mathbb{E}v = 0$  proves the claim about  $\mathbb{E}\eta$ . Arguing similarly we obtain

(A.2) 
$$\mathbb{E}|\eta^{2} - v^{2}| \leq \int_{|x+\hat{f}| \geq 2C} x^{2} d\pi(x) \leq \int_{|x| \geq C} x^{2} d\pi(x)$$

$$\leq \frac{1}{C} \int_{|x| \geq C} |x|^{3} d\pi(x) = O(\epsilon^{3/2}).$$

In the first inequality we used that  $|v^2 - \eta^2| = |x^2 - (-f)^2| \le x^2$  on the domain of integration. Combining with (3.4) proves the claim about  $\mathbb{E}\eta_j^2$ . Finally,

(A.3) 
$$|\eta - \bar{\eta}|^3 \le c(|\eta - v|^3 + |v - \bar{\eta}|^3) \le c(|\eta - v|^3 + |v|^3 + |\bar{\eta}|^3)$$
  
Similarly to (A.1),

(A.4) 
$$\mathbb{E}|\eta - v|^{3} \leq \int_{|x + \hat{f}| \geq 2C} |x + \hat{f}|^{3} d\pi(x) \leq 8 \int_{|x| \geq C} |x|^{3} d\pi(x)$$

$$\leq \frac{8}{C} \int_{|x| > C} x^{4} d\pi(x) \leq \frac{8}{C} \mathbb{E}v^{4} = O(\epsilon^{3/2}).$$

Taking the expectation on both sides of (A.3) and using the previously obtained results and (3.4) we prove the last claim of the lemma

(A.5) 
$$\mathbb{E}|\eta - \bar{\eta}|^3 = O(\epsilon^{3/2}).$$

Appendix B. Proof of Lemma 5.2

In view of the first line in (5.4), we introduce the function

(B.1) 
$$F_{\epsilon}(x) := \left( \operatorname{Op}(\tilde{B}_{\epsilon}(x,\xi)) \chi(z) \mathcal{R}^* g_{\epsilon} \right) (x).$$

Throughout this section, we assume that  $x \in \mathcal{U}_0$ .

B.1. Replacing  $\tilde{B}_{\epsilon}$  with a simpler amplitude in (B.1). In view of (B.1), (5.2), (5.5), introduce the symbols

(B.2) 
$$\tilde{B}'_0(x,\xi) := \frac{|\xi|}{\tilde{Q}_0(x,\xi) + \kappa \epsilon^3 |\xi|^3}, \ \Delta \tilde{B}_{\epsilon}(x,\xi) := \tilde{B}_{\epsilon}(x,\xi) - \tilde{B}'_0(x,\xi).$$

As is easily seen from (5.5) and the parametrix construction [33, Appendix to Section I.4],  $\Delta \tilde{B}_{\epsilon} \in S^0(\mathcal{U})$  with seminorms uniformly bounded for  $0 < \epsilon \ll 1$ .

Using (3.7), (3.9), (5.1) and (B.1), define

$$J_{\epsilon}(x; y_{j}) := \frac{\epsilon \mu}{(2\pi)^{2}} \int_{\mathbb{R}^{2}} \int_{\mathcal{U}} \tilde{B}_{\epsilon}(x, \xi) \chi(z) e^{i\xi \cdot (z - x)}$$

$$\times \int_{\mathbb{R}} W(z, (\alpha, p)) \varphi_{p} \left(\frac{p - p_{j_{2}}}{\epsilon}\right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} dz d\xi,$$

$$p = \Phi(z, \alpha), \ \alpha = \alpha_{j_{1}} + \epsilon \mu \check{\alpha}.$$

The factor  $\epsilon \mu$  appears due to the change of variables  $\alpha \to \check{\alpha} = (\alpha - \alpha_{j_1})/(\epsilon \mu)$ . Similarly, define

(B.4) 
$$\Delta J_{\epsilon}(x; y_{j}) := \frac{\epsilon \mu}{(2\pi)^{2}} \int_{\mathbb{R}^{2}} \int_{\mathcal{U}} \Delta \tilde{B}_{\epsilon}(x, \xi) \chi(z) e^{i\xi \cdot (z - x)} \times \int_{\mathbb{R}} W(z, (\alpha, p)) \varphi_{p} \left(\frac{p - p_{j_{2}}}{\epsilon}\right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} dz d\xi.$$

**Lemma B.1.** Under the assumptions of Theorem 4.2, one has  $|\Delta J_{\epsilon}(x; y_j)| \le c\epsilon$ ,  $x \in \mathcal{U}_0$ ,  $y_j \in \mathcal{V}$ ,  $0 < \epsilon \ll 1$ .

*Proof.* Our goal is to obtain a bound for  $\Delta J_{\epsilon}(x; y_j)$  as  $\epsilon \to 0$ . To this end we first omit the integral with respect to  $\check{\alpha}$  in (B.4) and drop the prefactor:

$$\Delta H_{\epsilon}(x; \check{\alpha}, y_{j}) := \int_{\mathcal{U}} \int_{\mathbb{R}^{2}} \Delta \tilde{B}_{\epsilon}(x, \xi) e^{i\xi \cdot (z - x)} d\xi$$

$$\times W_{1}(z, \check{\alpha}) \varphi_{p} \left( \frac{\Phi(z, \alpha) - p_{j_{2}}}{\epsilon} \right) dz,$$

$$W_{1}(z, \check{\alpha}) := \chi(z) W(z, (\alpha, p)).$$

The integral with respect to  $\xi$  is not absolutely convergent and understood in the sense of distributions. In (B.4) and (B.5), p and  $\alpha$  are the same as in (B.3). For convenience, in the rest of the proof we drop the dependence of various quantities on  $\check{\alpha}$  and  $y_j$ . By construction,  $\alpha - \alpha_{j_1} = O(\epsilon)$  in (B.5). In simplified notation, (B.5) becomes

(B.6) 
$$\Delta H_{\epsilon}(x) = \int_{\mathcal{U}} \int_{\mathbb{R}^2} \Delta \tilde{B}_{\epsilon}(x,\xi) e^{i\xi \cdot (z-x)} d\xi W_1(z) \varphi_p \left(\frac{\Phi(z) - p_{j_2}}{\epsilon}\right) dz.$$

Pick any point  $z_0 \in \mathcal{S}_{y_j}$ , i.e. so that  $\Phi(z_0, \alpha_{j_1}) = p_{j_2}$ . Using that  $\varphi_p$  is compactly supported and appealing to a partition of unity argument, we can assume without loss of generality that supp  $W_1 \subset \mathcal{U}_1$ , where  $\mathcal{U}_1$  is a small neighborhood of  $z_0$ .

Pick any  $x \in \mathcal{U}_0$  such that  $\operatorname{dist}(x,\mathcal{U}_1) \geq c_0 > 0$  for a sufficiently small  $c_0$ . Therefore, |x-z|,  $z \in \mathcal{U}_1$ , is bounded away from zero. Integrate by parts with respect to  $\xi$  multiple times and use that  $\Delta \tilde{B}_{\epsilon} \in S^0(\mathcal{U})$  to obtain an integral with respect to  $\xi$  that converges absolutely and uniformly in x. This implies that  $|\Delta H_{\epsilon}(x)| \leq c$  for all  $0 < \epsilon \ll 1$ . The constant c can be selected independently of  $x \in \mathcal{U}_0$  provided that  $\operatorname{dist}(x,\mathcal{U}_1) \geq c_0$ .

Suppose now that  $\operatorname{dist}(x, \mathcal{U}_1) \leq c_0$ . By decreasing  $c_0 > 0$  and shrinking  $\mathcal{U}_1$ , if necessary, we can assume that x is sufficiently close to  $z_0$ . Using that  $\operatorname{d}_x \Phi(z) \neq 0$ ,  $z \in \mathcal{U}_1$ , we can change variables  $z = \psi(w)$  so that  $\Phi(z) - p_{j_2} \equiv w_1$  and  $\operatorname{det}(\partial z/\partial w)$  is bounded. Using [33, Theorem 3.3, Chapter I] and [12,

Theorem 1.5], we can transform the expression for  $\Delta H_{\epsilon}$  to the form

(B.7) 
$$\Delta H_{\epsilon}(x) = \int_{\Xi} \int_{\mathbb{R}^2} \Delta \tilde{B}_{\epsilon}^{(1)}(v, w, \eta) e^{i\eta \cdot (w - v)} d\eta W_1(\psi(w)) \varphi_p(w_1/\epsilon) dw$$

with some  $\Delta \tilde{B}_{\epsilon}^{(1)} \in S^0(\Xi)$ , where  $\Xi = \psi^{-1}(\mathcal{U}_1)$  and  $v = \psi^{-1}(x)$ . The absolute value of the Jacobian  $|\det(\partial z/\partial w)|$  is included in  $\Delta \tilde{B}_{\epsilon}^{(1)}$ . Changing variables gives

(B.8) 
$$\Delta H_{\epsilon}(x) = \int_{\mathbb{R}^{2}} \left[ \int_{\mathbb{R}^{2}} \Delta \tilde{B}_{\epsilon}^{(1)} \left( v, w, (\hat{\eta}_{1}/\epsilon, \eta_{2}) \right) W_{1} \left( \psi(\epsilon \check{w}_{1}, w_{2}) \right) \right. \\ \left. \times e^{i\eta_{2}(w_{2}-v_{2})} \mathrm{d}w_{2} \, \varphi_{p}(\check{w}_{1}) e^{i\hat{\eta}_{1}(\check{v}_{1}-\check{w}_{1})} \mathrm{d}\check{w}_{1} \right] \mathrm{d}\hat{\eta}_{1} \mathrm{d}\eta_{2}.$$

Using that  $|\Delta \tilde{B}_{\epsilon}^{(1)}(v, w, \eta)| \leq c$ ,  $v, w \in \Xi$ ,  $\eta \in \mathbb{R}^2$ , and that  $W_1$ ,  $\varphi_p$  are sufficiently smooth and compactly supported, we can show that the integral in brackets decays sufficiently fast as  $\hat{\eta}_1, \eta_2 \to \infty$ .

Specifically, the symbol  $\Delta \tilde{B}_{\epsilon}^{(1)}(v, w, (\hat{\eta}_1/\epsilon, \eta_2))$  and  $W_1$  are smooth in  $\check{w}_1$  and  $w_2$ . Hence, the expression in brackets is absolutely integrable with respect to  $\eta_2$  uniformly with respect to  $x \in \mathcal{U}_0$  and  $\hat{\eta}_1 \in \mathbb{R}$ . To prove decay, it therefore suffices to require that  $\tilde{\varphi}_p \in L^1(\mathbb{R})$ , a property that follows directly from Assumption 3.5(1). Indeed, since  $\varphi'_p \in L^\infty(\mathbb{R})$  and  $\varphi_p$  is compactly supported, it follows that  $\tilde{\varphi}_p \in L^1(\mathbb{R})$ .

For convenience, the domain of integration with respect to  $(\check{w}_1, w_2)$  is extended to all of  $\mathbb{R}^2$ . This does not change the value of the integral due to the assumption supp  $W_1 \subset \mathcal{U}_1$ . The absolute convergence of the integral with respect to  $(\hat{\eta}_1, \eta_2)$  implies that  $|\Delta H_{\epsilon}(x)| \leq c$  for all  $x \in \mathcal{U}_0$  close to  $z_0$  (if they exist) and  $0 < \epsilon \ll 1$ .

The above argument proves that  $|\Delta H_{\epsilon}(x)| \leq c$  for all  $x \in \mathcal{U}_0$  and  $0 < \epsilon \ll 1$ . Integrating with respect to  $\check{\alpha}$  in (B.4) and taking into account the factor  $\epsilon$  proves the lemma.

B.2. Simplification of the leading order term. In view of Lemma B.1, replace  $\tilde{B}_{\epsilon}$  with  $\tilde{B}'_{0}$  in (B.3) and change variables to obtain

$$J_{\epsilon}(x; y_{j}) := \frac{\epsilon^{-2} \mu}{(2\pi)^{2}} \int_{\mathcal{U}} \int_{\mathbb{R}^{2}} \tilde{B}_{0}(x, \hat{\xi}) \chi(z) e^{i\hat{\xi} \cdot (z-x)/\epsilon} d\hat{\xi}$$

$$\times \int_{\mathbb{R}} W(z, (\alpha, p)) \varphi_{p} \left(\frac{\Phi(z, \alpha) - p_{j_{2}}}{\epsilon}\right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} dz,$$

$$p = \Phi(z, \alpha), \ \alpha = \alpha_{j_{1}} + \epsilon \mu \check{\alpha},$$

where we have redefined the meaning of  $J_{\epsilon}$ . Recall that  $\tilde{B}_0$  is defined in (5.8). As is easily checked,  $\epsilon \tilde{B}'_0(x,\hat{\xi}/\epsilon) = \tilde{B}_0(x,\hat{\xi})$ .

Introduce the kernel

(B.10) 
$$B_0(x, \check{w}) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \tilde{B}_0(x, \hat{\xi}) e^{-i\hat{\xi} \cdot \check{w}} d\hat{\xi}.$$

Later we will use the following result:

(B.11) 
$$B_0(x, \check{w}) = \begin{cases} O(-\ln|\check{w}|), & |\check{w}| \to 0, \\ O(|\check{w}|^{-3}), & |\check{w}| \to \infty. \end{cases}$$

The top case follows from [1, Theorem 5.12] because  $\tilde{B}_0 \in S^{-2}(\mathcal{U})$ . The bottom case is proven completely analogously to [23, Lemma 8.1].

Using  $B_0$ , represent  $J_{\epsilon}$  in the form

$$J_{\epsilon}(x; y_{j}) = \epsilon^{-2} \mu \int_{\mathcal{U}} B_{0}\left(x, \frac{x-z}{\epsilon}\right) \chi(z)$$

$$\times \int_{\mathbb{R}} W(z, (\alpha, p)) \varphi_{p}\left(\frac{\Phi(z, \alpha) - p_{j_{2}}}{\epsilon}\right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} dz,$$

$$p = \Phi(z, \alpha), \ \alpha = \alpha_{j_{1}} + \epsilon \mu \check{\alpha}.$$

Change variables  $z \to \check{w} = (x-z)/\epsilon$ :

(B.13)

$$J_{\epsilon}(x; y_{j}) = \mu \int_{\epsilon^{-1}(x-\mathcal{U})} B_{0}(x, \check{w}) \chi(x - \epsilon \check{w})$$

$$\times \int_{\mathbb{R}} W(x - \epsilon \check{w}, (\alpha, p)) \varphi_{p} \left( \frac{\Phi(x - \epsilon \check{w}, \alpha) - p_{j_{2}}}{\epsilon} \right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} d\check{w},$$

$$p = \Phi(x - \epsilon \check{w}, \alpha), \ \alpha = \alpha_{j_{1}} + \epsilon \mu \check{\alpha}.$$

Using (B.11) and that the domain of integration with respect to  $\check{w}$  is contained in a ball  $|\check{w}| \leq c/\epsilon$  gives

$$(B.14)$$

$$J_{\epsilon}(x; y_{j}) = \mu W(x, (\alpha_{j_{1}}, \Phi(x, \alpha_{j_{1}})))\chi(x)$$

$$\times \int_{\mathbb{R}^{2}} B_{0}(x, \check{w})\vartheta\left(\frac{\Phi(x, \alpha_{j_{1}}) - p_{j_{2}}}{\epsilon} - d_{x}\Phi(x, \alpha_{j_{1}})\check{w}\right)d\check{w} + O(\epsilon \ln(1/\epsilon)).$$

To obtain (B.14) from (B.13), we first incorporate the function  $\vartheta$ , introduced in (5.10), by using that

(B.15)
$$\frac{\Phi(x - \epsilon \check{w}, \alpha) - p_{j_2}}{\epsilon}$$

$$= \frac{\Phi(x - \epsilon \check{w}, \alpha_{j_1}) - p_{j_2}}{\epsilon} + \mu \Phi'_{\alpha}(x - \epsilon \check{w}, \alpha_{j_1}) \check{\alpha} + O(\epsilon)$$

$$= \left[\frac{\Phi(x, \alpha_{j_1}) - p_{j_2}}{\epsilon} - d_x \Phi(x, \alpha_{j_1}) \check{w} + O(\epsilon(1 + |\check{w}|))\right] + \mu \Phi'_{\alpha}(x, \alpha_{j_1}) \check{\alpha}.$$

Ignoring the big-O term in the last line in (B.15) leads to an overall error of  $O(\epsilon \ln(1/\epsilon))$  in (B.14). This follows from the properties  $\varphi_p' \in L^{\infty}(\mathbb{R})$  (see Assumption 3.5(1)),  $B_0(x, \check{w}) = O(|\check{w}|^{-3})$ ,  $|\check{w}| \to \infty$ , and the fact that the domain of integration in (B.13) is contained in a ball  $|\check{w}| \le c/\epsilon$ .

Then we extend the domain of integration from  $e^{-1}(x-\mathcal{U})$  to  $\mathbb{R}^2$ , which contributes a term of magnitude  $O(\epsilon)$ . This follows from the fact that the set  $x-\mathcal{U}$  is a neighborhood of zero,  $\vartheta$  is compactly supported in its first argument,  $|\mathrm{d}_x\Phi(x,\alpha_{j_1})| \geq c$ , and  $B_0(x,\check{w}) = O(|\check{w}|^{-3})$ ,  $|\check{w}| \to \infty$ . To simplify notation, the arguments  $x,\alpha_{j_1}$  of  $\vartheta$  are omitted.

Next, suppose  $x = x_0 + \epsilon \check{x}$ . It is easy to show that the partial derivatives  $\partial_{x_j} B_0(x, \check{w}), j = 1, 2$ , also satisfy (B.11). This can be established using the

same method of proof as for  $B_0$ . Using also that  $\chi(x) \equiv 1$  on  $\mathcal{U}_0$  we obtain

$$(B.16) J_{\epsilon}(x_0 + \epsilon \check{x}; y_j) = \mu W \left( x_0, (\alpha_{j_1}, \Phi(x_0, \alpha_{j_1})) \right)$$

$$\times \int_{\mathbb{R}^2} B_0(x_0, \check{w}) \vartheta \left( \frac{\Phi(x_0 + \epsilon \check{x}, \alpha_{j_1}) - p_{j_2}}{\epsilon} - d_x \Phi(x_0, \alpha_{j_1}) \check{w} \right) d\check{w}$$

$$+ O(\epsilon \ln(1/\epsilon)).$$

The integral with respect to  $\check{w}$  can be simplified as follows

(B.17) 
$$\int_{\mathbb{R}^2} B_0(x_0, \check{w}) \vartheta (\check{q} - d_x \Phi(x_0, \alpha_{j_1}) \check{w}) d\check{w}$$

$$= \int K_0(\check{t}; x_0, \alpha_{j_1}) \vartheta (\check{q} - \check{t}) d\check{t}, \ \check{q} = \frac{\Phi(x_0 + \epsilon \check{x}, \alpha_{j_1}) - p_{j_2}}{\epsilon},$$

where

(B.18) 
$$K_0(\check{t}; x_0, \alpha) := \int_{\mathbb{R}} B_0(x_0, \check{w}_1 e_1 + \check{w}_2 e_2) d\check{w}_2.$$

Here  $e_1 = |\mathrm{d}_x \Phi(x_0, \alpha)|^{-1} \mathrm{d}_x \Phi(x_0, \alpha)$ ,  $e_2$  is a unit vector orthogonal to  $e_1$ , and  $\check{w} = (\check{w}_1, \check{w}_2)$ . Using (B.10) shows that  $K_0$  satisfies (5.9). Clearly,

(B.19) 
$$\frac{\Phi(x_0 + \epsilon \check{x}, \alpha_{j_1}) - p_{j_2}}{\epsilon} = \frac{\Phi(x_0, \alpha_{j_1}) - p_{j_2}}{\epsilon} + d_x \Phi(x_0, \alpha_{j_1}) \check{x} + O(\epsilon).$$

We therefore approximate  $\check{q}$  in (B.17) and (B.16) by the right-hand side of the last equation with the term  $O(\epsilon)$  neglected. Arguing as above and using that  $\vartheta'_{\check{t}} \in L^{\infty}(\mathbb{R})$  in its first argument, we conclude that this approximation introduces an error of magnitude  $O(\epsilon)$  in (B.16).

B.3. End of proof. Consider now the second and third terms on the right in (5.4).

The operator  $\operatorname{Op}(\tilde{B}_{\epsilon}(x,\xi))(1-\chi(z), x \in \mathcal{U}_0$ , is smoothing. Arguing analogously to the proof of Lemma B.1 we establish that

(B.20) 
$$\left(\operatorname{Op}\left(\tilde{B}_{\epsilon}(x,\xi)\right)(1-\chi(z))\mathcal{R}^*\varphi_{\epsilon}(\cdot-y_j)=O(\epsilon),\ x\in\mathcal{U}_0,\ 0<\epsilon\ll 1.\right)$$

Clearly,  $\|\varphi_{\epsilon}(\cdot - y_j)\|_{L^2(\mathcal{V})} = O(\epsilon)$ . Set  $h(y) = \varphi_{\epsilon}(y - y_j)$  in (5.3), and let  $\psi_{\epsilon,y_j} \in H^1_0(\mathcal{U}_b)$  be the corresponding solution to (5.2). The inequality (5.3) implies  $\|\psi_{\epsilon,y_j}\|_{H^{-1/2}(\mathcal{U})} \leq c\epsilon$ . By the linearity of (5.2),  $N_{\epsilon}^{\text{rec}}(x) = \sum_j \psi_{\epsilon,y_j}(x)\eta_j$ . Applying  $\mathcal{T}_{\epsilon}$  on both sides gives

(B.21) 
$$(\mathcal{T}_{\epsilon} N_{\epsilon}^{\text{rec}})(x) = \sum_{j \in J} A_{\epsilon, y_j}(x) \eta_j, \ x \in \mathcal{U}_0,$$

for some  $A_{\epsilon,y}(x) \in C^{\infty}(\mathcal{U}_0)$  which satisfy  $|A_{\epsilon,y}(x)| \leq c\epsilon$  uniformly in  $x \in K \subset \mathcal{U}_0$ , where K is any compact set,  $y \in \mathcal{V}$ , and  $0 < \epsilon \ll 1$ . This follows because  $\mathcal{T}_{\epsilon} : H^{-1/2}_{comp}(\mathcal{U}) \to H^s_{loc}(\mathcal{U}_0)$  is bounded for any  $s \in \mathbb{R}$ . Selecting any s > 1 and using the bound  $||f||_{L^{\infty}(\mathbb{R}^2)} \leq c||f||_{H^s(\mathbb{R}^2)}$  [32, Chapter 4, Proposition 1.3] gives the desired result.

Combining all the results, and using (3.9) along with the fact that  $p_{j_2} = \epsilon j_2$ , we finish the proof.

#### Appendix C. Proof of Lemma 6.1

We begin with a simple lemma, which is needed in the proof.

Lemma C.1. One has

(C.1) 
$$|K_0(\check{q}; x, \alpha)| \le c(1 + |\check{q}|)^{-2}, \ \check{q} \in \mathbb{R},$$

and

(C.2) 
$$\left| \partial_{\check{q}} \int K_0(\check{q} - \check{t}; x, \alpha) \vartheta(\check{t}; x, \alpha) d\check{t} \right| \le c(1 + |\check{q}|)^{-3},$$

for all  $x \in \mathcal{U}$  and  $\alpha \in \mathcal{I}_{\alpha}$ .

The bound (C.1) follows from (5.9) by integration by parts and the fact that  $\tilde{B}_0(x,\hat{\xi}) = O(|\hat{\xi}|)$  as  $\hat{\xi} \to 0$ , see (5.8). The integral is absolutely convergent and  $K_0$  is bounded because  $\tilde{B}_0(x,\lambda d_x \Phi(x,\alpha)) = O(\lambda^{-2})$  as  $\lambda \to \infty$ .

The bound (C.2) likewise follows from integration by parts. The decay at infinity is faster due to the extra factor  $\lambda$  in the integrand, which arises in the Fourier domain from the derivative  $\partial_{\tilde{q}}$ . The derivative is bounded due to the convolution with  $\vartheta$ , which satisfies  $\vartheta'_{\tilde{t}} \in L^{\infty}(\mathbb{R})$  and is compactly supported in the first argument.

In this section, we omit  $x_0$  from the arguments of G for simplicity. By (C.1) and the fact that  $\vartheta$  is compactly supported in  $\check{t}$ , we have

(C.3) 
$$|G(\alpha, \check{q})| \le c(1+|\check{q}|)^{-2}, \ \alpha \in \mathcal{I}_{\alpha}, \ \check{q} \in \mathbb{R}.$$

The assertion (6.2) follows from (3.6), (5.12), and (C.3).

C.1. **Proof of** (6.3) **in the case** L = 1**.** Begin by setting L = 1 and  $\theta_l = 1$  in (6.1). Denote:

(C.4) 
$$D_{\epsilon} := \sum_{j \in J} \operatorname{Var}(\zeta_{\epsilon,j}),$$

i.e.  $D_{\epsilon}$  is the expression in brackets in the denominator in (6.3). By (3.6),

(C.5) 
$$\sum_{j_{2} \in J_{p}} \left| (G(\alpha, r - j_{2}) + R_{j})^{2} - G^{2}(\alpha, r - j_{2}) \right| \mathbb{E}\eta_{j}^{2}$$
$$\leq c\epsilon \sum_{j_{2} \in J_{p}} \left[ R_{j}^{2} + |G(\alpha, r - j_{2})R_{j}| \right].$$

Using (C.3) and (5.12), straightforward calculations show that

(C.6) 
$$\sum_{j_2 \in J_p} R_j^2 \le c\epsilon \ln^2(1/\epsilon), \sum_{j_2 \in J_p} |G(\alpha, r - j_2)R_j| \le c\epsilon \ln(1/\epsilon).$$

Also, by (3.6), (5.12), and (6.1),

(C.7) 
$$\mathbb{E}\zeta_{\epsilon,j} = \left[G(\alpha_{j_1}, r_{j_1} - j_2) + O(\epsilon \ln(1/\epsilon))\right]O(\epsilon^{3/2}).$$

Therefore,

(C.8)  

$$\sum_{j \in J} (\mathbb{E}\zeta_{\epsilon,j})^{2}$$

$$\leq c\epsilon^{3} \sum_{j \in J} \left[ G^{2}(\alpha_{j_{1}}, r_{j_{1}} - j_{2}) + \epsilon \ln(1/\epsilon) \left| G(\alpha_{j_{1}}, r_{j_{1}} - j_{2}) \right| + \epsilon^{2} \ln^{2}(1/\epsilon) \right]$$

$$\leq c\epsilon^{3} \sum_{j_{1} \in J_{\epsilon}} \left[ 1 + \epsilon \ln(1/\epsilon) + \epsilon \ln^{2}(1/\epsilon) \right] = O(\epsilon^{2}).$$

Hence, combining (C.5)–(C.8) yields

(C.9)  

$$D_{\epsilon} = \sum_{j \in J} \mathbb{E}\zeta_{\epsilon,j}^{2} - \sum_{j \in J} (\mathbb{E}\zeta_{\epsilon,j})^{2}$$

$$= \Delta \alpha \sum_{j_{1} \in J_{\alpha}} \left[ \sum_{j_{2} \in J_{p}} G^{2}(\alpha_{j_{1}}, r_{j_{1}} - j_{2}) \sigma^{2}(\alpha_{j_{1}}, p_{j_{2}}) + O(\epsilon \ln^{2}(1/\epsilon)) \right]$$

$$+ O(\epsilon^{2})$$

$$= \Delta \alpha \sum_{j_{1} \in J_{\alpha}} \sum_{\substack{j_{2} \in J_{p} \\ |r_{j_{1}} - j_{2}| \leq \epsilon^{-1/3}}} G^{2}(\alpha_{j_{1}}, r_{j_{1}} - j_{2}) \sigma^{2}(\alpha_{j_{1}}, p_{j_{2}}) + O(\epsilon \ln^{2}(1/\epsilon)).$$

Here we have used that  $\sigma$  is bounded and, by (C.3), limiting the sum with respect to  $j_2$  to the set  $|r_{j_1}-j_2| \leq \epsilon^{-1/3}$  leads to an overall error of magnitude  $O(\epsilon)$  in (C.9).

Define

(C.10) 
$$\psi(\alpha, r) := \sum_{j_2 \in \mathbb{Z}} G^2(\alpha, r - j_2), \ \alpha \in \mathcal{I}_{\alpha}, r \in \mathbb{R}.$$

By (C.3) the series is absolutely convergent. It is easy to see that

(C.11) 
$$\psi(\alpha, r + m) = \psi(\alpha, r), \ \alpha \in \mathcal{I}_{\alpha}, r \in \mathbb{R}, m \in \mathbb{Z}.$$

Replace  $p_{j_2}$  with  $\Phi(x_0, \alpha_{j_1})$  in the arguments of  $\sigma^2$  in (C.9) and extend the sum with respect to  $j_2$  to all of  $\mathbb{Z}$  to obtain

(C.12) 
$$D_{\epsilon} = \Delta \alpha \sum_{j_1 \in J_{\alpha}} \psi(\alpha_{j_1}, r_{j_1}) \sigma^2(\alpha_{j_1}, \Phi(x_0, \alpha_{j_1})) + O(\epsilon \ln^2(1/\epsilon)).$$

Indeed, consider the sum with respect to  $j_2$ . By the definition of  $r_{j_1}$  in (5.12) we have  $p_{j_2} - \Phi(x_0, \alpha_{j_1}) = \epsilon(j_2 - r_{j_1} + O(1))$ . By (C.3) and the Lipschitz continuity of  $\sigma$  in a neighborhood of the set  $\Gamma_{x_0}$  (Assumption 3.7(2)):

continuity of 
$$\sigma$$
 in a neighborhood of the set  $\Gamma_{x_0}$  (Assumption 3.7(2))
$$\sum_{\substack{j_2 \in J_p \\ |r-j_2| \le \epsilon^{-1/3}}} G^2(\alpha, r-j_2) |\sigma^2(\alpha, \epsilon j_2) - \sigma^2(y, \Phi(x_0, \alpha_{j_1}))|$$
(C.13)
$$\stackrel{|r-j_2| \le \epsilon^{-1/3}}{\le c} \le c \sum_{|r-j_2| \le \epsilon^{-1/3}} \epsilon (1 + |r-j_2|) (1 + |r-j_2|)^{-4} = O(\epsilon).$$

Here we use the fact that all  $y_j$  satisfying  $|r_{j_1} - j_2| \le \epsilon^{-1/3}$  lie in a neighborhood of  $\Gamma_{x_0}$  for all  $0 < \epsilon \ll 1$ .

Again, by (C.3), we have

(C.14) 
$$\sum_{j_2 \in \mathbb{Z} \setminus J_p} G^2(\alpha, r - j_2) \le c \sum_{|r - j_2| \ge 1/\epsilon} G^2(\alpha, r - j_2) = O(\epsilon^{-3}),$$

where  $r = r_{j_1}$ . In the first inequality, we used that Assumption 3.1(3) implies the existence of  $\delta > 0$  such that, for each  $(x_0, \alpha_{j_1}) \in \mathcal{U}_0 \times \mathcal{I}_{\alpha}$ , all  $p_{j_2}$  satisfying  $|\Phi(x_0, \alpha_{j_1}) - p_{j_2}| \leq \delta$  are in the data, i.e. they belong to  $\mathcal{I}_p$ .

Summing over  $\alpha_{j_1} \in \mathcal{I}_{\alpha}$  and accounting for the factor  $\Delta \alpha$  in front of the sum in (C.9) proves (C.12).

Using Assumption 3.7 and arguments very similar to [21, Appendix D], we obtain the following result

(C.15) 
$$\lim_{\epsilon \to 0} D_{\epsilon} = \int_{\mathcal{I}_{\alpha}} \left( \int_{0}^{1} \psi(\alpha, r) dr \right) \sigma^{2}(\alpha, \Phi(x_{0}, \alpha)) d\alpha.$$

Indeed, the argument in [21, Appendix D] is based on Assumption 2.8 in that paper, which consists of three items. Our Assumption 3.7(1) ensures that Assumption 2.8(1) holds, and Assumption 3.7(3) guarantees Assumption 2.8(2). Finally, Assumption 2.8(3) requires that the equation  $d_x\Phi(x_0,\alpha) \cdot \check{x} = 0$  for  $\alpha \in \mathcal{I}_{\alpha}$  has a solution set of Lebesgue measure zero for each  $\check{x} \in \mathbb{R}^2 \setminus 0_2$ . This follows from Assumption 3.2(1). Indeed, if  $d_x\Phi(x_0,\alpha) \cdot \check{x} = 0$  for some  $\alpha \in \mathcal{I}_{\alpha}$ , then  $\partial_{\alpha}(d_x\Phi(x_0,\alpha) \cdot \check{x}) \neq 0$ , since otherwise the vectors  $d_x\Phi(x_0,\alpha)$  and  $d_x\Phi'_{\alpha}(x_0,\alpha)$  would both be orthogonal to  $\check{x}$  and hence linearly dependent. This contradiction shows that solutions to  $d_x\Phi(x_0,\alpha) \cdot \check{x} = 0$  are isolated points.

Furthermore,

(C.16) 
$$\int_0^1 \psi(\alpha, r) dr = \sum_{j_2 \in \mathbb{Z}} \int_0^1 G^2(\alpha, r - j_2) dr = \int_{\mathbb{R}} G^2(\alpha, r) dr =: H(\alpha).$$

By Parseval's theorem and (5.9), (5.11),

(C.17) 
$$\int_{\mathbb{R}} G^{2}(\alpha, r) dr$$

$$= \frac{\mu^{2} W^{2}(x_{0}, \Phi(x_{0}, \alpha))}{2\pi} \int_{\mathbb{R}} |\tilde{B}_{0}(x_{0}, \lambda d_{x} \Phi(x_{0}, \alpha)) \tilde{\vartheta}(\lambda; x_{0}, \alpha)|^{2} d\lambda.$$

Combined with (C.15)–(C.17), Assumption 3.7(2) yields

(C.18) 
$$\lim_{\epsilon \to 0} D_{\epsilon} = \int_{\mathcal{I}_{\alpha}} H(x_0, \alpha) \sigma^2(\alpha, \Phi(x_0, \alpha)) d\alpha > 0.$$

Consider now the numerator in (6.3) (with L=1). It is bounded by

(C.19) 
$$c \sum_{j \in J} |G(\alpha_{j_1}, r_{j_1} - j_2) + R_j|^3 \mathbb{E} |\eta_j - \bar{\eta}_j|^3.$$

Using the last result in (3.6) and the estimates for G and  $R_j$  we get that the numerator is  $O(\epsilon^{1/2})$ . Together with (C.18) this proves Lemma 6.1 in the case L=1.

C.2. **Proof of** (6.3) **in the case**  $L \ge 1$ **.** The proof of (6.3) in the general case  $L \ge 1$  is similar to the above, so we only highlight key new points. Instead of (C.12), we now have

(C.20) 
$$D_{\epsilon} = \Delta \alpha \sum_{j \in J} \left[ \sum_{l=1}^{L} \theta_l \left( G(\alpha_{j_1}, r_{j_1}^{(l)} - j_2) + R_j^{(l)} \right) \right]^2 \sigma^2(\alpha_{j_1}, \Phi(x_0, \alpha_{j_1})) + O(\epsilon \ln^2(1/\epsilon)).$$

See (6.1) for the definitions of  $r_{j_1}^{(l)}$  and  $R_j^{(l)}$ . Here we have used the same substitution  $p_{j_2} \cong \Phi(x_0, \alpha_{j_1})$  (cf. (C.13)) and (3.6). Estimating the contribution of the remainder  $R_j^{(l)}$  similarly to (C.6) and passing to the limit as in (C.15) gives

(C.21) 
$$\lim_{\epsilon \to 0} D_{\epsilon} = \int_{\mathcal{I}_{\alpha}} \int_{\mathbb{R}} f^{2}(\alpha, r) dr \, \sigma^{2}(\alpha, \Phi(x_{0}, \alpha)) d\alpha,$$
$$f(\alpha, r) := \sum_{l=1}^{L} \theta_{l} G(\alpha, r + d_{x} \Phi(x_{0}, \alpha) \cdot \check{x}_{l}).$$

Multiplying out the sum in the definition of f yields

(C.22) 
$$\lim_{\epsilon \to 0} D_{\epsilon} = \sum_{l_1=1}^{L} \sum_{l_2=1}^{L} \theta_{l_1} \theta_{l_2} C(\check{x}_{l_1} - \check{x}_{l_2}),$$
$$C(\check{w}) := \int_{\mathcal{I}_{\alpha}} (G \star G) (\alpha, d_x \Phi(x_0, \alpha) \cdot \check{w}) \sigma^2(\alpha, \Phi(x_0, \alpha)) d\alpha.$$

Observe that if L = 1,  $\theta_1 = 1$ , and  $w = 0_2$ , then (C.21) coincides with (C.15), (C.16) after a change of variables  $r + d_x \Phi(x_0, \alpha) \cdot \check{x}_1 \to r$ . By (5.11),

(C.23) 
$$f(\alpha, r) = \frac{\mu W(x_0, \Phi(x_0, \alpha))}{2\pi} \int_{\mathbb{R}} \tilde{B}_0(x_0, \lambda d_x \Phi(x_0, \alpha)) \tilde{\vartheta}(\lambda; x_0, \alpha) \times \left[ \sum_{l=1}^L \theta_l e^{-i\lambda d_x \Phi(x_0, \alpha) \cdot \check{x}_l} \right] e^{-i\lambda r} d\lambda.$$

Suppose the limit in (C.21) equals zero. By Assumption 3.7(2), this implies that  $f(\alpha, r) \equiv 0$  for all  $r \in \mathbb{R}$  and  $\alpha \in I_{\alpha}$ . The set  $I_{\alpha}$  is introduced in Assumption 3.7(2). Furthermore,  $\vartheta$  is compactly supported in  $\check{t}$ , so  $\check{\vartheta}$  is analytic in  $\lambda$  and cannot be zero on an open set. Additionally,  $W(x_0, \Phi(x_0, \alpha)) > 0$  and  $\tilde{B}_0(x_0, \lambda d_x \Phi(x_0, \alpha)) \neq 0$  (see Assumption 3.2(4) and (5.8)). Therefore

(C.24) 
$$\sum_{l=1}^{L} \theta_l e^{-i\lambda d_x \Phi(x_0, \alpha) \cdot \check{x}_l} \equiv 0, \ \lambda \in \mathbb{R}, \ \alpha \in I_{\alpha}.$$

Arguing very similarly to [21, Appendix E], we conclude that all  $\theta_l$  are zero. For convenience, this argument is presented in Section C.3.

Finally, the argument to show that the numerator in (6.3) is  $O(\epsilon^{1/2})$  is very similar to (C.19).

C.3. **Proof of the fact that**  $\theta_l = 0$ ,  $1 \leq l \leq L$ . Consider the vectors  $e_k := \check{x}_{l_1} - \check{x}_{l_2} \in \mathbb{R}^2$ ,  $1 \leq l_1, l_2 \leq L$ ,  $l_1 \neq l_2$ , and the sets  $I_{\alpha,k} := \{\alpha \in I_\alpha : d_x \Phi(x_0, \alpha) \cdot e_k = 0\}$ . The  $\check{x}_l$  are distinct, so  $e_k \neq 0$  for each k. By Assumption 3.2(1), if  $\alpha$  satisfies  $d_x \Phi(x_0, \alpha) \cdot e_k = 0$ , then  $\partial_\alpha (d_x \Phi(x_0, \alpha) \cdot e_k) \neq 0$ . Hence each set  $I_{\alpha,k}$  is a collection of isolated points. There are finitely many  $I_{\alpha,k}$ , so their union is not all of  $I_\alpha$ . Therefore there exists  $\alpha_0 \in I_\alpha \setminus (\cup_k I_{\alpha,k})$  such that  $d_x \Phi(x_0, \alpha_0) \cdot e_k \neq 0$  for each k.

We found  $\alpha_0 \in I_\alpha$  such that  $s_l := d_x \Phi(x_0, \alpha_0) \cdot \check{x}_l$ ,  $1 \le l \le L$ , are pairwise distinct. Differentiate (C.24) with respect to  $\lambda$  multiple times and set  $\lambda = 0$ . This gives

(C.25) 
$$\sum_{l=1}^{L} \theta_l s_l^k = 0, \ k = 0, 1, \dots, L-1.$$

Since all  $s_l$  are pairwise distinct, using the properties of the Vandermond determinant we conclude from (C.25) that all  $\theta_l = 0$ . This contradiction proves that the limit in (C.21) is not zero unless  $\vec{\theta} = 0$ .

#### Appendix D. Proof of Lemma 7.3

To prove the lemma we use results from the theory of Sobolev spaces. Consider the sets

(D.1) 
$$\Gamma_{\delta} := \{ f \in C : \|f\|_{W^{2,2}(D)}^2 \le 1/\delta \},$$

where D is a bounded Lipschitz domain. We use the Sobolev space  $W^{k,p}(D)$ , which is the closure of  $C^{\infty}(\bar{D})$  in the norm:

(D.2) 
$$||f||_{k,p} := \left( \int_D \sum_{|m| \le k} |\partial_x f(x)|^p dx \right)^{1/p}, \ m \in \mathbb{N}_0^2, \ f \in C^{\infty}(\bar{D}).$$

By [14, eq. (7.8), p. 146 and Theorem 7.26, p. 171], the imbedding  $W^{2,2}(D)\hookrightarrow C(\bar{D})$  is compact. More precisely, we use here that the imbedding  $W^{2,2}(D)\hookrightarrow W^{2,p}(D)$ ,  $1\leq p\leq 2$  is continuous [14, eq. (7.8), p. 146], and the imbedding  $W^{2,p}(D)\hookrightarrow C(\bar{D})$ ,  $1\leq p<2$ , is compact [14, Theorem 7.26, p. 171]. Hence the set  $\Gamma_{\delta}\subset C$  is compact for every  $\delta>0$ .

The proof parallels that of Lemma 5.2. Begin with  $F_{\epsilon}$ . Similarly to (B.4),

$$\Delta J_{\epsilon}(x_{0} + \epsilon \check{x}; y_{j}) := \frac{\epsilon \mu}{(2\pi)^{2}} \int_{\mathbb{R}^{2}} \int_{\mathcal{U}} \Delta \tilde{B}_{\epsilon}(x_{0} + \epsilon \check{x}, \xi) \chi(z) e^{-i\xi \cdot w}$$

$$\times \int_{\mathbb{R}} W(z, (\alpha, p)) \varphi_{p} \left(\frac{p - p_{j_{2}}}{\epsilon}\right) \varphi_{\alpha}(\check{\alpha}) d\check{\alpha} dw d\xi,$$

$$p = \Phi(z, \alpha), \ \alpha = \alpha_{j_{1}} + \epsilon \mu \check{\alpha}, \ z = x_{0} + \epsilon \check{x} - w.$$

Arguing as in Appendix B.1, we prove that  $\partial_{\tilde{x}}^m \Delta J_{\epsilon}(x_0 + \epsilon \check{x}; y_j) = O(\epsilon)$ ,  $|m| \leq 2$ . In particular, differentiating  $\varphi_p((p-p_{j_2})/\epsilon)$  with respect to  $\check{x}$  does not introduce factors that grow with  $\epsilon$ , since the  $\epsilon$  in the numerator and denominator cancel out.

Based on this argument, the function  $\varphi_p''$  must satisfy the smoothness conditions imposed on  $\varphi_p$  in Appendix B. The most restrictive of these is the requirement  $\varphi_p' \in L^{\infty}(\mathbb{R})$  (see, e.g., the paragraph following (B.14)), which in turn implies that  $\varphi_p^{(3)} \in L^{\infty}(\mathbb{R})$ . The analog of the other condition

from Appendix B, namely  $\tilde{\varphi}_p \in L^1(\mathbb{R})$  (see the paragraph following (B.8)), now becomes  $\tilde{\varphi}_p'' \in L^1(\mathbb{R})$ . As before, this requirement is less restrictive and is automatically satisfied because  $\varphi_p$  is compactly supported and  $\varphi_p^{(3)} \in L^{\infty}(\mathbb{R})$ .

Substituting  $x = x_0 + \epsilon \check{x}$  into (B.13) and differentiating  $J_{\epsilon}(x_0 + \epsilon \check{x}; y_j)$  with respect to  $\check{x}$ , we conclude that the analysis in the remainder of Appendix B.2 applies equally to the derivatives. Note that this differentiation does not introduce boundary terms, despite the domain of integration depending on x, because  $\chi \in C_0^{\infty}(\mathcal{U})$ . Therefore, we can represent  $\partial_{\bar{x}}^m J_{\epsilon}(x_0 + \epsilon \check{x}; y_j)$  in a form analogous to (5.12). The analog of G has the same decay at infinity, i.e.  $O((r_{j_1} - j_2)^{-2})$ . This gives

(D.4) 
$$|\partial_{\check{x}}^m J_{\epsilon}(x_0 + \epsilon \check{x}; y_j)| \le c [(1 + |r_{j_1} - j_2|)^{-2} + \epsilon \ln(1/\epsilon)], |m| \le 2.$$

Recall that  $r_{j_1}$  are defined in (5.12).

The final contribution to  $N_{\epsilon}^{\text{rec}}$  comes from the difference  $N_{\epsilon}^{\text{rec}} - F_{\epsilon}$ , see (5.4) and (B.1). The magnitude of its derivatives with respect to  $\check{x}$  can be estimated similarly to section B.3.

Combining the results proves that

(D.5) 
$$\partial_{\check{x}}^{m} N_{\epsilon}^{\text{rec}}(x_{0} + \epsilon \check{x}) = \sum_{j \in J} A_{m,j}(\check{x}) \eta_{j}, \\ |A_{m,j}(\check{x})| \le c \left[ (1 + |r_{j_{1}} - j_{2}|)^{-2} + \epsilon \ln(1/\epsilon) \right], \ |m| \le 2.$$

Therefore (cf. (C.9))

$$(D.6) \qquad \mathbb{E}(\partial_{\check{x}}^{m} N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}))^2 = c\epsilon \sum_{j \in J} A_{m,j}^2 \sigma^2(y_j) \le c, \ \check{x} \in D, \ |m| \le 2.$$

As opposed to (C.9), here we do not need an asymptotic expression but a simple upper bound. This implies that  $\mathbb{E}||N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}))||_{W^{2,2}(D)}^2 \leq c_0$  for some  $c_0$  and all  $0 < \epsilon \ll 1$ .

Let  $\pi_{\epsilon}(x)$  be the probability distribution of the non-negative random variable  $\xi_{\epsilon} = ||N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x})||_{W^{2,2}(D)}^2$ . By an elementary inequality,

(D.7) 
$$\mathbb{P}(N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}) \notin \Gamma_{\delta}) = \mathbb{P}(\|N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x})\|_{W^{2,2}(D)}^2 \ge 1/\delta)$$
$$= \int_{1/\delta}^{\infty} d\pi_{\epsilon}(x) \le \delta \int_{1/\delta}^{\infty} x d\pi_{\epsilon}(x) \le c_0 \delta.$$

Therefore  $(N_{\epsilon}^{\text{rec}}(x_0 + \epsilon \check{x}))$  is a tight family.

# ACKNOWLEDGMENT

The author acknowledges using ChatGPT for minor language editing (spelling, grammar, and general style).

#### References

- H. Abels, Pseudodifferential and Singular Integral Operators, De Gruyter, Berlin/Boston, 2011.
- [2] A. ABHISHEK, A. KATSEVICH, AND J. W. WEBBER, Local reconstruction analysis of inverting the Radon transform in the plane from noisy discrete data, SIAM Journal on Imaging Sciences, 18 (2025), pp. 936–962.

- [3] ——, Statistical microlocal analysis in two-dimensional X-ray CT, June 2025. arXiv:2506.05113 [math].
- [4] R. J. Adler, The Geometry of Random Fields, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2010.
- [5] K. B. ATHREYA AND S. N. LAHIRI, Measure Theory and Probability Theory, Springer Texts in Statistics, Springer, New York, 2006.
- [6] N. BISSANTZ, H. HOLZMANN, AND K. PROKSCH, Confidence regions for images observed under the Radon transform, Journal of Multivariate Analysis, 128 (2014), pp. 86–107.
- [7] T. Blu, P. Thévenaz, and M. Unser, Complete parameterization of piecewisepolynomial interpolation kernels, IEEE Transactions on Image Processing, 12 (2003), pp. 1297–1309.
- [8] L. CAVALIER, Asymptotically efficient estimation in a problem related to tomography, Mathematical Methods of Statistics, 7 (1998), pp. 445–456.
- [9] ——, Efficient estimation of a density in a problem of tomography, Annals of Statistics, 28 (2000), pp. 630–647.
- [10] S. E. DIVEL AND N. J. PELC, Accurate Image Domain Noise Insertion in CT Images, IEEE Transactions on Medical Imaging, 39 (2020), pp. 1906–1916.
- [11] Y. V. EGOROV AND B.-W. SCHULZE, Pseudo-Differential Operators, Singularities, Applications, Birkhauser, Basel, 1997.
- [12] YU. V. EGOROV AND M. A. SHUBIN, Linear Partial Differential Equations. Elements of the Modern Theory, in Partial Differential Equations II, 1994, pp. 1–120.
- [13] C. L. EPSTEIN, Introduction to the Mathematics of Medical Imaging, SIAM, Philadelphia, 2008.
- [14] D. GILBARG AND N. S. TRUDINGER, Elliptic Partial Differential Equations of Second Order, Classics in Mathematics, Springer, Berlin, Heidelberg, 2001.
- [15] A. GRIGIS AND J. SJÖSTRAND, Microlocal Analysis for Differential Operators, Cambridge University Press, 1994.
- [16] S. Holman, Generalized Radon transforms, in Microlocal Analysis and Inverse Problems in Tomography and Geometry, T. Quinto, P. Stefanov, and G. Uhlmann, eds., Walter de Gruyter GmbH, Berlin/Boston, 2024, pp. 1–34.
- [17] A. C. KAK AND M. SLANEY, Principles of Computerized Tomographic Imaging, Classics in Applied Mathematics, SIAM, 2001.
- [18] A. Katsevich, A local approach to resolution analysis of image reconstruction in tomography, SIAM Journal on Applied Mathematics, 77 (2017), pp. 1706–1732.
- [19] —, Resolution Analysis of Inverting the Generalized N-Dimensional Radon Transform in  $\mathbb{R}^n$  from Discrete Data, Journal of Fourier Analysis and Applications, 29 (2023).
- [20] ——, Resolution of 2 Dimensional Reconstruction of Functions With Nonsmooth Edges From Discrete Radon Transform Data, SIAM Journal on Applied Mathematics, 83 (2023), pp. 695–724.
- [21] ——, Analysis of reconstruction from noisy discrete generalized Radon data, arXiv:2405.13269, (2024).
- [22] ——, Analysis of reconstruction of functions with rough edges from discrete Radon data, Journal of Fourier Analysis and Applications, 31, article 38 (2025).
- [23] ——, Local analysis of iterative reconstruction from discrete generalized Radon transform data in the plane, SIAM Journal on Imaging Sciences, to appear (2025).
- [24] R. G. KEYS, Cubic Convolution Interpolation for Digital Image Processing, IEEE Transactions on Acoustics, Speach, and Signal Processing, ASSP-29 (1981), pp. 1153–1160
- [25] D. Khoshnevisan, Multiparameter Processes: An Introduction to Random Fields, Springer-Verlag, New York, 2002.
- [26] A. KOROSTELEV AND A. TSYBAKOV, Asymptotically minimax image reconstruction problems, in Topics in Nonparametric Estimation, Amer. Math. Soc., Providence, RI, 1992, pp. 45–86.

- [27] A. P. KOROSTELEV AND A. B. TSYBAKOV, Optimal rates of convergence of estimates in the stochastic problem of computerized tomography, Problems of information transmission, 27 (1991), pp. 73–81.
- [28] F. MONARD, R. NICKL, AND G. P. PATERNAIN, Efficient nonparametric Bayesian inference for X-ray transforms, Annals of Statistics, 47 (2019), pp. 1113–1147.
- [29] E. T. QUINTO, The dependence of the generalized Radon transform on defining measures, Transactions of the American Mathematical Society, 257 (1980), pp. 331–346.
- [30] S. SILTANEN, V. KOLEHMAINEN, S. J. RVENP, J. P. KAIPIO, P. KOISTINEN, M. LAS-SAS, J. PIRTTIL, AND E. SOMERSALO, Statistical inversion for medical X-ray tomography with few radiographs: I. General theory, Physics in Medicine and Biology, 48 (2003), pp. 1437–1463.
- [31] P. Stefanov and S. Tindel, Sampling linear inverse problems with noise, Asymptotic Analysis, 132 (2023), pp. 331–382.
- [32] M. E. TAYLOR, Partial Differential Equations I: Basic Theory, Applied Mathematical Sciences, Springer, New York, NY, 2011.
- [33] F. Treves, Introduction To Pseudodifferential and Fourier Integral Operators Vol. 1: Pseudodifferential Operators, Plenum Press, New York and London, 1980.
- [34] ——, Introduction To Pseudodifferential and Fourier Integral Operators. Vol. 2: Fourier Integral Operators, Plenum Press, New York and London, 1980.
- [35] S. Vanska, M. Lassas, and S. Siltanen, Statistical X-ray tomography using empirical Besov priors, International Journal of Tomography & Statistics, 11 (2009), pd. 3-32.
- [36] A. Wunderlich and F. Noo, Image covariance and lesion detectability in direct fan-beam X-ray computed tomography, Physics in Medicine and Biology, 53 (2008), pp. 2471–2493.