Department of Applied Physics

Final Year Project Report

Towards Understanding and Optimising Ptychography for High-Performance Electron Microscopy

Student Name:	LU Shan
Student Number:	22098858D
Programme Code:	11443-XAI
Supervisor:	Prof. ZHU Ye
Co-supervisor:	Dr. CHEN Changsheng
Academic Year:	2025–2026

The Hong Kong Polytechnic University

Abstract

This thesis analyses PtyRAD as an automatic-differentiation-based reconstruction system for electron ptychography. It first explains the forward model, trainable variables, objective terms, and implementation choices used in modern ptychographic reconstruction. It then examines a controlled set of optimisation studies on a reduced $SrTiO_3$ baseline, focusing on objective design, staged activation, batch geometry, and mixed-state flexibility. The thesis shows how these choices change reconstruction behaviour and how the resulting changes should be interpreted. The results show that different objective terms regulate different scales of agreement; staged activation changes the optimisation path and the distribution of adaptation across parameter blocks; batch geometry changes how overlap information enters each update; and mixed-state flexibility improves fit but does not by itself identify a unique physical mechanism. Together, these findings show that PtyRAD behaves as a coupled nonconvex inverse problem whose practical behaviour depends on modelling choices, update schedule, and data partitioning, and they provide a clearer basis for future optimisation and regularisation studies.

Acknowledgement

The completion of this thesis marks, on the surface, the end of a technical inquiry into electron ptychography, multislice propagation, and optimisation. To me, however, it also reflects a quieter journey: learning how to enter an unfamiliar field, endure uncertainty, and continue forward even when understanding seemed always just beyond reach.

First and foremost, I owe a profound debt to **Prof. ZHU Ye**, my supervisor. It was in the quiet spaces of our conversations that this project first drew breath, and it was his guidance that kept it from losing its way. He did not simply point me toward answers; he taught me how to sit with a difficult question, to let its inner logic reveal itself at its own pace. His insight was a light, his patience a steady anchor, and his trust a quiet permission to grow.

I am equally grateful to **Dr. CHEN Changsheng**, whose support formed the bedrock of this work. The data that give this thesis its empirical weight, and the experimental acquisitions that brought it to life, were made possible only through his generous help.

I am also thankful to **Dr. Cheung Yat-ming**, a friend whose expertise in optimisation theory offered me another lens through which to see this project.

I am likewise especially grateful to **Ms. CHEN Xin** for her guidance in writing. Her advice did more than refine the language of this thesis; it taught me to take care to weigh each word, to let ideas breathe on the page. She showed me that writing is not merely a vessel for thought, but a way of thinking more clearly, more kindly, and more honestly.

Finally, I would like to thank my family. During many moments of difficulty, frustration, and self-doubt that accompanied this project, they gave me patience, encouragement, and emotional steadiness. Their support carried me through the times when the work felt heavy, and reminded me that perseverance is sometimes simply another form of hope.

To have completed this work is therefore not only a matter of gratitude to all those who guided and supported me, though that gratitude is profound. It is also, perhaps, a quiet gratitude to the earlier version of myself who chose not to turn away from a difficult problem simply because it was difficult. May this spring, which closes one chapter, also open the way to every spring yet to come.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Electron microscopy seeks quantitative information about matter at the atomic scale, yet standard detectors record only intensity rather than the full complex electron wave. This creates a fundamental limitation for imaging: although the measured intensity reveals how much signal arrives at the detector, it does not directly retain the phase information that describes how the wave has been modified by its interaction with the specimen. In many physically relevant regimes, the missing phase carries information of direct scientific interest, including the specimen's projected electrostatic potential and, under suitable approximations, other weak interaction effects. Recovering phase is therefore essential for making electron microscopy a quantitatively meaningful tool for materials characterisation.

A compact expression of the difficulty is

$$I(\mathbf{q}) = |\Psi(\mathbf{q})|^2, \qquad \Psi(\mathbf{q}) = \mathcal{F}\{\psi(\mathbf{r})\},$$

where $\psi(\mathbf{r})$ denotes the complex electron wave in real space and $\Psi(\mathbf{q})$ its diffraction-space representation. The detector measures only $I(\mathbf{q})$, so the phase of $\Psi(\mathbf{q})$ is lost in the measurement process. This loss makes naive inversion impossible in general: many distinct complex waves may produce the same measured intensity. From the viewpoint of inverse problems, the reconstruction task is therefore ill-posed unless additional diversity, prior structure, or physical constraints are introduced. This point is especially important in electron imaging, where the scientifically meaningful information is often encoded precisely in the phase that the detector does not directly observe.

Electron ptychography addresses this difficulty by replacing a single intensity measurement with a structured family of measurements. In a 4D-STEM experiment, a focused probe is scanned across the specimen, and a diffraction pattern is recorded at each probe position. The key idea, however, is not simply that many diffraction patterns are

acquired, but that neighbouring probe positions overlap in real space. Because adjacent measurements share parts of the specimen while weighting them differently through the probe, the dataset acquires strong redundancy. This overlap-generated diversity transforms phase retrieval from an underconstrained single-measurement problem into a more tractable inverse problem with rich internal consistency. Ptychography has therefore emerged as a powerful route towards high-resolution and quantitatively meaningful electron imaging.

Modern ptychographic reconstruction typically requires joint estimation of multiple coupled variables, including the specimen transmission function, the illumination probe, scan positions, detector effects, partial coherence, and, in more advanced models, multislice or geometric parameters.This is one reason why automatic-differentiation-based frameworks have become so important. Rather than relying exclusively on hand-derived update rules for narrowly specified models, automatic differentiation makes it practical to define a differentiable forward model and jointly optimise multiple physical variables within a unified computational graph.

This thesis focuses on PtyRAD as a particularly useful case study in this modern reconstruction paradigm. PtyRAD provides a concrete setting in which object, probe, positions, thickness, tilt, and other physically meaningful variables can be estimated within the same reconstruction pipeline. Yet precisely because of this flexibility, the theoretical structure of the framework is not immediately transparent. For a new researcher, it can be easy to use such a system as a black-box reconstruction tool while remaining uncertain about what is actually being optimised, how the forward model is structured, which variables are coupled, and why convergence may plateau or fail. Accordingly, this thesis analyses PtyRAD as a structured inverse problem and studies its physical model, computational graph, and optimisation behaviour in a systematic way.

This perspective also explains why optimisation occupies an important role in the present work. Automatic differentiation makes larger and more realistic reconstruction models implementable and jointly optimisable. The underlying problem nevertheless remains nonconvex and can still be ill-conditioned. In PtyRAD, reconstruction is defined jointly by the forward model, the objective, the staged update schedule, and the iteration-wise constraints. These components together form a coupled nonconvex reconstruction system. Understanding such a system therefore requires analysing which parameter blocks absorb residual error during fitting, how that residual shifts between

object, probe, and geometry variables, and how objective choice and batching expose different update information to the optimiser.

Against this background, the motivation of this thesis is twofold. First, it aims to develop a rigorous and physically intuitive understanding of electron ptychography as implemented in PtyRAD, with particular attention to the forward model and the structure of the optimisation problem. Second, it examines a small number of controlled optimisation studies on a practical reduced experimental baseline. The goal is to draw conditional conclusions about how PtyRAD behaves under the present baseline, which optimisation choices materially change that behaviour, and which directions for improvement remain most credible on physical and numerical grounds.

1.2 Project Objectives

Motivated by the considerations outlined above, this thesis pursues two tightly connected objectives. The first is to develop a rigorous and physically intuitive understanding of electron ptychography as implemented in PtyRAD. This involves clarifying the physical and mathematical structure of the forward model, identifying the role of the principal reconstruction variables, and formulating the overall reconstruction process as a structured inverse problem. In particular, the thesis seeks to make explicit how quantities such as the object transmission function, illumination probe, scan positions, multislice and mixed-state components, and selected system-side variables enter the differentiable forward model, and how their associated gradients participate in iterative reconstruction within an automatic-differentiation-based framework.

The second objective is to examine a small set of scientifically interpretable optimisation choices within PtyRAD, especially objective design, staged activation of parameter blocks, batch geometry, and limited late-stage refinement. These are studied under a common practical reduced baseline, so that any claims of improvement or diagnostic significance remain tied to observable reconstruction behaviour rather than to heuristic preference alone.

These two objectives make the thesis both explanatory and evaluative. The thesis studies PtyRAD as a physical model embedded in a nonconvex optimisation problem and uses that understanding to evaluate which optimisation choices most clearly influence practical behaviour. The broader goal is to determine what kind of reconstruction system

PtyRAD is, why it behaves as it does, and how its behaviour can be studied in a principled and experimentally accountable way.

1.3 Research Questions

The objectives of this thesis are pursued through three central research questions. Together, these questions specify the analytical problems that organise the chapters that follow and provide the basis for both theoretical discussion and controlled experimental evaluation.

(a) How can electron ptychography be formulated as a physics-constrained nonconvex optimisation problem, and what does this formulation reveal about identifiability, parameter coupling, and reconstruction failure modes?

(b) How does an automatic-differentiation-based reconstruction framework such as PtyRAD change the modelling and implementation paradigm relative to the classical PIE-family approach based on hand-derived update rules?

(c) Under a practical reduced experimental baseline, how do objective design, stage-dependent loss scheduling staged activation, batch geometry, and limited late-stage refinement influence reconstruction behaviour, and what do these controlled studies reveal about freedom allocation, residual mismatch, and optimisation path?

These questions position the thesis at the intersection of physical modelling and optimisation. The first question establishes the mathematical framing of the reconstruction problem, the second clarifies the conceptual significance of the automatic-differentiation paradigm, and the third links this understanding to the controlled studies through which optimisation behaviour is examined in practice.

1.4 Scope and Delimitations

The scope of this thesis is deliberately defined so as to balance theoretical depth with methodological realism. Its primary focus is the study of electron ptychography through the specific case of PtyRAD, with emphasis on the physical forward model, the mathematical structure of the reconstruction problem, and the optimisation behaviour induced by an automatic-differentiation-based implementation.

Accordingly, the central analysis is restricted to questions that are both scientifically meaningful and feasible within the present project. These include the formulation of reconstruction as a structured inverse problem, the interpretation of the principal trainable variables, loss terms, and constraint-enforcement steps, and the study of a limited set of controlled optimisation interventions that can be justified on physical or numerical grounds. In particular, empirical part of thesis is centred on practical reduced baseline for the experimental $SrTiO_3$ dataset, together with a small number of controlled comparisons involving objective design, batch geometry, and limited late-stage refinement.

At the same time, several important delimitations should be made explicit. First, the thesis does not aim to provide an exhaustive survey of all ptychographic algorithms or all variants of 4D-STEM reconstruction. Classical iterative schemes such as PIE, ePIE, rPIE, and related developments are discussed only to the extent necessary to situate PtyRAD within a broader methodological lineage.

Second, the project does not focus on low-level kernel engineering, large-scale software acceleration, or hardware-specific optimisation. While computational efficiency is certainly relevant in practice, such issues are not the principal object of study here.

Third, the thesis does not attempt an exhaustive multi-dataset benchmark or a broad hyperparameter search. Its experimental claims are intentionally tied to one reduced experimental baseline and a common comparison horizon, so that the resulting conclusions remain interpretable in optimisation terms.

Finally, the thesis does not claim to establish a universally optimal reconstruction pipeline. Ptychographic reconstruction is highly sensitive to experimental conditions, model choice, initialisation, parameterisation, and data quality, so any proposed improvement must be interpreted within context rather than as a context-free prescription. For this reason, the optimisation component of the project is framed as an investigation of credible and interpretable strategies rather than as a search for a single best method. Where improvements are identified, they are evidence-based and conditional, supported by controlled experimental settings rather than by sweeping generalisation.

1.5 Structure of the Thesis

This thesis is organised as a progressive argument. It begins from the physical motivation of phase retrieval, develops the mathematical structure of the corresponding forward

model, reformulates reconstruction as a structured optimisation problem, and finally tests that viewpoint through controlled experimental study and evidence-tiered discussion.

Chapter 2 establishes the conceptual and historical background of the problem. It reviews the phase problem in electron microscopy, explains the basic mechanism of ptychography and 4D-STEM ptychography, and situates PtyRAD within the broader algorithmic lineage from PIE to ePIE, rPIE, and related developments.

Chapter 3 then turns to the physics of the PtyRAD forward model itself. Its purpose is to make explicit how the principal ingredients of reconstruction, including the object, probe, scan geometry, diffraction formation, and selected extensions such as multislice propagation, mixed-state modelling, and detector-side measurement formation, fit together within a unified formulation.

Once this structure has been clarified, Chapter 4 reformulates the reconstruction task as a physics-constrained nonconvex optimisation problem. This chapter identifies the main parameter groups, interprets the role of the objective, regularisation, constraint enforcement, and gradients, and analyses the sources of coupling, ill-conditioning, and plateau behaviour that later motivate the experimental study.

With this formulation in place, the thesis then turns from understanding to controlled evaluation. Chapter 5 establishes the experimental baseline and diagnostic framework, documents the practical reduced $SrTiO_3$ configuration used throughout the study, analyses its behaviour and limitations, and formulates the working hypotheses.

Chapter 6 builds directly on this baseline by examining a selected set of controlled optimisation studies under fixed conditions. Its main comparisons concern objective design, including stage-dependent loss schedule, staged activation, batch geometry, and limited late-stage refinement, while the ambiguity of mixed-state probe freedom is retained as a cross-family interpretive axis.

Chapter 7 then provides an evidence-tiered discussion. It separates what is directly supported by present experiments from what remains plausible but not yet fully validated, and reflects on what these results reveal about PtyRAD as an optimisation problem.

Finally, Chapter 8 concludes the thesis by drawing together its two main contributions, rigorous understanding and optimisation-aware evaluation, and by outlining future directions that follow naturally from the limits and opportunities identified in the preceding chapters.

Chapter 2

Literature Review and Conceptual Background

2.1 The Phase Problem in Electron Microscopy

In electron microscopy, phase matters because it carries physically meaningful information about the specimen that intensity-only imaging does not preserve directly [1]. In particular, the phase of the electron wave is closely related to specimen properties such as projected electrostatic potential, internal electric or magnetic fields, and other structural features that may be only weakly expressed in intensity contrast alone [1]. This point is especially important for weakly scattering or light-element materials, where the physically relevant signal may be encoded more clearly in phase than in conventional image intensity [2]. From this perspective, the phase problem should not be regarded as a merely formal limitation of imaging theory, but as a genuine obstacle to recovering physically informative specimen content.

This physical significance also explains why phase retrieval is closely connected to the broader goal of quantitative, rather than merely qualitative, electron imaging [1]. Under thin-sample or weak-phase conditions, the phase of the transmission function is closely related to the projected potential of the specimen, so loss of phase information directly limits quantitative interpretation of the underlying material [2]. More generally, access to phase can improve sensitivity in dose-limited imaging and in settings where conventional contrast mechanisms are weak or difficult to interpret reliably [2]. The practical importance of this issue is reflected in the development of modern electron ptychography, where phase-sensitive reconstruction has enabled high-resolution and high-contrast imaging in experimentally relevant conditions [3]. For the present thesis, however, the key point is not yet the success of any particular reconstruction method, but the more fundamental fact that the physically informative part of the wave is not directly available from standard intensity measurements.

The origin of the phase problem is therefore rooted in measurement structure. In a conventional diffraction or imaging measurement, the detector records intensity,

which may be written schematically as $I(\mathbf{q}) = |\Psi(\mathbf{q})|^2$, rather than the full complex wavefunction $\Psi(\mathbf{q})$ itself [1]. As a result, the amplitude may be observed, but the associated phase is not directly retained. This means that the mapping from specimen to measurement is intrinsically incomplete: different complex waves may be consistent with the same recorded intensity, so the underlying object cannot in general be recovered by naive inversion alone. In this sense, the phase problem is naturally understood not only as a limitation of measurement, but also as an inverse-problem difficulty, since recovering the missing phase requires information beyond that contained in a single intensity measurement.

Phase retrieval in electron microscopy is not synonymous with ptychography; different approaches introduce different forms of measurement diversity and therefore define different inverse problems. Examples include off-axis electron holography and transport-of-intensity-based methods, each of which supplements intensity data in a different way and therefore imposes a different reconstruction structure [4, 5]. The present thesis focuses on ptychography, not because it is the only route to phase retrieval, but because it offers a particularly powerful framework in which additional recoverability is created through structured redundancy across multiple overlapping diffraction measurements. The central question is therefore not simply how to invert intensity, but what additional structure, constraints, or redundancy can render phase retrieval recoverable. It is this question that leads naturally to the basic mechanism of ptychography discussed in the next section [6].

2.2 4D-STEM and Basic Mechanism of Ptychography

What makes ptychography powerful is not merely the collection of many diffraction patterns, but the structured relationship among them created by shifted probe illumination [1, 7]. In the electron microscopy setting, this mechanism is commonly realised through 4D-STEM data acquisition, in which a two-dimensional diffraction pattern is recorded at each probe position on a two-dimensional scan grid [6]. It is therefore useful to distinguish clearly between the two ideas: 4D-STEM is an acquisition modality, whereas ptychography is a reconstruction method that operates on such data [6]. The significance of this distinction is that ptychography does not derive its strength simply from having more data, but from how those data are related to one another through the measurement

design itself.

At the core of this design is the interaction between a localised probe and the specimen at a sequence of shifted scan positions. If the object is denoted by $O(\mathbf{r})$ and the probe by $P(\mathbf{r})$, then at scan position $\mathbf{R}_j$ the resulting exit wave may be written schematically as

$$\psi_j(\mathbf{r}) = P(\mathbf{r} - \mathbf{R}_j)\, O(\mathbf{r}),$$

and the detector records the corresponding far-field intensity

$$I_j(\mathbf{q}) = \left|\mathcal{F}\{\psi_j\}(\mathbf{q})\right|^2 .$$

These expressions are sufficient at the present stage to show the central idea: each scan position does not produce an entirely independent measurement, but a diffraction pattern associated with a locally illuminated region of the same specimen [1, 8]. In this sense, ptychography replaces a single underdetermined intensity measurement with a structured family of related measurements whose differences are governed by controlled probe translations.

The crucial point is that overlap occurs in real space, through the illumination footprints of neighbouring probe positions. When adjacent probe positions partially cover the same region of the object, the corresponding diffraction patterns are not independent snapshots, but different encodings of shared specimen information [7, 9]. It is this overlap that creates redundancy across the dataset and, more importantly, enforces self-consistency: reconstructions inferred from one measurement must agree with those inferred from neighbouring measurements on the regions they share. Ptychographic recoverability therefore does not arise simply because more diffraction patterns are collected, but because overlap makes phase retrieval more recoverable by creating redundant measurements that must remain mutually consistent [1, 8]. Compared with a single diffraction measurement, where many complex wavefields may remain compatible with the same recorded intensity, an overlapping scan set constrains the admissible solution much more strongly.

This is also why 4D-STEM provides such a natural setting for modern electron ptychography. Since ptychographic reconstruction requires the full diffraction distribution at each scan position rather than an integrated detector signal, pixelated detectors are especially important: they make it possible to retain the spatially resolved diffraction

information needed for overlap-based phase retrieval [2, 6]. The practical rise of electron ptychography in recent years is therefore closely linked to the emergence of detector technology capable of recording such datasets efficiently, although detector development itself is not the focus of the present thesis [6].

The importance of overlap is not only conceptual but also operational. As shown in the early ptychographic literature, the overlap parameter can materially influence convergence behaviour and reconstruction quality, which indicates that redundancy is not a cosmetic feature of the method but one of the main sources of its practical recoverability [9]. More broadly, ptychography includes both direct and iterative reconstruction families, but in all cases the central challenge is how effectively the reconstruction scheme exploits the structured redundancy present in the data [1]. The measurement design therefore explains why ptychography is recoverable in principle, but practical reconstruction still depends on how object and probe are updated in computation.

2.3 From PIE to ePIE, rPIE, and mPIE

The development of the PIE family can be read as a sequence of responses to increasingly realistic reconstruction challenges [1]. Once overlap-generated redundancy is available, the central computational question is no longer whether phase retrieval is possible in principle, but how that redundancy should be converted into stable and informative updates of the unknown quantities. From PIE to ePIE, rPIE, and mPIE, the key change is a progressively clearer account of what must be updated and how.

The original ptychographical iterative engine (PIE) provided the basic update philosophy on which much of the later literature was built [7]. Its key idea was that each diffraction measurement could be used to correct the corresponding locally illuminated region of the object by comparing the predicted exit wave with one whose Fourier amplitude had been replaced by the measured data. In this sense, PIE offered a concrete computational mechanism for exploiting overlap, since information obtained at one scan position could be propagated back into the object estimate and subsequently checked for consistency at neighbouring positions. At the same time, however, classical PIE typically assumed that the illuminating probe was sufficiently well known. This made PIE an elegant and foundational starting point, but also exposed an important limitation: in realistic experiments, the probe is often not known with the precision that the algorithm

ideally requires.

This limitation motivated the development of the extended ptychographical iterative engine (ePIE), which is the most important algorithmic anchor for the present thesis [10]. The conceptual shift introduced by ePIE is not simply that another update rule was proposed, but that the reconstruction problem was reformulated to allow the object and the probe to be updated jointly. In other words, ePIE relaxes the restrictive assumption of a perfectly known probe and treats illumination uncertainty as part of the reconstruction task itself. This is especially important in electron microscopy, where the probe is shaped by experimental conditions that may not be known exactly in advance [11]. From this perspective, ePIE marks a decisive step toward a more realistic and practically useful reconstruction framework. For this reason, ePIE is the central bridge between early overlap-based reconstruction and later optimisation-oriented frameworks.

Subsequent developments within the PIE lineage responded less to modelling assumptions than to numerical behaviour during reconstruction. In particular, it became increasingly clear that standard PIE- or ePIE-style updates could become inefficient or unstable in weakly illuminated regions, and that difficult datasets could lead to slow convergence or stagnation [12]. The regularized PIE (rPIE) addresses the first of these problems by modifying the update normalization so that corrections are better balanced across regions of different probe intensity. The effect is not merely heuristic: it reflects a more careful treatment of how strongly different parts of the object or probe should be corrected when the local illumination is uneven. The momentum-accelerated variant mPIE, in turn, addresses the second problem by introducing an accumulated update direction that can reduce oscillatory behaviour and improve convergence speed on challenging data [12]. More recent work such as LoRePIE suggests that this logic of failure-mode-driven refinement remains active, with regularisation continuing to be explored for low-dose and otherwise difficult reconstruction regimes [13].

Taken together, the PIE family clarifies a central point that is highly relevant for this thesis: ptychographic reconstruction is shaped not only by the measurement design, but also by the choice of update rule and by the assumptions built into that rule. Although PIE-family methods constitute only one important branch of the broader ptychographic literature, alongside, for example, projection-based and objective-based formulations [1, 8, 14], they make especially clear what quantities must be updated, why those updates can become unstable, and how algorithm design mediates between physical

modelling and computational performance. The historical evolution of PIE-family algorithms therefore helps explain why more flexible optimisation frameworks later became attractive.

2.4 Automatic Differentiation and AD-based Reconstruction

Automatic differentiation became attractive in ptychography as reconstruction models grew more complex and began to include a wider set of variables and physical effects [1, 15]. As discussed previously, the PIE family clarified how overlap-generated redundancy could be exploited through iterative updates of the object and, later, the probe. However, as ptychographic models became more realistic, the reconstruction task began to involve a wider set of unknowns and physical effects than could be handled conveniently through bespoke update design alone. In this setting, the central issue was no longer only how to craft a physically meaningful update rule, but also how to implement increasingly rich forward models without having to re-derive an entirely new reconstruction scheme each time the model was extended.

A natural bridge toward this modern view was the growing use of objective-based formulations, in which reconstruction is expressed explicitly as the minimization of a data-fidelity term, often supplemented by constraints or regularization [8, 14]. In schematic form, one may write

$$\min_{\theta} \; \mathcal{L}\big(y, \hat{y}(\theta)\big) + R(\theta),$$

where θ denotes the collection of optimisable variables, $\hat{y}(\theta)$ is the predicted measurement generated by the forward model, $\mathcal{L}$ measures disagreement with the data, and R denotes any regularizing or constraint-enforcing term. This formulation makes explicit that ptychographic reconstruction can be posed as an optimisation problem in which modelling uncertainty, experimental imperfections, and additional latent variables enter the same framework [14]. In this sense, objective-based methods provided an important conceptual bridge between classical update-rule design and later AD-based reconstruction.

Against this background, the main contribution of automatic differentiation is to shift the implementation focus toward the differentiable forward model itself [15]. Once the predicted measurements are written as a differentiable function of the chosen

variables, gradients of the reconstruction objective can be generated automatically rather than being derived and coded manually for each new model variant. Physical insight remains essential, but in AD-based reconstruction it is expressed mainly through the forward model, the choice of optimisable variables, and the loss design rather than through hand-derived update rules. The resulting framework is especially attractive when one wishes to incorporate additional physical effects or jointly optimize multiple parameter groups, since new variables can in principle be introduced by extending the computational graph rather than rebuilding the reconstruction method from scratch [15].

At the same time, the advantages of AD should not be overstated. Automatic differentiation does not remove the nonconvexity of ptychographic reconstruction, nor does it automatically resolve conditioning difficulties, stagnation, or hyperparameter sensitivity. What it changes is the way gradients are obtained and the way reconstruction models are expressed. The value of AD lies less in guaranteeing easy optimisation than in making complex ptychographic models easier to formulate, implement, and extend in a unified reconstruction workflow. It is precisely this shift that helps explain why modern AD-based frameworks became attractive in electron ptychography, and for this reason, PtyRAD is worth studying as a concrete example of the move from hand-derived update rules to AD-based reconstruction.

2.5 Positioning PtyRAD in the Literature

PtyRAD is useful for this thesis because it allows object, probe, scan corrections, tilt, slice thickness, and mixed-state variables to enter one differentiable reconstruction model [16]. For this thesis, PtyRAD is valuable because it turns the AD-based reconstruction paradigm into a concrete system that can be analysed in physical, computational, and optimisation terms.

Its main practical strength is that it allows multiple physically meaningful parameter groups to be optimised together within the same reconstruction. It also accommodates more advanced modelling components such as multislice propagation, regularization or constraint terms, and workflow-level procedures including hyperparameter tuning [16]. These features matter because they bring experimentally relevant model complexity into the reconstruction itself rather than leaving it outside the optimisation loop. Its reported efficiency also matters because the practical value of joint optimisation depends

on whether these larger coupled models can still be run at useful cost [16].

PtyRAD should nevertheless be positioned carefully within the broader literature. It is not the only modern software environment to reflect the trend toward more flexible and extensible ptychographic reconstruction, and recent work such as *phaser* likewise points toward unified frameworks that combine conventional and gradient-based reconstruction engines within a broader software architecture [17]. Even so, PtyRAD remains particularly suitable for the present thesis because it offers a focused, electron-ptychography-specific, AD-first system in which the interaction between forward modelling and optimisation can be examined concretely. Its value here lies not in replacing the earlier algorithmic literature, but in providing a modern reconstruction framework through which that literature can be reinterpreted and connected to current practice. The following chapters therefore use PtyRAD as the concrete framework for analysing the forward model and optimisation structure of modern electron ptychography.

Chapter 3

Physics of the PtyRAD Forward Model

3.1 Probe, Object, and Scanning Geometry

The forward model of electron ptychography begins with three ingredients: an incident probe, a specimen transmission function, and a scan geometry that organises repeated local interactions across the dataset. Before propagation, multislice effects, or optimisation variables can be discussed, these quantities must be specified in a physically meaningful way, because every later stage of the forward model is built on them [1, 6, 18].

The probe is represented as a complex wavefield in the specimen plane,

$$P(\mathbf{r}) \in \mathbb{C}, \tag{3.1}$$

where $\mathbf{r} = (x, y)$ denotes the transverse specimen-plane coordinate. Its amplitude describes how strongly the specimen is illuminated across the probe footprint, while its phase records the local wavefront structure inherited from the electron optics. The probe is a structured illumination wavefield whose amplitude and phase directly shape the recorded diffraction patterns.

The specimen is represented by a complex transmission function

$$O(\mathbf{r}) = A(\mathbf{r})e^{i\phi(\mathbf{r})} \in \mathbb{C}, \tag{3.2}$$

where $A(\mathbf{r})$ describes amplitude modulation and $\phi(\mathbf{r})$ describes phase modulation. This factorisation separates attenuation from phase retardation and will remain useful throughout the thesis. Later sections will show that it also provides the natural bridge from physical modelling to computational representation and, eventually, to the object-related parameter blocks introduced in Chapter 4.

The third ingredient is the scan geometry. Let $\mathbf{R}_j$ denote the j-th probe coordinate

in the scan. The corresponding shifted probe is written as

$$P_j(\mathbf{r}) = P(\mathbf{r} - \mathbf{R}_j). \tag{3.3}$$

At this stage, the scan index labels a local probe–object configuration[1]; the corresponding detector-plane measurement is introduced in the next section.

What makes ptychography distinctive is the deliberate real-space overlap between neighbouring probe positions. Adjacent diffraction patterns are therefore not independent observations of unrelated specimen regions; they are coupled measurements of partially shared object content under shifted probe illumination. This overlap is the physical origin of the inter-measurement redundancy on which ptychographic phase retrieval relies [10, 19].

To make this acquisition geometry more concrete, Figure 3-1 provides a schematic summary of the experimental arrangement underlying ptychographic measurement. The figure is included here not as a complete forward-model diagram, but as a visual aid for the three ingredients introduced in this section: the focused probe, the scan-position indexing, and the deliberate real-space overlap between neighbouring illuminations.

From a modelling viewpoint, the scan positions serve three simultaneous roles:

(a) they define the acquisition geometry in real space;

(b) they determine which local object region enters each forward evaluation; and

(c) they generate the overlap-based coupling that links neighbouring measurements.

Scan geometry therefore enters the forward operator itself rather than appearing as passive metadata. Once the probe, object, and scan positions are fixed, the next step is to determine how one such local configuration is converted into a detector-plane measurement.

3.2 From Exit Wave to Recorded Diffraction Intensity

With the local ingredients now specified, a single diffraction measurement is generated by a short but consequential chain: a local probe–object interaction produces an exit

[1] In the current PtyRAD implementation, this distinction later appears computationally as a decomposition of scan coordinates into integer crop locations and sub-pixel residual shifts. The continuous notation $\mathbf{R}_j$ is retained here because it is better suited to the physical discussion.

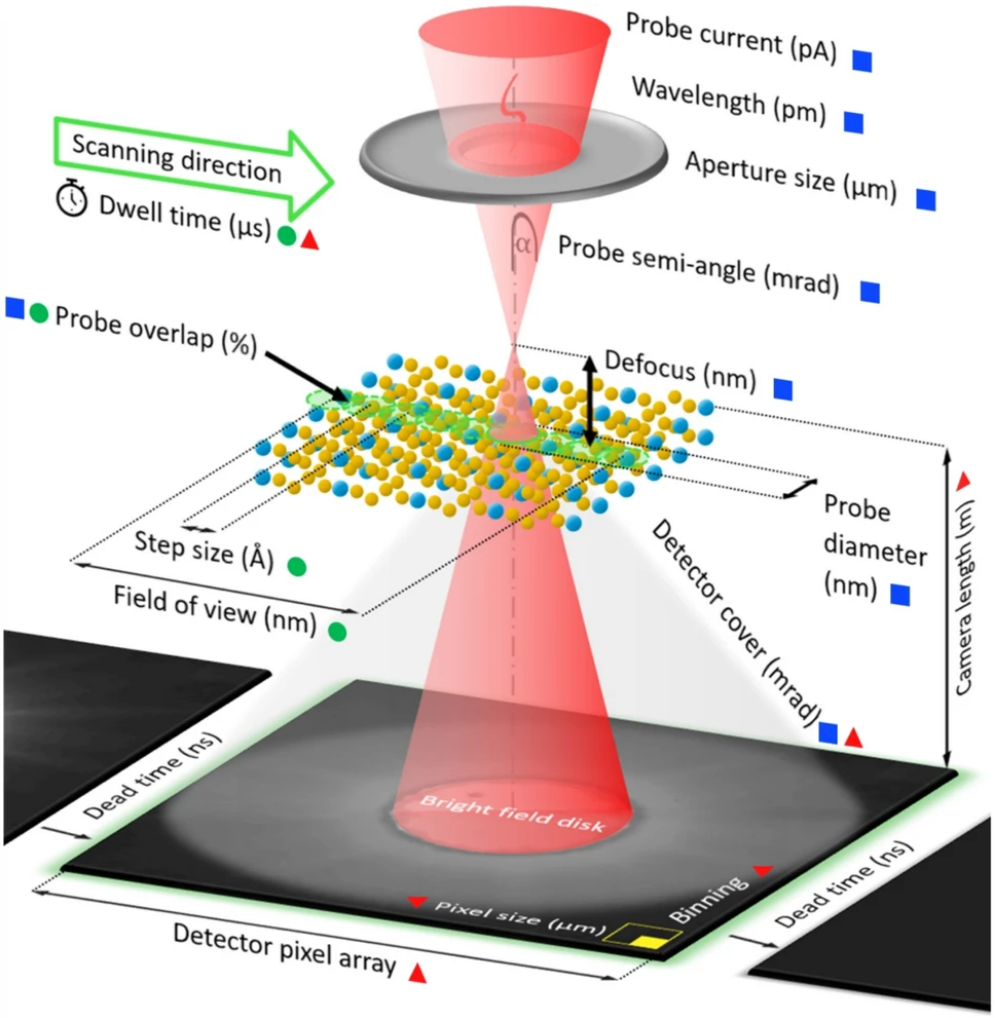


Figure 3-1. Acquisition-side schematic of probe formation and scan geometry in electron ptychography. A focused incident probe is rastered across the specimen, generating overlapping illuminated regions whose associated diffraction patterns form the ptychographic dataset. The figure visually summarises the roles of probe size, scan step, field of view, and probe overlap in organising the real-space measurement geometry. Detector-related quantities are also indicated to show how the acquisition setup connects to the recorded diffraction data.

wave, propagation maps that wave to the detector plane, and the detector records only intensity. In the simplest single-slice picture, the exit wave associated with scan position $\mathbf{R}_j$ is

$$\psi_j(\mathbf{r}) = O(\mathbf{r})\, P(\mathbf{r} - \mathbf{R}_j). \tag{3.4}$$

This quantity remains a real-space complex wavefield. It is indexed by the scan position, and it already contains the multiplicative probe–object coupling that later becomes central to the optimisation problem.

Figure 3-2 summarises the measurement chain. The important distinction is between the real-space exit wave $\psi_j(\mathbf{r})$, the reciprocal-space detector-plane wave $\Psi_j(\mathbf{q})$, and the recorded intensity $I_j(\mathbf{q})$.

Under the far-field approximation, the detector-plane complex amplitude is given by the Fourier transform of the exit wave,

$$\Psi_j(\mathbf{q}) = \mathcal{F}\{\psi_j(\mathbf{r})\}(\mathbf{q}), \tag{3.5}$$

where $\mathbf{q}$ denotes the reciprocal-space detector coordinate. The corresponding ideal diffraction intensity is then

$$I_j(\mathbf{q}) = \left|\Psi_j(\mathbf{q})\right|^2 = \left|\mathcal{F}\{\psi_j(\mathbf{r})\}(\mathbf{q})\right|^2. \tag{3.6}$$

Eqs. (3.4)–(3.6) define the single-position chain

$$P(\mathbf{r} - \mathbf{R}_j),\ O(\mathbf{r}) \longrightarrow \psi_j(\mathbf{r}) \longrightarrow \Psi_j(\mathbf{q}) \longrightarrow I_j(\mathbf{q}).$$

This chain is compact, but it already contains the two structural difficulties that later motivate Chapter 4. First, the exit wave depends multiplicatively on the probe and object. Second, the detector preserves only the modulus square of the propagated wave, not its phase. A single diffraction pattern therefore does not determine a unique complex wavefield [1, 18].

The forward model is accordingly not one diffraction pattern but a family of scan-indexed intensities

$$\{I_j(\mathbf{q})\}_{j=1}^{J},$$

whose redundancy comes from overlap across measurements rather than from any

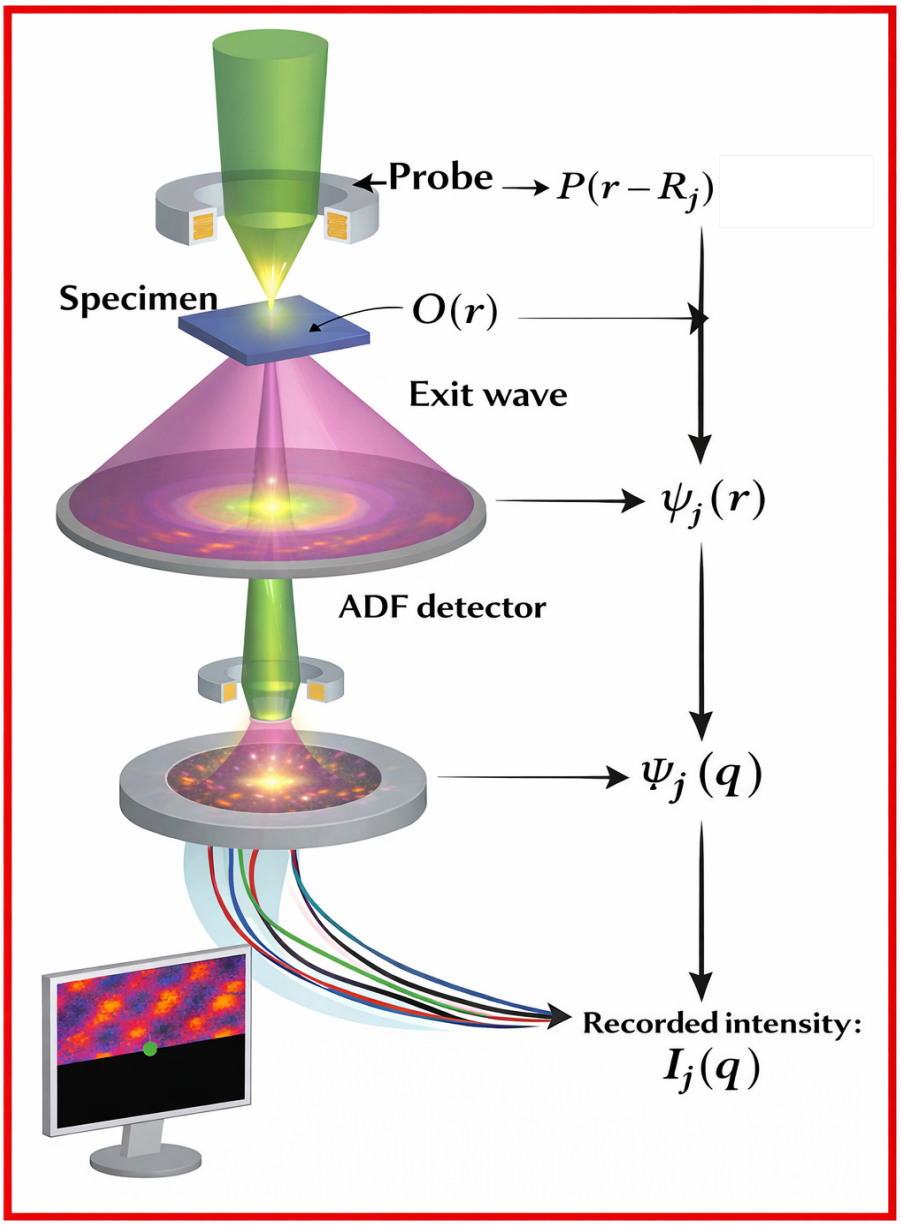


Figure 3-2. Schematic illustration of the forward measurement chain in electron ptychography. A shifted probe $P(\mathbf{r} - \mathbf{R}_j)$ interacts with the specimen transmission function $O(\mathbf{r})$ to produce the exit wave $\psi_j(\mathbf{r})$ in real space. This wave is then propagated to the detector plane, giving the reciprocal-space wave $\Psi_j(\mathbf{q})$, whose intensity $I_j(\mathbf{q})$ is finally recorded.

individual pattern in isolation [10, 19]. In implementation, the same stage later appears as an FFT-based mapping to reciprocal space followed by an absolute-square operation, before additional refinements such as multislice propagation, mixed-state summation, and detector response are introduced.[2]

For the purposes of the present chapter, Eqs. (3.4)–(3.6) establish the basic measurement backbone. The next refinement is physical rather than numerical: the single transmission event must be replaced by depth-resolved propagation when multiple scattering becomes significant.

3.3 Multislice Propagation Model

The single-slice picture developed in the previous section provides a useful starting point for understanding how a local probe–object interaction gives rise to a diffraction measurement. It nevertheless assumes that the specimen can be compressed into one effective transmission layer, so that incident wave is modified only once before propagating to the detector. For sufficiently thin samples this projected description may be adequate, but once the specimen has appreciable thickness the electron wave undergoes successive interactions at different depths and multiple scattering can no longer be neglected. In that regime, the forward model must resolve propagation through specimen rather than absorb the entire sample into a single multiplicative object [20, 21].

The multislice model does this by discretising the specimen along the beam direction into an ordered family of thin slices. Instead of one transmission function $O(\mathbf{r})$, one introduces slice-wise transmissions $O_n(\mathbf{r})$, so that specimen interaction becomes a depth-ordered composition of local modulations separated by inter-slice propagation [21, 22]. The key physical point is that, in a thick specimen, the field reaching slice $n+1$ already carries the cumulative effect of the preceding slices.

To make this approximation more intuitive, Figure 3-3 illustrates the central idea of the multislice model. A specimen of finite thickness is replaced by a stack of thin slices along the beam direction. Each slice is treated, to first approximation, as a thin transmission layer, while neighbouring slices are connected by free-space propagation through the corresponding inter-slice spacing. In this way, a volumetric specimen is

[2] This implementation viewpoint is deferred until Sections 3.6 and 3.7, where the physical chain is recast as tensor operations and a differentiable operator map.

converted into a sequential operator chain that remains computationally tractable while still capturing depth-dependent wave evolution.

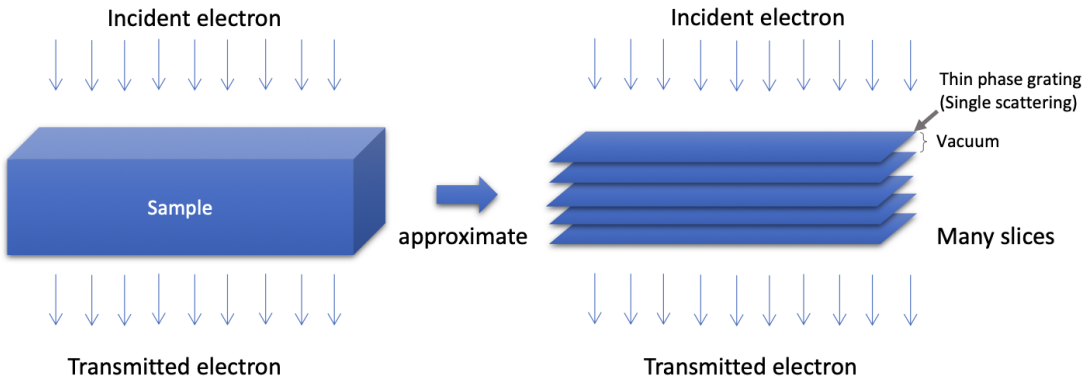


Figure 3-3. Schematic illustration of the multislice approximation. A specimen of finite thickness is represented as a sequence of thin slices along the beam direction. Each slice is treated as a thin transmission layer, and adjacent slices are connected by free-space propagation across the inter-slice gaps. This provides the physical basis for modelling depth-dependent wave evolution in the multislice forward model.

For a probe positioned at $\mathbf{R}_j$, the multislice forward model begins from incident wave

$$\psi_j^{(0)}(\mathbf{r}) = P(\mathbf{r} - \mathbf{R}_j), \tag{3.7}$$

and then evolves recursively according to

$$\psi_j^{(n+1)}(\mathbf{r}) = \mathcal{P}_{\Delta z}\left[O_n(\mathbf{r})\,\psi_j^{(n)}(\mathbf{r})\right], \tag{3.8}$$

where $O_n(\mathbf{r})$ denotes the transmission function of the n-th slice and $\mathcal{P}_{\Delta z}$ denotes propagation over the inter-slice spacing Δz[3]. Writing the slice transmission as $O_n(\mathbf{r}) = A_n(\mathbf{r})e^{i\phi_n(\mathbf{r})}$ already makes clear which physically meaningful quantities enter the forward chain: slice-wise amplitude and phase govern local interaction, whereas Δz enters through the propagation block itself rather than through specimen content alone.

After the final slice, the resulting wave is propagated to the detector plane and converted into an intensity prediction as in Section 3.2. The forward model now becomes

[3]Appendix A gives the reciprocal-space form of $\mathcal{P}_{\Delta z}$ and the short derivational bridge from a paraxial electron-wave equation to the multislice recursion. When propagation-geometry corrections such as tilt are introduced, the same operator should be understood more generally as carrying additional geometric dependence rather than as a purely formal label.

an ordered sequence of transmission and propagation steps through the specimen thickness.

This internal propagation is the decisive difference from the single-slice picture. In a projected model, propagation enters only when the exit wave is mapped to the detector plane. In the multislice model it acts between slices and therefore couples depth to scattering: what happens at one slice modifies the incident field for the next. The measured diffraction pattern thus depends on the entire operator chain rather than on one isolated transmission event. A single-slice model forced onto genuinely thick-specimen data would risk absorbing thickness effects into an unphysical effective transmission, thereby introducing model mismatch into the reconstruction [20, 23].

This point also matters in Chapter 4. Once the multislice block is embedded in a differentiable forward model, slice amplitude, slice phase, slice spacing Δz, and any included geometric corrections all become variables that change the predicted diffraction data. Therefore, multislice is both the physically appropriate model for thick specimens and the reason depth-resolved parameters later enter the inverse problem explicitly [16].

3.4 Mixed-State Probe and Object Modes

The multislice formulation developed in the previous section resolves depth-dependent scattering, but it still retains an idealisation: namely, that the measured diffraction data arise from a single coherent probe interacting with a single effective object state. In practice, residual probe instability, partial coherence of the illumination, and other sources of state variability can violate that pure-state picture. A more flexible forward model therefore allows the measured intensity to arise from an incoherent mixture of modes rather than from one perfectly coherent probe–object configuration [24, 25].

Let $P_m(\mathbf{r})$ denote probe modes and $O_n(\mathbf{r})$ denote object modes. The scan-indexed wavefield generated by the pair (m, n) is written $\psi_j^{(m,n)}$, where, when multislice propagation is enabled, $\psi_j^{(m,n)}$ should be understood as the output of the multislice operator chain introduced in Section 3.3 rather than of a single multiplicative interaction. The recorded detector intensity is then written as an occupancy-weighted incoherent sum,

$$I_j(\mathbf{q}) = \sum_{m=1}^{M} \sum_{n=1}^{N} \omega_{mn} \left| \mathcal{F}\{\psi_j^{(m,n)}(\mathbf{r})\}(\mathbf{q}) \right|^2, \tag{3.9}$$

with non-negative weights

$$\omega_{mn} \geq 0, \qquad \sum_{m=1}^{M} \sum_{n=1}^{N} \omega_{mn} = 1. \tag{3.10}$$

Eq. (3.9) makes the physical assumption explicit: the detector records an incoherent sum of intensities, not a coherent sum of amplitudes. The weights ω_{mn} therefore encode modal participation in the effective measurement model rather than serving as arbitrary fitting coefficients.[4]

Figure 5-5 provides a qualitative illustration of this idea. The redistribution of probe-mode amplitudes during reconstruction shows that the forward model need not remain locked to one fixed coherent illumination profile; instead, modal structure can adapt to absorb partially coherent or unstable contributions. Object modes are better treated as an additional modelling layer that allows the reconstruction to represent sample-side variability or unresolved structured mismatch, without implying that the specimen literally occupies a small set of discrete physical states [24].

Mixed-state modelling increases flexibility, but it also enlarges the number of admissible explanations for the same data. Introducing additional modes can improve numerical fit, but it can also produce *mode leakage*, whereby one mode absorbs structure better explained elsewhere in the model. For PtyRAD this is not a cosmetic extension: mixed-state modelling enlarges the tensor structure of the forward model and foreshadows the mode-related regularisation, parameter coupling, and identifiability issues analysed in Chapter 4 [25].

3.5 Detector Blur and Measurement Formation

The preceding sections have progressively enriched the forward model at the level of the specimen and illumination: multislice propagation accounts for depth-dependent multiple scattering, while mixed-state modes allow the model to represent partial coherence and mode variability. Even so, the intensity predicted at this stage is still an idealised detector-plane quantity. In experiment, the detector does not record this ideal intensity

[4] In the current PtyRAD implementation, mode participation is handled more conservatively than the main probe and object fields themselves. This distinction becomes important again in Chapter 4 when trainable variables are separated from forward-participating but fixed quantities.

field directly. Instead, the measured diffraction pattern is further shaped by detector-side response mechanisms such as point-spread effects, charge sharing, and finite pixel integration. The final measurement should therefore not be identified with the ideal intensity generated by the specimen alone [26].

A minimal way to express this distinction is to model the detector response through a blur kernel acting on the ideal predicted intensity. Writing the ideal detector-plane intensity from the previous sections as $I_j^{\mathrm{ideal}}(\mathbf{q})$, the measured intensity may be approximated by

$$I_j^{\mathrm{meas}}(\mathbf{q}) \approx \left(h * I_j^{\mathrm{ideal}}\right)(\mathbf{q}), \tag{3.11}$$

where h denotes the detector point-spread function and $*$ denotes convolution.

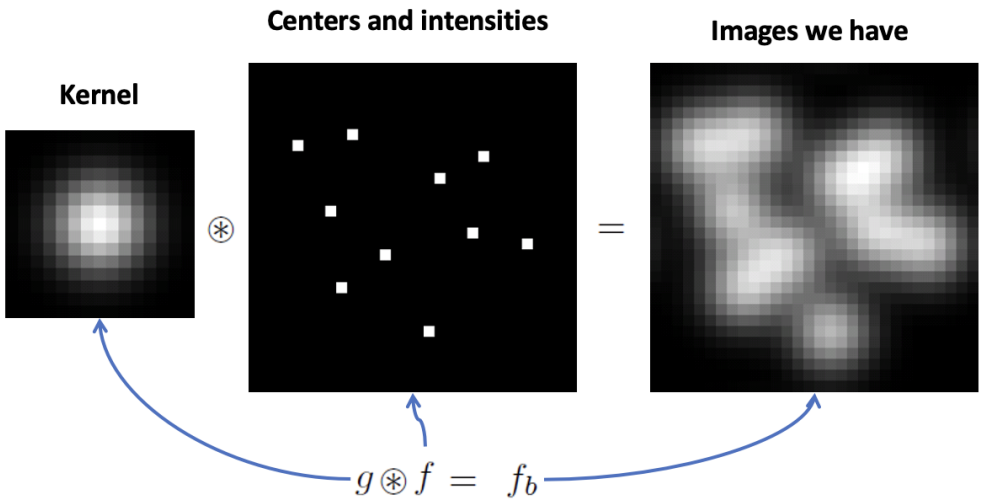


Figure 3-4. Schematic illustration of detector blur as convolution. A detector kernel h acts on an idealised intensity distribution, producing a spatially blurred measured image. This visualises the measurement model in Eq. (3.11), where detector response is represented by convolution with a point-spread function.

In the simplest treatment adopted here, h may be approximated by a Gaussian kernel, so that detector blur enters as a spatial smoothing of the ideal diffraction intensity. More detailed detector characterisations are often expressed in terms of point-spread functions or modulation transfer functions, but the central modelling idea is the same: the detector modifies the measurement after the specimen-generated wavefield has already been converted into intensity [27].

This distinction is physically important. Detector blur belongs to the measurement

model. It should be applied before the simulated data are compared with experiment, otherwise detector response can be misattributed to probe or object structure, forcing the inverse problem to explain instrumental effects in terms of specimen physics. In this sense, detector blur is therefore part of measurement formation, instead of a post-processing step.

This interpretation is reflected directly in PtyRAD. After the forward model generates the predicted diffraction intensity, detector blur is applied within the forward pass itself rather than being treated as an external image-processing step. This is a conceptually important modelling choice: the reconstruction is asked to match what the detector actually measures, not an idealised intensity that would only exist before detector response.

At the same time, detector blur is only one part of measurement realism. Real detectors also introduce noise and other statistical effects, but those belong more naturally to the discussion of objective functions and data fidelity in Chapter 4. For the present chapter, the essential point is that the forward model should terminate not at ideal intensity formation, but at a physically motivated prediction of the measured detector signal. This completes the progression from specimen physics to measurement formation and leads naturally to the next section, where these physical ingredients are recast in tensor form and aligned with the computational structure used in PtyRAD.

3.6 Tensor Structure and Computational Representation

In implementation, the physical ingredients introduced so far must be represented in discretised numerical form. This discretisation is part of the model specification, because the numerical representation must preserve the structure of the underlying wave-optics operators, where fields, coordinates, and propagators are encoded computationally must remain aligned with [16].

At the most basic level, the model lives on sampled real-space and reciprocal-space grids. Probe and object are therefore manipulated not as continuous fields but as arrays whose entries represent sampled values on those grids. This discretisation is what allows the physical sequence of shifting, multiplication, propagation, Fourier transformation, modulus-square formation, and summation to be implemented as tensor operations.

The important point is that these arrays remain organised by physical role: object

amplitude and phase, probe, scan-dependent shifts, and slice-dependent propagation are kept explicit rather than collapsed into one undifferentiated numerical state. In this way, the computational representation preserves the same distinctions already introduced at the physical level: local specimen transmission, local probe illumination, scan-dependent pairing, and depth propagation.

The tensor structure also makes visible how physical realism enlarges the computational problem. Measurements are naturally indexed by scan position and detector pixel, multislice propagation introduces a slice axis, and mixed-state modelling introduces probe-mode and object-mode axes. As a result, the dimensionality of the tensors reflects the physical richness of the model. This is not cosmetic: memory usage, batching strategy, and feasible update schemes are all affected by how many physical axes are retained simultaneously.

This representation prepares the optimisation formulation in Chapter 4 by keeping the physically meaningful quantities explicit in the computation. Once object amplitude and phase, probe, scan corrections, and propagation-related quantities survive as explicit tensor blocks rather than being absorbed into one opaque numerical state, they can later be promoted to optimisation variables inside the differentiable forward chain. The tensor representation should therefore be understood as part of the modelling argument of this thesis rather than as a purely software-engineering detail.

3.7 From Physical Model to Computational Graph

Chapter 3 has progressively assembled the forward model of PtyRAD from physically meaningful ingredients to increasingly realistic measurement formation. Probe and object define the local interaction, scan geometry organises that interaction across measurements, multislice propagation introduces depth-dependent scattering, mixed-state modes account for incoherent variability, and detector blur refines the final measurement prediction. Section 3.6 then showed how these ingredients are encoded computationally as tensors. Collectively, these components define a forward map from physically structured variables to predicted diffraction data [16].

This structure may be expressed schematically as

$$\theta \;\mapsto\; I^{\mathrm{model}}(\theta), \tag{3.12}$$

where θ denotes the collection of variables entering the forward prediction, including probe and object representations, scan geometry, multislice parameters, mixed-state quantities, and detector-response settings where applicable. The notation emphasises that the forward model is a composition of operators that maps a structured parameter set to predicted measurements.

Figure 3-5 presents this structure as a reconstruction workflow: the model variables are encoded as tensors, propagated through the forward simulation, compared with the measurements through the loss, and differentiated backward to obtain updates.

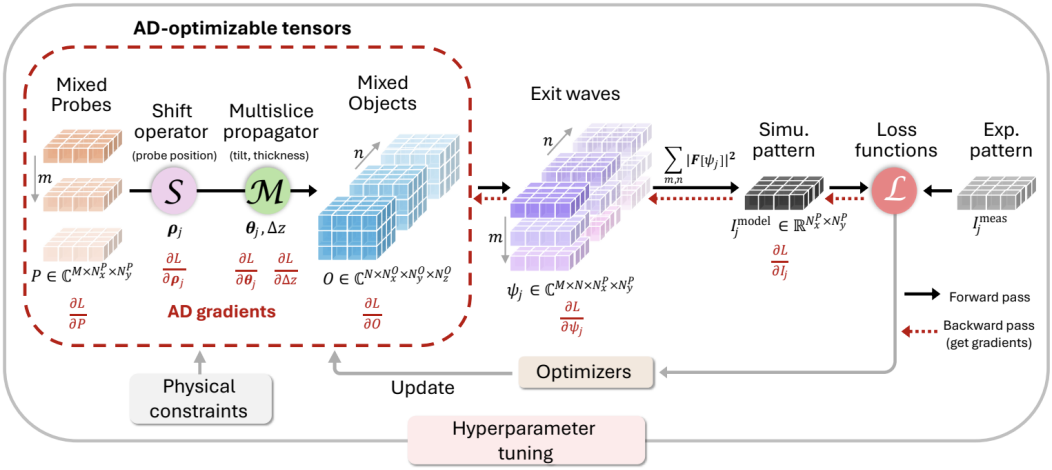


Figure 3-5. Overall computational graph of PtyRAD. Physically meaningful variables such as mixed probes, scan-position shifts, multislice propagation parameters, and mixed objects are represented as differentiable tensors and propagated through the forward model to produce predicted diffraction intensities. These are then compared with experimental measurements through a loss function, after which gradients are propagated backward for optimisation. This diagram makes explicit how the physical forward model is converted into an automatic-differentiation-based reconstruction pipeline, and for hyperparameter see Appendix B.

Chapter 4 does not introduce a new class of variables; it asks which of the quantities already present in the forward model are treated as trainable. Once a physically meaningful quantity participates in the construction of $I^{\mathrm{model}}(\theta)$, it becomes, in principle, a candidate for gradient-based update. That is why later chapters can treat object amplitude and phase, probe, scan corrections, and multislice propagation parameters as optimisation variables rather than as external metadata. In practice, the computational graph is the ordered set of tensor operations by which object amplitude and phase, probe, scan corrections, tilt, and propagation parameters generate predicted diffraction

intensities and therefore receive gradients during backpropagation.

In automatic differentiation frameworks, such a forward model is implemented as a composition of elementary tensor operations, and derivatives with respect to trainable variables are obtained through the chain rule applied to this composition [28, 29]. The present thesis does not require a detailed treatment of automatic differentiation internals. What matters here is the conceptual consequence: once the forward model has been organised as a differentiable[5] operator chain, reconstruction naturally becomes the task of finding parameter values for which the predicted measurements are consistent with the experimental data.

This completes the role of Chapter 3. Beginning from physically motivated ingredients, the chapter has shown how the forward model of PtyRAD is built, refined, and finally organised into a computational form suitable for inference. Chapter 4 can therefore begin from this point and recast reconstruction explicitly as a nonconvex inverse problem, in which the variables entering θ are treated as optimisation targets and the agreement between $I^{\mathrm{model}}(\theta)$ and measured data is formalised through objective functions, regularisation, and gradient-based optimisation.

[5]Here, "differentiable" means that the predicted diffraction intensities are computed through tensor operations for which gradients with respect to selected model variables can be obtained automatically.

Chapter 4

Optimisation Formulation of PtyRAD

4.1 From Physical Variables to Trainable Parameters

Chapter 3 established the forward model of PtyRAD as a composition of physically meaningful operations, beginning from probe–object interaction and scan geometry, and proceeding through propagation, intensity formation, mixed-state summation, and measurement formation. Once the forward model is implemented in a differentiable form, any variable that changes the predicted diffraction data can in principle be estimated from the data [15, 28, 29]. This statement is the natural continuation of the map

$$\theta \mapsto I^{\mathrm{model}}(\theta), \tag{4.1}$$

introduced at the end of Chapter 3. The forward model is no longer only a physical description of how diffraction patterns are generated; it also becomes the operator through which candidate parameter values are judged against the measurements.

This principle should, however, be stated carefully. Entering the forward model is a necessary condition for estimation, but it does not guarantee identifiability, numerical stability, or reliable recovery from finite noisy data. For the present chapter, the trainable parameters are organised as

$$\Theta_{\mathrm{train}} = \Theta_{\mathrm{obj}} \cup \Theta_{\mathrm{illum}} \cup \Theta_{\mathrm{geom/sys}}, \tag{4.2}$$

with

$$\Theta_{\mathrm{obj}} = \{A, \phi\}, \qquad \Theta_{\mathrm{illum}} = \{P\}, \qquad \Theta_{\mathrm{geom/sys}} = \{\delta\mathbf{r},\ \boldsymbol{\tau},\ \Delta z\}. \tag{4.3}$$

Here A and ϕ denote object amplitude and phase, P denotes the probe, $\delta\mathbf{r}$ denotes scan-position correction variables, $\boldsymbol{\tau}$ denotes object-tilt variables, and Δz denotes the slice-thickness parameter.[1]

[1] In the current PtyRAD snapshot examined in this thesis, the principal trainable tensors correspond precisely to object amplitude, object phase, probe, probe-position shifts, object tilts, and slice thickness

The first block, Θ_{obj}, is associated with the specimen content. In the forward model the object enters as a complex transmission function,

$$O(\mathbf{r}) = A(\mathbf{r}) \exp\big(i\phi(\mathbf{r})\big). \tag{4.4}$$

Amplitude and phase are treated separately because they correspond to different physical channels through which the specimen modifies the incident wave.

The second block, Θ_{illum}, contains the probe, which is estimated jointly with the object and a structured complex wavefield whose shape, phase, and mode content directly influence the recorded diffraction patterns [24, 25]. For this reason it is natural to treat the probe as its own optimisation block. Doing so also reflects a central feature of ptychographic inference: object and probe are typically inferred jointly, because each measurement is generated by their local interaction rather than by either factor in isolation [10, 14].

The third block, $\Theta_{\text{geom/sys}}$, gathers variables that modify how the physical interaction is embedded into the forward operator[16, 20, 23]:

(a) $\delta\mathbf{r}$: corrected scan positions alter which local object region is paired with the probe at each measurement index;

(b) $\boldsymbol{\tau}$: tilt variables modify the effective propagation geometry;

(c) Δz: slice thickness enters the multislice propagator itself, thereby changing the way the wavefield evolves through depth.

These quantities do not describe specimen content as A and ϕ, yet they are not external metadata either. Because they influence the generated diffraction patterns through the same differentiable operator chain, they can also be posed as estimation targets.[2]

Not every quantity that participates in the forward model must, however, be included in Θ_{train} for every reconstruction. It is therefore useful to introduce a second collection,

$$\Xi_{\text{forward}} = \{\text{quantities that remain fixed or externally prescribed}\} \tag{4.5}$$

[16, 30]. This implementation-level narrowing is a design choice of the reconstruction framework, rather than a logical limit of the modelling language itself.

[2]For notational clarity, the main text writes the scan geometry in terms of corrected positions, e.g. $\mathbf{R}_j + \delta\mathbf{r}_j$. In the current implementation this is realised more specifically as integer patch extraction together with a sub-pixel probe-shift correction [30]. The abstract notation is retained here because it is better suited to the modelling discussion.

This set may include, depending on the reconstruction setup, mode occupancies, detector-response settings, or measurement-side configuration choices that shape the forward prediction without being actively updated. The distinction between Θ_{train} and Ξ_{forward} is conceptually important. It makes clear that the differentiable forward model defines a space of possible optimisation variables that is generally broader than the specific subset chosen for a given reconstruction schedule.

With these definitions in place, the forward model can be written more explicitly as a map from trainable parameter blocks and fixed forward settings to predicted diffraction data,

$$\left\{I_j^{\text{model}}\right\}_{j=1}^{N} = \mathsf{F}(\Theta_{\text{train}};\, \Xi_{\text{forward}})\,, \tag{4.6}$$

where N denotes the number of scan positions and F abbreviates the full operator chain from local probe–object interaction to recorded detector-plane intensity.

Once physically meaningful variables have been organised into trainable blocks, reconstruction can be recast as the task of estimating those blocks from measured diffraction data.

4.2 Reconstruction as a Physics-Constrained[3] Nonconvex Optimisation Problem

With the parameter blocks of Eqs. (4.2)–(4.6) now specified, reconstruction can be formulated as an estimation problem. Let the measured 4D-STEM dataset be denoted by

$$\mathcal{D} = \left\{I_j^{\text{meas}}\right\}_{j=1}^{N}, \tag{4.7}$$

and let the corresponding model prediction at scan index j be written as

$$I_j^{\text{model}} = \mathsf{F}_j(\Theta_{\text{train}};\, \Xi_{\text{forward}})\,, \qquad j = 1, \ldots, N. \tag{4.8}$$

Under this notation, reconstruction is naturally formulated as the search for parameter values whose forward-generated diffraction patterns agree, in a suitably defined sense,

[3] Here, physics-constrained means that admissible iterates are shaped jointly by the physical forward operator, the chosen loss terms, and explicit post-step enforcement such as recentering, thresholding,smoothing, or mode-structure control.

with the measured dataset [14–16].

One formulation is obtained by introducing a penalised objective,

$$\mathcal{J}(\Theta_{\text{train}}) = \mathcal{L}_{\text{data}}\left(\left\{I_j^{\text{model}}(\Theta_{\text{train}}; \Xi_{\text{forward}})\right\}_{j=1}^{N}, \left\{I_j^{\text{meas}}\right\}_{j=1}^{N}\right) + \sum_q \lambda_q R_q(\Theta_{\text{train}}), \quad (4.9)$$

where $\mathcal{L}_{\text{data}}$ measures agreement between predicted and measured diffraction data, R_q denotes regularisation terms, and λ_q denotes their associated weights. Reconstruction then returns an estimate

$$\hat{\Theta}_{\text{train}} \in \arg\min_{\Theta_{\text{train}}} \mathcal{J}(\Theta_{\text{train}}). \quad (4.10)$$

The notation in Eq. (4.9) is kept deliberately generic at this stage. The physical forward model alone does not determine what should count as a good fit; that role is played by the choice of data-fidelity term and regularisation, whose statistical motivation and computational consequences will be examined in Section 4.3.

Under this formulation, reconstruction can be viewed as a nonconvex optimisation problem. The nonconvexity is not an incidental numerical inconvenience, but a structural consequence of how the measurements are generated:

(a) First, the exit wave depends multiplicatively on the object and probe, so the inverse problem is already coupled at the level of local wavefield formation [10, 14].

(b) Second, the detector records intensities rather than complex fields, introducing the modulus-square nonlinearity between latent wavefield and measurement.

(c) Third, multislice propagation makes the predicted diffraction depend recursively on depth propagation, so parameters such as tilt and slice thickness influence the data indirectly through the propagation operator rather than through a simple additive term [20, 23].

(d) Finally, mixed-state or multi-mode formulations add further latent structure through incoherent summation, increasing modelling flexibility but also enlarging the space of ambiguities and couplings [24, 25].

A more detailed discussion of these couplings and their optimisation consequences is deferred to Section 4.5.

Eq. (4.10) does not by itself describe the whole reconstruction procedure. In practice, the search trajectory is also shaped by post-step constraints and staged parameter

schedules. In abstract form, one may write iterative update as

$$\Theta_{\text{train}}^{(k+1)} = \Pi_k \left[\text{OptStep}\left(\Theta_{\text{train}}^{(k)}, \nabla \mathcal{J}\left(\Theta_{\text{train}}^{(k)} \right) \right) \right], \tag{4.11}$$

where $\text{OptStep}(\cdot)$ denotes the chosen optimiser step and $\Pi_k(\cdot)$ denotes iteration-dependent constraint or projection-like operations.[4] This distinction is important for the argument of the chapter: the inverse problem is defined by the forward model and the objective, but the actual search trajectory is additionally shaped by physically motivated enforcement steps and reconstruction schedules.

4.3 Loss Functions, Regularisation, and Constraint Enforcement

Equation (4.9) introduced the reconstruction problem in generic penalised form. The next question is how the notion of "fit" is instantiated in PtyRAD itself. A useful starting point is to separate the role of the forward map from the role of the objective. The purpose of F is to generate a physically meaningful prediction from the current parameter estimate, whereas the loss only specifies how disagreement between prediction and measurement is quantified [15, 16].

In the current PtyRAD framework, the objective can be written schematically as

$$\mathcal{J}_{\text{PtyRAD}}(\Theta_{\text{train}}) = w_{\text{G}}\,\mathcal{L}_{\text{G}} + w_{\text{P}}\,\mathcal{L}_{\text{P}} + w_{\text{C}}\,\mathcal{L}_{\text{PACBED}} + \lambda_{\text{s}} R_{\text{s}} + \lambda_{\text{m}} R_{\text{m}}, \tag{4.12}$$

where $\mathcal{L}_{\text{G}}$, $\mathcal{L}_{\text{P}}$, and $\mathcal{L}_{\text{PACBED}}$ denote data-fidelity terms, R_{s} and R_{m} denote regularisation terms, and the associated weights w_{G}, w_{P}, w_{C}, λ_{s}, and λ_{m} control their relative influence [16, 30]. This expression should be understood as an implementation-aware instantiation of Eq. (4.9), not as a claim that all reconstructions must use exactly this combination. The active terms do not ask for the same kind of agreement: `loss_single` penalises scan-resolved detector mismatch, `loss_pacbed` penalises the position-averaged diffraction statistic, and `loss_sparse` penalises object-side phase structure.

[4] In the current PtyRAD implementation, such post-update operations can include recentering of position corrections, smoothing of tilt fields, positivity-like clipping or thresholding of object channels, mode-related orthogonalisation or sorting, and other filtering or stabilisation steps imposed outside the main loss evaluation [30].

4.3.1 Data fidelity terms

PtyRAD compares the reconstruction with the measurements in several ways, because different loss terms emphasise different detector statistics and different levels of aggregation [14, 16, 31]. To keep the notation unambiguous, let

$$I_{j,k}^{\text{model}},\ I_{j,k}^{\text{meas}} \in \mathbb{R}_+ \tag{4.13}$$

denote the detector intensity at scan index j and detector pixel k. Let $\langle\cdot\rangle_{j,\mathbf{k}}$ denote averaging over scan index j and detector pixel $\mathbf{k}$, and let $\langle\cdot\rangle_{\mathbf{k}}$ denote averaging over detector pixels only. Introducing the power-transform notation

$$\mathcal{T}_p(I) = I^p, \tag{4.14}$$

For PACBED-type quantities, position averaging is written as

$$\bar{I}_k^{\text{model}} = \frac{1}{N}\sum_{j=1}^{N} I_{j,k}^{\text{model}}, \qquad \bar{I}_k^{\text{meas}} = \frac{1}{N}\sum_{j=1}^{N} I_{j,k}^{\text{meas}}. \tag{4.15}$$

With this convention, the three data-fidelity terms can be presented separately.

Gaussian-like pattern fidelity. A clean abstraction of the Gaussian-like term is

$$\mathcal{L}_{\text{G}} = \frac{\left\langle \left| \mathcal{T}_{p_{\text{G}}}\left(I_{j,k}^{\text{model}}\right) - \mathcal{T}_{p_{\text{G}}}\left(I_{j,k}^{\text{meas}}\right) \right|^2 \right\rangle_{j,k}^{1/2}}{\left\langle \mathcal{T}_{p_{\text{G}}}\left(I_{j,k}^{\text{meas}}\right) \right\rangle_{j,k}}. \tag{4.16}$$

Its statistical motivation is least-squares matching on the comparison variable selected by $\mathcal{T}_{p_{\text{G}}}$. If $\mathcal{T}_{p_{\text{G}}}$ is the identity, Eq. (4.16) reduces to a normalised intensity-domain least-squares or RMSE-type objective. In the implemented form, this term is best read as transformed-domain least-squares matching rather than as an exact Gaussian likelihood on raw detector counts[14]. A fuller connection to least-squares and transformed-domain Gaussian modelling is given in Appendix C.2.

Poisson-like pattern fidelity. The Poisson-like term may be written as

$$\mathcal{L}_{\mathrm{P}} = -\frac{\left\langle \mathcal{T}_{p_{\mathrm{P}}}\left(I_{j,k}^{\mathrm{meas}}\right) \log\left(\mathcal{T}_{p_{\mathrm{P}}}\left(I_{j,k}^{\mathrm{model}}\right) + \varepsilon\right) - \mathcal{T}_{p_{\mathrm{P}}}\left(I_{j,k}^{\mathrm{model}}\right)\right\rangle_{j,k}}{\left\langle \mathcal{T}_{p_{\mathrm{P}}}\left(I_{j,k}^{\mathrm{meas}}\right)\right\rangle_{j,k}}. \tag{4.17}$$

This term is motivated by Poisson counting statistics and is therefore most naturally associated with low-dose or counting-limited regimes [31]. If the transform is removed, that is, if $\mathcal{T}_{p_{\mathrm{P}}}$ is the identity, and if the extra normalisation is ignored, Eq. (4.17) reduces to the familiar negative Poisson log-likelihood up to additive constants that do not affect optimisation[5]. In the implemented form, with power transforms and numerical stabilisation included, this term is better understood as a Poisson-inspired fitting term than as an exact count model.

PACBED as an auxiliary calibration term. The third data term is not a per-pattern likelihood but a position-averaged agreement measure,

$$\mathcal{L}_{\mathrm{PACBED}} = \frac{\left\langle \left| \mathcal{T}_{p_{\mathrm{C}}}\left(\bar{I}_{k}^{\mathrm{model}}\right) - \mathcal{T}_{p_{\mathrm{C}}}\left(\bar{I}_{k}^{\mathrm{meas}}\right)\right|^2 \right\rangle_{k}^{1/2}}{\left\langle \mathcal{T}_{p_{\mathrm{C}}}\left(\bar{I}_{k}^{\mathrm{meas}}\right)\right\rangle_{k}}. \tag{4.18}$$

Here the averaging is over scan positions j only, so $\bar{I}_k$ remains a detector-plane pattern indexed by k. Physically, this is precisely the PACBED idea: local scan dependence is collapsed, while the global diffraction distribution is retained as a summary statistic. This makes the term especially useful as an auxiliary calibration term for variables such as thickness, tilt, or other system-sensitive quantities whose influence may be diffuse or strongly coupled when only per-pattern mismatch is considered [6, 16, 32].

4.3.2 Regularisation terms

In addition to data fidelity, PtyRAD introduces terms that act directly on the admissible structure of the reconstruction. These should be described as regularisation terms in the objective, although some of them can also be interpreted in prior-like language when

[5] More details see Appendix C.3

read from a statistical viewpoint [16, 33]. A clean abstraction of two main regularisers is

$$R_{\mathrm{s}} = \sum_{m=1}^{M_{\mathrm{o}}} \alpha_m \left(\langle |\phi_m|^{q_{\mathrm{s}}} \rangle_{\mathrm{ROI}} \right)^{1/q_{\mathrm{s}}}, \tag{4.19a}$$

$$R_{\mathrm{m}} = \sum_{\chi \in C} \left\langle \mathrm{Std}_m \left[\mathcal{B}_\chi \left(X_m^{(\chi)} \right) \right] \right\rangle_{\mathbf{r}}, \tag{4.19b}$$

where ϕ_m denotes the phase of object mode m, q_{s} denotes the chosen norm order, α_m denotes fixed mode-weighting coefficients, $X_m^{(\chi)}$ denotes the selected object channel (amplitude, phase, or both), and $\mathcal{B}_\chi$ denotes an optional blur/downsampling operator used to compare modes at a coarser scale.

The sparsity term R_{s} of Eq. (4.19a) is the more classical regulariser. It penalises large or widespread phase excursions over the reconstructed region of interest, thereby discouraging high-frequency or noise-like fluctuations that are weakly supported by the data [16, 30]. From a statistical perspective, such a penalty can be read in MAP-like terms as the imposition of a sparsity-favouring prior on the chosen representation [33]. That interpretation is useful, but it should not be overstated. The benefit of sparsity is regime- and object-dependent: it can stabilise low-dose reconstructions, yet it can also bias extended or genuinely diffuse structures if imposed too strongly [33].

The similarity term R_{m} of Eq. (4.19b) plays a different role. Rather than promoting sparsity in a single channel, it imposes coarse-scale structural consistency across mixed-state object modes. In the present implementation this is done after optional blurring and downsampling, so that the penalty acts primarily on large-scale inter-mode divergence rather than on fine pixelwise differences [30]. This is why the term is best understood as a mode-regularising stabiliser: it discourages spurious splitting of the reconstruction across modes while still allowing mixed-state modelling to absorb physically meaningful incoherent variability [24, 25, 34]. As with sparsity, the interpretation should remain conditional. Enforcing similarity across modes can reduce ambiguity and improve stability, but it also encodes a structural hypothesis about what kinds of mode variation are acceptable.

The regularisation terms in Eq. (4.19) do more than penalise large values. They encode which kinds of object and mode structure are preferred when the data do not determine a unique solution.

4.3.3 Constraint enforcement beyond the objective

Regularisation terms should not be conflated with all other physically motivated operations used during reconstruction. Equation (4.11) already indicated that the optimisation trajectory may be modified by an iteration-dependent operator $\Pi_k(\cdot)$ in addition to gradient-based minimisation of $\mathcal{J}_{\text{PtyRAD}}$. In PtyRAD, this distinction is not cosmetic: some operations are added to the loss, whereas others act directly on the iterates after an optimiser step [16, 30]. The latter are best described as constraint-enforcement mechanisms or projection-like operations, rather than as ordinary penalty terms.

At an abstract level, the principal categories are:

(a) **recentring operations**, which remove accumulated global offsets or redistribute structure into a more physically interpretable reference frame;

(b) **smoothing or filtering operations**, which suppress implausible high-frequency variation in selected parameter blocks;

(c) **clipping or thresholding operations**, which impose positivity-like or boundedness conditions on amplitude- or phase-related channels; and

(d) **mode-structure operations**, which enforce structured behaviour across probe or object modes, for example through orthogonalisation-like actions or mode-order control.

This distinction between regularisation and constraint enforcement matters for two reasons. First, it keeps the formulation honest: not every physically motivated stabilisation step is most naturally written as an additive term in the objective. Second, it prepares the discussion of optimisation dynamics. Penalty terms change the scalar objective and therefore the gradient field, whereas projection-like operations can reshape the search trajectory even when they do not correspond to the gradient of any single smooth functional [16, 30, 34]. The reconstruction process is therefore better understood as the joint action of a weighted objective and an algorithmic enforcement scheme.

4.4 Physical Interpretation of Gradients

Section 4.3 specified how agreement between predicted and measured diffraction data is quantified in PtyRAD. Once that objective is fixed, the next question is no longer what should count as mismatch, but how that mismatch is propagated backward through the

forward operator F so that each parameter block receives an update direction. The key point is that the same operator chain that generates the prediction in the forward pass also determines how sensitivity is transported in the backward pass [15, 16, 29]. Let

$$\mathcal{R} := \lambda_{\mathrm{s}} R_{\mathrm{s}} + \lambda_{\mathrm{m}} R_{\mathrm{m}}, \tag{4.20}$$

denote the regularisation part of Eq. (4.12). The gradient of the implemented objective may then be written abstractly as

$$\nabla_{\Theta_{\mathrm{train}}} \mathcal{J}_{\mathrm{PtyRAD}} = \left(\frac{\partial \mathsf{F}}{\partial \Theta_{\mathrm{train}}} \right)^{*} \nabla_{I^{\mathrm{model}}} \mathcal{L}_{\mathrm{data}} + \nabla_{\Theta_{\mathrm{train}}} \mathcal{R}, \tag{4.21}$$

where the first term represents the adjoint action of the linearised forward map on the data-mismatch signal, and the second term collects the direct contributions from regularisation. At implementation level, these sensitivities are obtained through automatic differentiation across complex-valued wavefield operations under a real-valued objective.[6]

A useful starting point is therefore the detector-plane sensitivity induced by the data term,

$$G_j := \frac{\partial \mathcal{L}_{\mathrm{data}}}{\partial I_j^{\mathrm{model}}}, \qquad j = 1, \ldots, N. \tag{4.22}$$

The detector-plane gradient is not simply the raw intensity residual. Its form depends on which data term is active: Gaussian-like, Poisson-like, and PACBED-oriented terms all induce different detector-plane sensitivities. In schematic form, the sensitivity signal is propagated through the reverse of the forward chain:

(a) detector-plane intensity sensitivity;

(b) measurement formation and, where relevant, mode aggregation;

(c) Fourier/intensity mapping between wavefield and measured pattern;

(d) multislice propagation through depth; and

(e) local probe–object interaction together with geometric corrections.

Backpropagation therefore has a physical interpretation: it identifies which stage of the forward model a change in the objective is attributed to [20, 23].

The first genuinely local interpretation appears at probe–object interaction. At a

[6] For PyTorch, this is framed in terms of conjugate-Wirtinger-style gradients for complex tensors under real-valued losses [28].

given scan index, a small perturbation of the exit wave may be written schematically as

$$\delta\psi_j = (\delta O_j) \odot P_j + O_j \odot (\delta P_j), \tag{4.23}$$

where O_j denotes the relevant local object transmission and P_j denotes the corresponding probe patch. Equation (4.23) makes the central ptychographic coupling explicit: the same local mismatch can, in principle, be reduced either by changing the object or by changing the probe [10, 14]. This is why probe and object gradients should not be interpreted as independent channels. Rather, they are two projections of the same backpropagated sensitivity onto two multiplicatively interacting factors. For the object block, the amplitude–phase parameterisation introduced earlier gives

$$\delta O_j = \exp(i\phi_j) \odot \delta A_j + iO_j \odot \delta\phi_j. \tag{4.24}$$

This decomposition shows why amplitude and phase gradients carry distinct physical meanings. The amplitude gradient indicates where the transmission magnitude should be strengthened or weakened in order to reduce mismatch, whereas the phase gradient indicates how phase retardation should be redistributed. The probe gradient, by contrast, tells us how the illumination state itself should be reshaped so that the full set of local interactions across scan positions becomes more consistent with the measurements [16, 24, 25].

A second class of gradients concerns geometric and propagation-related variables. These do not act by directly altering specimen content, but by changing how wavefields are paired, shifted, or propagated. In abstract form, their sensitivities may be written as

$$\frac{\partial \mathcal{J}_{\text{PtyRAD}}}{\partial \delta \mathbf{r}_j} = \left\langle \nabla_{\psi_j} \mathcal{J}_{\text{PtyRAD}}, \frac{\partial \psi_j}{\partial \delta \mathbf{r}_j} \right\rangle \tag{4.25a}$$

$$\frac{\partial \mathcal{J}_{\text{PtyRAD}}}{\partial \boldsymbol{\tau}} = \left\langle \nabla_{\psi} \mathcal{J}_{\text{PtyRAD}}, \frac{\partial \psi}{\partial \boldsymbol{\tau}} \right\rangle \tag{4.25b}$$

$$\frac{\partial \mathcal{J}_{\text{PtyRAD}}}{\partial \Delta z} = \left\langle \nabla_{\psi} \mathcal{J}_{\text{PtyRAD}}, \frac{\partial \psi}{\partial \Delta z} \right\rangle \tag{4.25c}$$

where the brackets indicate the appropriate inner-product contraction over the relevant wavefield degrees of freedom. Although Eq. (4.25) is schematic, its physical content is clear. The position gradient indicates whether the local probe–object alignment should

shift. The tilt gradient indicates how the effective propagation geometry should be corrected. The slice-thickness gradient indicates how the wavefield evolution through depth should be modified, precisely because thickness enters the multislice propagator rather than the specimen transmission alone [20, 23, 30].

The physical meanings of the principal gradient blocks can therefore be summarised as follows:

(a) **Amplitude gradients** regulate how strongly the specimen attenuates or transmits the incident wave;

(b) **phase gradients** regulate how phase retardation should be redistributed across the reconstructed field;

(c) **probe gradients** regulate how the illumination state should be reshaped so that all local interactions become jointly consistent with the data;

(d) **position gradients** regulate geometric alignment between probe and object patches;

(e) **tilt gradients** regulate the effective propagation geometry experienced by the wavefield;

(f) **thickness gradients** regulate the depth-evolution operator in multislice propagation; and

(g) **detector/system gradients**, when such variables are estimated, regulate how measurement formation itself should be adjusted.

A nonzero gradient indicates local sensitivity, not necessarily stable identifiability or reliable recoverability. In ptychography, the same mismatch can often be redistributed across multiplicatively coupled variables, weakly constrained geometric corrections, or propagation-related parameters whose effects are indirect and highly correlated. The gradients therefore tell us *how* the objective responds locally, but not yet whether those responses define a well-conditioned or trustworthy search direction.

4.5 Parameter Coupling, Ill-Conditioning, and Plateau

Section 4.4 showed that each trainable block in PtyRAD receives a physically meaningful gradient. This, however, does not yet guarantee easy reconstruction. A nonzero gradient shows that the objective is locally sensitive to a parameter; it does not show that the parameter can be recovered cleanly or independently.

(a) **parameter coupling**, in which distinct parameter blocks produce partially overlapping or compensating changes in the predicted data;

(b) **ill-conditioning**, in which locally informative directions exist but are very unequally scaled or highly correlated; and

(c) **plateau behaviour**, in which optimisation progress becomes slow or apparently stagnant, often as a consequence of the first two effects rather than as a primitive property of the optimiser itself.

These effects do not imply that recovery is impossible, but they do imply that convergence behaviour cannot be inferred from gradient existence alone.

A convenient local language for this section is provided by the Jacobian of the predicted measurements with respect to the trainable parameter set. Around a current iterate, one may write the first-order perturbation model

$$\delta I^{\mathrm{model}} \approx J_\Theta \, \delta\Theta, \qquad J_\Theta = \left[J_A, J_\phi, J_P, J_{\delta\mathbf{r}}, J_{\boldsymbol{\tau}}, J_{\Delta z} \right], \tag{4.26}$$

where the block structure reflects the present implementation, in which object amplitude, object phase, probe, position shifts, tilts, and slice thickness are updated within a joint automatic-differentiation loop [16, 30]. Equation (4.26) is a local sensitivity model that makes block interactions explicit.

The first difficulty is coupling. If two parameter blocks u and v both influence the prediction, then the induced perturbation can be written schematically as

$$\delta I^{\mathrm{model}} \approx J_u \, \delta u + J_v \, \delta v, \qquad J_u \, \delta u + J_v \, \delta v \approx 0, \tag{4.27}$$

for directions in which the two blocks compensate each other. In ptychography, the most immediate example is the bilinear probe–object interaction: a change in the local probe can, to some extent, be traded against a change in the local object transmission while producing a similar effect on the exit wave and hence on the measured intensity [10, 16]. This does not mean that probe and object are interchangeable, but it does mean that their updates are not naturally independent.

A second source of coupling arises from propagation-related variables. In the present framework, tilt and slice thickness act through the propagator rather than by directly editing specimen content. Their effect is therefore mediated by how the wavefield

evolves between slices, which makes them especially susceptible to partial trade-offs with object contrast or with one another when depth information transfer is limited [20, 23, 30]. Position variables introduce a different kind of entanglement: they alter geometric registration and the sampling of local probe–object interaction, so residual mismatch may be reduced either by adjusting the scan geometry or by distorting content-related variables to absorb the misalignment [35, 36]. Coupling is further strengthened by structural ambiguities. Global phase offsets, scale-like trade-offs, and mixed-state rotations can generate equivalent or near-equivalent intensity predictions, so some shallow directions reflect non-uniqueness rather than mere numerical accident [24, 37].

Coupling becomes especially harmful when it is accompanied by poor conditioning. A useful local object here is the Gauss–Newton-like normal matrix

$$G_\Theta := J_\Theta^* W J_\Theta, \tag{4.28}$$

where W denotes the weighting induced by the chosen mismatch model. The off-diagonal blocks of G_Θ quantify sensitivity overlap between parameter groups, while its scale variation reflects how unevenly different directions are informed by the data. In ptychography, reduced overlap lowers data redundancy, geometric uncertainty weakens the reliability of registration-related corrections, and depth-related variables may be only weakly excited by the measurements. The resulting problem is not simply that updates are coupled, but that the informative directions can become strongly anisotropic and badly scaled [9, 20, 23, 36].

This point can be expressed in the local quadratic model

$$\mathcal{J}_{\mathrm{PtyRAD}}(\Theta + \delta\Theta) \approx \mathcal{J}_{\mathrm{PtyRAD}}(\Theta) + \Re\langle g, \delta\Theta\rangle + \frac{1}{2}\,\delta\Theta^* G_\Theta\,\delta\Theta, \tag{4.29}$$

where $g = \nabla_\Theta \mathcal{J}_{\mathrm{PtyRAD}}$. In this language, ill-conditioning means that some directions are steep and well informed whereas others are flat, weak, or highly correlated. Step sizes that are safe for one block may therefore be ineffective for another, and progress can become dominated by the best-conditioned modes of variation rather than by the parameters one would most like to recover [38, 39].

From this viewpoint, plateau is best understood as a symptom of weakly informed

or compensating directions than as a primitive property of the optimiser.[7] In thick or strongly dynamical regimes, for example, limited depth transfer and modelling approximations can leave a residual floor that the algorithm keeps redistributing across object, probe, and geometry-related blocks without finding a decisive descent direction [16, 23]. Plateau therefore should not be described as a vague "flat landscape" in the abstract; it is more useful to ask which parameter directions are weakly informed, which blocks are compensating, and which part of the mismatch may in fact arise from model inadequacy rather than recoverable parameter error.

4.6 From Gradients to Updates: Optimisation Methods for PtyRAD

Sections 4.4 and 4.5 have already established two points that are central for the remainder of the chapter. First, the present forward model yields gradients with respect to object amplitude and phase, probe, scan shifts, tilt, and slice thickness in forms that can be interpreted through their roles in the measurement chain. Second, the mere existence of those gradients does not guarantee that the resulting updates will be independent, well scaled, or rapidly convergent. The natural next question is therefore not how sensitivities are computed, but how they are converted into actual parameter updates. In the current PtyRAD framework these are distinct stages of the reconstruction loop: gradients are obtained by backpropagating through F, update directions are produced by an optimiser, and the iterates may then be further modified by post-step constraint enforcement [15, 16, 30].

4.6.1 First-order adaptive methods as the practical baseline

Adaptive first-order methods serve as the default baseline because they use only gradients, scale to large parameter sets, and fit naturally into an automatic-differentiation reconstruction loop [40, 41]. A compact expression of the Adam update may be written

[7] In this sense, a small $\|\nabla \mathcal{J}_{\mathrm{PtyRAD}}\|$ or a small update norm $\|\Theta^{k+1} - \Theta^{k}\|$ does not, by itself, imply that the reconstruction has reached a globally satisfactory explanation of the data. Weak or compensating sensitivities can produce the same numerical symptom.

as

$$m_{k+1} = \beta_1 m_k + (1-\beta_1) g_k, \qquad v_{k+1} = \beta_2 v_k + (1-\beta_2)(g_k \odot g_k), \tag{4.30a}$$

$$\widehat{m}_{k+1} = \frac{m_{k+1}}{1-\beta_1^{k+1}}, \qquad \widehat{v}_{k+1} = \frac{v_{k+1}}{1-\beta_2^{k+1}}, \tag{4.30b}$$

$$\Theta_{k+1} = \Theta_k - \alpha \frac{\widehat{m}_{k+1}}{\sqrt{\widehat{v}_{k+1}} + \varepsilon}, \tag{4.30c}$$

where $g_k = \nabla_\Theta \mathcal{J}_{\mathrm{PtyRAD}}(\Theta_k)$, m_k and v_k denote first- and second-moment estimates, β_1 and β_2 are decay parameters, and ε is a stabilising constant. In this form the method combines directional information from the current gradient with a coordinate-wise estimate of gradient scale, so that strongly and weakly varying components can be normalised before the update is applied [40]. This combination is particularly convenient when a single optimisation loop must accommodate object-related, probe-related, and geometry-related parameter blocks of very different numerical scales.

In the context of PtyRAD, this practical appeal is strengthened by implementation simplicity[8]. The reconstruction loop is already organised around automatic differentiation and explicit optimiser calls, so adaptive first-order updates can be inserted without rewriting the forward model or deriving bespoke analytic derivative expressions [15, 30].

4.6.2 Why first-order methods can stall in PtyRAD

The advantages of adaptive first-order methods should not be confused with immunity to slow convergence. In PtyRAD, late-stage stagnation is better understood as the interaction between a first-order update rule and the structural properties analysed in Section 4.5. When the inverse problem is strongly coupled, anisotropic, and only partially informative in some directions, coordinate-wise gradient scaling does not by itself produce a trustworthy local search model. A first-order method can stabilise updates, but it cannot create missing curvature information, disentangle compensating parameter blocks, or remove residual floors caused by model mismatch [38, 39].

This point is especially important for the present thesis, because plateau should

[8] Present implementation does not rely on a bare optimiser alone: parameter blocks are activated and deactivated through iteration-dependent `requires_grad` control, meaning that the practical baseline is better described as *adaptive first-order optimisation within a staged schedule* rather than as unconstrained Adam in isolation.

not be described as a mysterious defect of Adam alone. Rather, it is the observable consequence of using local gradient information in a problem where some directions are weakly informed, some parameter blocks can partially compensate one another, and some portions of the residual may arise from imperfect model fit rather than from recoverable parameter error [16, 30]. In that sense, adaptive first-order methods remain a sensible baseline, but their late-stage behaviour is limited by the same block-dependent sensitivity structure that was exposed in Sections 4.4 and 4.5.

4.6.3 Curvature refinement: Gauss–Newton / Levenberg–Marquardt and L-BFGS

Once first-order stagnation is understood as a limitation of local gradient-only updates, the next question is whether curvature-aware refinement can make better use of the local landscape. Building on the local quadratic picture of Eq. (4.29), a smooth curvature-aware step can be written schematically as

$$\left(H_k + \lambda_k I\right) d_k = -g_k, \qquad H_k \approx J_\Theta^* W J_\Theta, \tag{4.31}$$

where $g_k = \nabla_\Theta \mathcal{J}_{\mathrm{PtyRAD}}(\Theta_k)$, $J_\Theta^* W J_\Theta$ denotes a Gauss–Newton-like normal matrix, and $\lambda_k \geq 0$ is a damping parameter. This expression captures the logic behind Gauss–Newton and Levenberg–Marquardt type methods: rather than scaling gradients coordinate-wise, they attempt to use a local quadratic model of how residuals respond to parameter perturbations [38, 39]. For ptychographic reconstruction this is especially natural because the objective is built from residual-like mismatch terms, so the Gauss–Newton approximation is often more tightly connected to the problem structure than a generic Newton step.

That said, curvature-aware refinement is not automatically superior. Its usefulness depends on whether the local model is informative, whether the active parameter block is sufficiently well behaved, and whether the additional computational burden can be justified. In the ptychographic setting, matrix-free Levenberg–Marquardt is particularly attractive as a conceptual reference because it respects the residual structure and can exploit automatic differentiation without explicitly assembling full matrices [38]. This makes Gauss–Newton / LM a natural smooth refinement idealisation for the thesis, even though it is not the method directly implemented in the current PtyRAD code.

Within the present implementation, L-BFGS is the available curvature-aware

refinement method. In generic form, its search direction may be written as

$$d_k = -\widetilde{H}_k g_k, \tag{4.32}$$

where $\widetilde{H}_k$ denotes an inverse-Hessian surrogate built from a limited history of step and gradient differences [39]. Compared with Gauss–Newton or Levenberg–Marquardt, L-BFGS is less explicitly tied to residual structure, but it is often attractive as a practical curvature-aware refinement because it avoids forming or storing full Hessian approximations. In the present framework this relevance is immediate: PtyRAD contains an explicit L-BFGS branch, implemented through closure-based updates, and therefore treats L-BFGS as the concrete quasi-Newton option available within the current reconstruction loop [30].

4.6.4 Beyond smooth methods

A final method family worth mentioning is semismooth Newton. Its relevance here is real, but conditional. Semismooth Newton methods are most natural when the optimisation problem is reformulated so that nonsmooth or constrained structure is made explicit, for example through a nonsmooth equation, a KKT system, or a composite objective with an explicit nonsmooth block [42, 43]. Under the current PtyRAD loop, many physically motivated restrictions are still handled through post-step enforcement rather than through such an explicit nonsmooth formulation. For that reason, semismooth Newton should not be presented as a drop-in replacement for Adam or L-BFGS in the present code path. It is better understood as a reformulation-dependent direction that may become meaningful only if positivity-like constraints, clipping rules, sparsity-promoting terms, or other structured blocks are promoted from heuristic enforcement to explicit nonsmooth subproblems.

Overall, the optimiser landscape in PtyRAD is layered. Adaptive first-order methods provide the default update mechanism. Curvature-aware refinements become useful only when the local model is informative and the added computational cost is acceptable. Semismooth methods require a stronger nonsmooth reformulation than the current reconstruction loop provides.

Chapter 5

Baseline Reconstruction and Experimental Framework

This chapter establishes the experimental baseline and the corresponding evidential framework for the $SrTiO_3$ case study developed in this thesis. It documents the dataset, preprocessing choices, and baseline reconstruction configuration, and then introduces the diagnostic criteria used to evaluate optimisation behaviour, forward-model consistency, and reconstructed object quality. On this basis, the chapter analyses the principal behaviours and limitations exhibited by the baseline reconstruction and formulates the working hypotheses that motivate the controlled studies in Chapter 6.

5.1 Dataset and Preprocessing

This section records the experimental dataset used in baseline study and the measurement-side operations applied before reconstruction. It therefore defines the input to the reconstruction pipeline, while the baseline configuration itself is specified in Section 5.2.

5.1.1 Experimental dataset

The experimental case study used in Chapters 5 and 6 is a 4D-STEM dataset acquired from $SrTiO_3$ (STO), a cubic perovskite oxide. At the acquisition level, the raw dataset consists of a 200×200 scan grid, with one 128×128 diffraction pattern recorded at each probe position, giving a total of 40,000 scan-indexed measurements. The principal acquisition parameters, including the acceleration voltage, probe semi-convergence angle, and scan step size, are summarised later together with the baseline reconstruction settings in Table 5-1. For crystallographic reference, Fig. 5-1 shows schematic unit-cell and projected views of $SrTiO_3$.

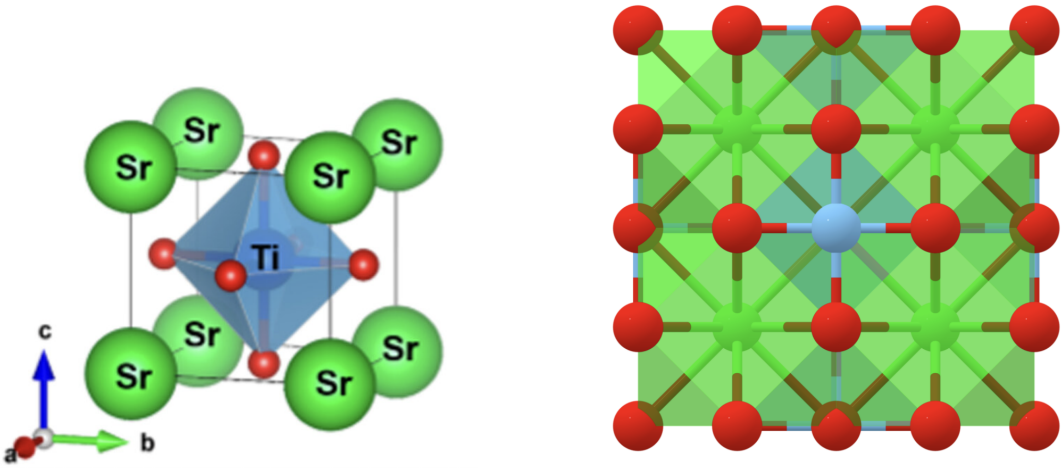


Figure 5-1. Crystallographic reference views of $SrTiO_3$ used to contextualise the baseline dataset. The left panel shows a schematic unit-cell representation of the cubic perovskite structure, while the right panel shows a projected structural view.

5.1.2 Preprocessing and calibration choices

The diffraction data were loaded from the original `.raw` measurement file as a stack of 40,000 patterns of size 128×128. Before reconstruction, negative measurement values were clipped to zero, and the measurements were normalised by the maximum intensity of the mean diffraction pattern. No additional measurement-side source-size blur, detector blur, or synthetic Poisson noise was introduced at this stage.

The measurements were further resampled on the detector plane by a factor of 0.5×0.5 in an on-the-fly manner. Consequently, the raw acquisition and the effective solver input are not identical: although the loaded measurement stack remained at 128×128, the effective detector sampling used by the solver corresponded to $N_{\text{pix}} = 64$ after resampling. Detector-space calibration was handled through RBF-based initialisation pathway in PtyRAD and updated consistently under this resampling step.

5.2 Defining the Practical Reduced Baseline

Starting from the preprocessed measurements defined in Section 5.1, the baseline reconstruction analysed in this chapter is implemented in PtyRAD with the configuration summarised in Table 5-1.

Several features of Table 5-1 should be noted.

Table 5-1. Practical reduced baseline configuration used for the $SrTiO_3$ case study.

Category	Baseline setting
Dataset / acquisition	$SrTiO_3$ 4D-STEM dataset; raw acquisition 200×200 scan positions, 128×128 detector pixels; 300 kV acceleration voltage; 28.9 mrad semi-convergence angle; 0.592 Å scan step size.
Meas preprocessing	Raw `.raw` measurement source; negative values clipped to zero; intensity normalised by `max_at_one`; detector-plane resampling `[0.5, 0.5]` applied in `on_the_fly` mode; no additional measurement-side detector blur, source-size blur, or synthetic Poisson noise.
Effective data exposure	`INDICES_MODE = center` with default half-size sub-scan, giving an effective 100×100 central scan subset; `GROUP_MODE = random`; batch size = 100 with `grad_accumulation` = 1, so the effective batch size per update is also 100.
Forward model	Multislice forward model with 32 slices and fixed slice thickness Δz = 10 Å; four probe modes; single object mode; simulated initial probe; simulated scan positions with affine correction and small random displacement; global object-tilt parameter initialised at zero; forward-model detector blur with `detector_blur_std = 1`.
Optimiser and update schedule	Adam optimiser; configured iteration budget $N_{\text{iter}} = 200$; save interval of 50 iterations. Object amplitude and phase are trainable from iteration 1, probe updates begin at iteration 5, and object-tilt together with probe-position-shift updates begin at iteration 10. Slice-thickness optimisation is disabled.
Active objective terms	`loss_single` (weight 1.0, `dp_pow`=0.5), `loss_pacbed` (weight 0.5, `dp_pow`=0.2), and `loss_sparse` (weight 0.1, L_1-type phase regularisation).
Iter-wise constraints	Probe-mode orthogonalisation; probe-intensity fixing; lateral object blur; depth-direction object blur; mirrored-amplitude constraint; object-amplitude thresholding; object-phase positivity; position recentering.

(a) First, the present reconstruction does not expose the optimiser to the full raw acquisition. The active scan set is restricted to the central 100×100 region, while detector-plane measurements are further reduced by on-the-fly resampling.

(b) Second, the forward model is already nontrivial: multislice propagation, mixed-state probe modelling, forward-model detector blur, and a global tilt variable are all included from the outset.

(c) Third, the optimisation is not performed with all variables released simultaneously; instead, the object is updated first, probe updates are enabled later, and geometric corrections are introduced only after a further delay.

(d) Finally, the baseline objective is accompanied by persistent iteration-wise constraints that directly regularise probe, object, and positional behaviour throughout the run.

Accordingly, the configuration analysed in this chapter is best understood as a practical reduced baseline. It is sufficiently rich to produce diagnostically meaningful behaviour on the experimental $SrTiO_3$ dataset, yet sufficiently constrained to serve as a stable reference configuration for the baseline analysis of Chapter 5 and the controlled comparisons developed later in Chapter 6.

5.3 Evaluation Framework and Metrics

For the present $SrTiO_3$ baseline, reconstruction quality is not assessed through a single scalar score. Because the dataset is experimental rather than simulated, an exact ground-truth object is not available for direct comparison. The baseline is therefore evaluated through a layered framework of evidence, designed to separate detector-side agreement, optimisation behaviour, and object-side interpretability. In this chapter, that framework is organised into three complementary groups: data-consistency signals, optimisation and physical diagnostics, and image-side and volumetric proxies.

5.3.1 Data-consistency signals

The first layer of evaluation concerns agreement between the reconstructed forward model and the measured diffraction data. At the scalar level, this includes the total objective together with the active loss components, so that detector-side fitting terms

and regularisation terms are not hidden inside a single aggregated curve. Under the present baseline, the active objective consists of `loss_single`, `loss_pacbed`, and `loss_sparse`.

The total loss is useful for identifying coarse optimisation regimes, such as rapid initial descent and later slowing-down, but it is not sufficient by itself. In particular, the total objective mixes detector-space agreement with regularising pressure on the reconstructed object. For this reason, the active loss components are reported separately wherever possible. Their relative behaviour is later used to determine whether improvement is dominated by single-pattern fitting, by position-averaged diffraction consistency, or by object-side regularisation.

Scalar losses are complemented by representative forward-pass comparisons between simulated and measured diffraction patterns. These comparisons are treated here primarily as data-consistency evidence, since they reveal where the model already matches the measured diffraction morphology and where structured mismatch remains. At the same time, they also prepare later optimisation diagnosis, because persistent detector-side residuals may indicate limitations in the objective, the measurement model, or the allocation of freedom between specimen-side and probe-side variables.

5.3.2 Optimisation and physical diagnostics

The second layer concerns how the optimisation proceeds and which physically interpretable parameter blocks appear to be active under the baseline. This includes the qualitative rate of loss decrease, the presence or absence of later-stage slowing-down, and runtime quantities such as average iteration time. These quantities are not direct measures of reconstruction quality, but they are necessary for understanding how the baseline behaves as an optimisation process.

Several parameter-side diagnostics are also retained because they provide physically meaningful evidence about behaviour of reconstruction. In the present study, these include probe modal occupancies, the magnitude of scan-position correction, and average magnitude and direction of object-tilt updates. Probe-mode redistribution is particularly important because the current baseline already includes mixed-state probe modelling; substantial redistribution may indicate either useful partial-coherence modelling or the absorption of residual mismatch into probe-side freedom. Likewise, scan-position

refinement and object-tilt correction help indicate whether the baseline is relying primarily on geometry-side correction or more strongly on object/probe adaptation.

5.3.3 Image-side and volumetric proxies

The third layer concerns the reconstructed object itself. In the absence of exact ground truth, these quantities are treated as diagnostic proxies rather than absolute image-quality metrics. At the 2D level, the depth-summed object phase serves as the primary visual summary of the baseline reconstruction. It is assessed through lattice visibility, local contrast, and the cleanliness or otherwise of the background and boundary regions.

Because the present baseline is multislice rather than single-slice, image-side judgement cannot be based on the depth-summed view alone. Representative depth slices and the corresponding z-stack are therefore inspected as complementary volumetric evidence. Their role is deliberately limited: they are used to check whether the reconstruction forms a nontrivial and qualitatively plausible depth-dependent volume, rather than to claim strong depth resolution in a strict quantitative sense.

This separation between 2D summary and volumetric sanity check is important for the present thesis. A visually readable z-sum is necessary, but it is not by itself sufficient to establish that the multislice reconstruction behaves plausibly along depth. Conversely, a slice-resolved volume may be numerically nontrivial while remaining difficult to interpret physically if the lateral projection remains excessively noisy or artefactual. For this reason, both projected and slice-resolved object outputs are retained within the evaluation framework.

Table 5-2 summarises layered criteria used throughout the remainder of Chapter 5.

5.4 Standard Diagnostic Figures

The figures collected below define the fixed visual evidence base for the remainder of this chapter. They do not serve equivalent roles. Some figures make the effective reconstruction problem explicit, such as the scan-subset and grouping diagnostic. Others record detector-side agreement or mechanism-level behaviour, including the loss trajectory, representative forward-pass comparisons, probe-mode evolution, and geometry-side corrections. Only the final object-output figure is intended primarily as a specimen-side representation. Unless otherwise stated, all figures in this section

Table 5-2. Layered evaluation framework used for the baseline $SrTiO_3$ reconstruction.

Layer	Quantity or diagnostic	What it indicates	Main figure anchor
Data Cons	Total loss	Overall behaviour of the active baseline objective; useful for identifying coarse optimisation phases, but not sufficient as a standalone quality claim.	Fig. 5-3
Data Cons	Active loss: (`loss_single`, `loss_pacbed`, `loss_sparse`)	Relative behaviour of detector-side fitting and regularisation terms; helps reveal objective-level tension hidden by the total-loss trace.	Fig. 5-3
Data Cons	Representative forward/data comparison	Spatially interpretable detector-side agreement; shows which diffraction morphology is already reproduced and where structured mismatch remains.	Fig. 5-4
Opt trajectory	Early descent / later slowing-down	Qualitative regime structure of the optimisation trajectory.	Fig. 5-3
Phy plausibility	Probe modal occupancies	Degree to which mixed-state probe freedom is being used during fitting; may indicate either beneficial modelling flexibility or mismatch absorption.	Fig. 5-5
Phy plausibility	Scan-position correction magnitude	Whether geometry-side correction is dominant active mechanism or secondary refinement.	Fig. 5-6
Phy plausibility	Average object-tilt magnitude / direction	Whether the baseline is learning a substantial or only mild global tilt correction.	Fig. 5-7
Img-side proxy	Depth-summed object phase (z-sum)	Primary 2D visual summary of reconstructed object; assessed through lattice visibility, contrast, and background cleanliness.	Fig. 5-8a
Vol proxy	Representative z-slices / z-stack	Complementary volume-side evidence for qualitatively plausible multislice behaviour; not treated as proof of depth resolution.	Fig. 5-8b
Effi context	Average iteration time / wall-clock	Computational cost context for baseline; retained for completeness and for later comparison with controlled interventions.	Table only / runtime log

correspond to the iteration-200 checkpoint of the present practical reduced baseline. Detailed interpretation is deferred to Section 5.5.

5.4.1 Scan subset and grouping diagnostics

Figure 5-2. Scan-subset and grouping diagnostic for the present baseline. The left panel shows the effective scan subset used in the reconstruction, while the right panel shows one representative group under `random` batching together with a 90% probe-intensity footprint marker.

Figure 5-2 makes the effective reconstruction problem explicit , and is included first because later optimisation behaviour can only be interpreted correctly once the active field of view and the single-update sampling geometry have been fixed.

5.4.2 Loss trajectory

Figure 5-3 provides the primary trajectory-level diagnostic. It records the total objective together with a magnified view of the late-stage iterations, so that early descent and later slowing-down can be distinguished within one plot. At this stage, the figure is used only to establish the qualitative structure of the optimisation trajectory; its interpretation is reserved for Section 5.5.

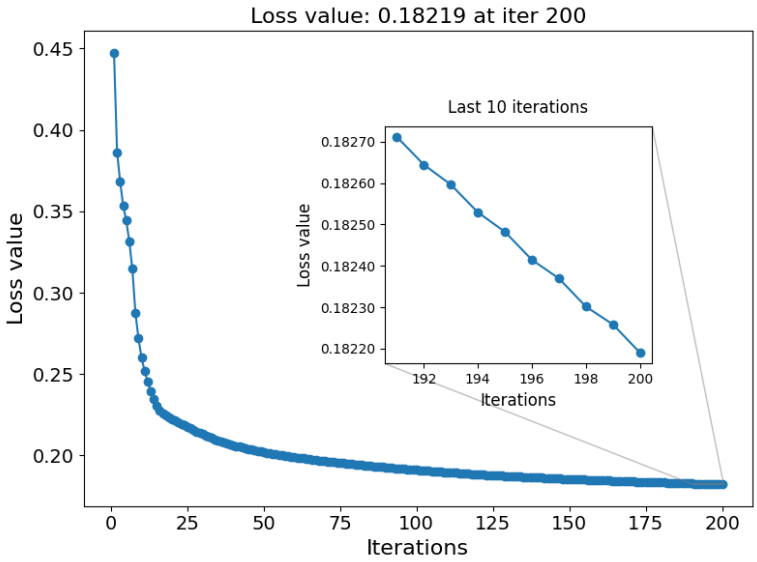


Figure 5-3. Loss trajectory of the baseline reconstruction at the iteration-200 checkpoint. The inset enlarges the final part of the run in order to record the late-stage behaviour more clearly.

5.4.3 Forward-pass comparison

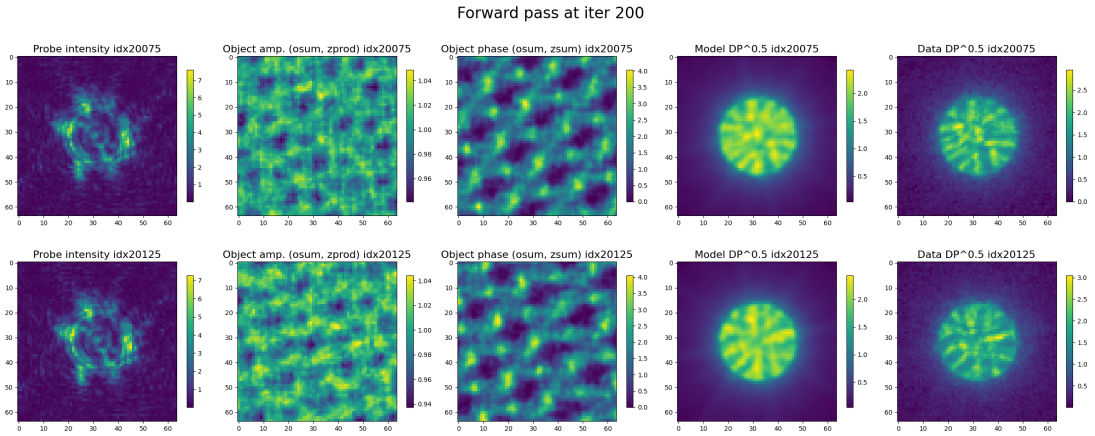


Figure 5-4. Representative forward-pass comparisons between simulated and measured diffraction patterns at two scan positions. For each representative position, the figure shows the probe intensity, local object amplitude and phase patches, the modelled diffraction pattern, and the corresponding measured diffraction pattern.

The forward-pass diagnostic is the main detector-side comparison figure in the baseline sequence. For each selected position, it records local probe and object context together with the modelled and measured diffraction patterns. Its purpose is to localise agreement

and residual mismatch in detector space, not to provide a direct specimen image.

5.4.4 Probe evolution and mixed-state behaviour

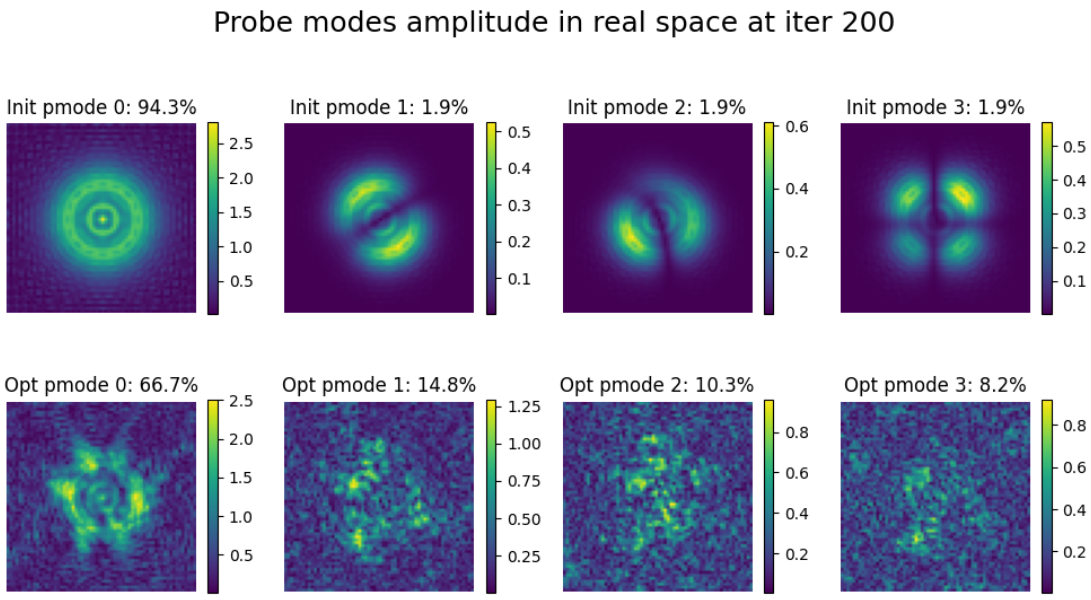


Figure 5-5. Real-space probe-mode amplitudes before and after optimisation at the iteration-200 checkpoint. The corresponding modal occupancies are shown above each mode.

Probe behaviour is monitored in the main text through the real-space amplitudes of the initial and optimised probe modes together with their reported modal occupancies. This figure is retained in the core diagnostic sequence because the present baseline already includes mixed-state probe modelling, so qualitative mode evolution is part of the evidential record rather than a secondary implementation detail.

5.4.5 Geometry diagnostics: scan positions and object tilt

Geometry-side behaviour is therefore recorded through two complementary diagnostics. Figure 5-6 overlays the initial and optimised scan coordinates in object space, while Fig. 5-7 records the behaviour of the learned object tilt through both its iteration history and its spatial visualisation across the active scan set. These figures are included to separate geometry correction from specimen-side output and to provide a fixed basis for later judging whether geometry variables act as dominant corrections or only as comparatively mild refinements.

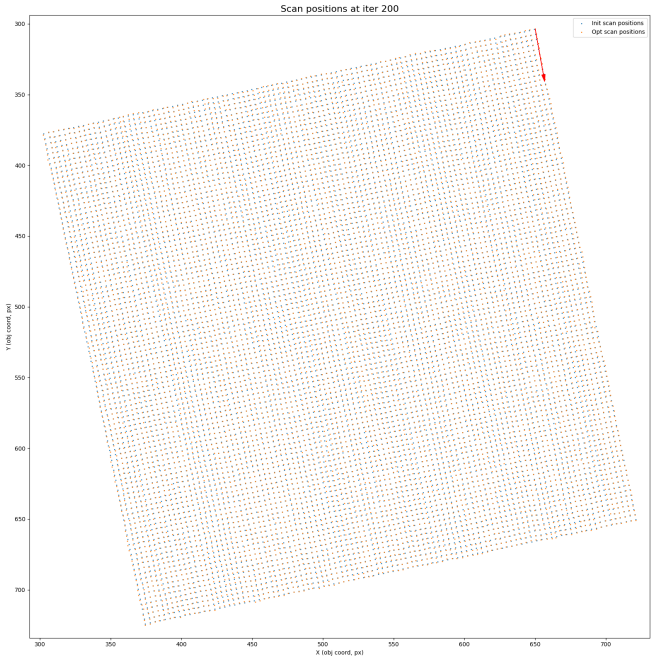
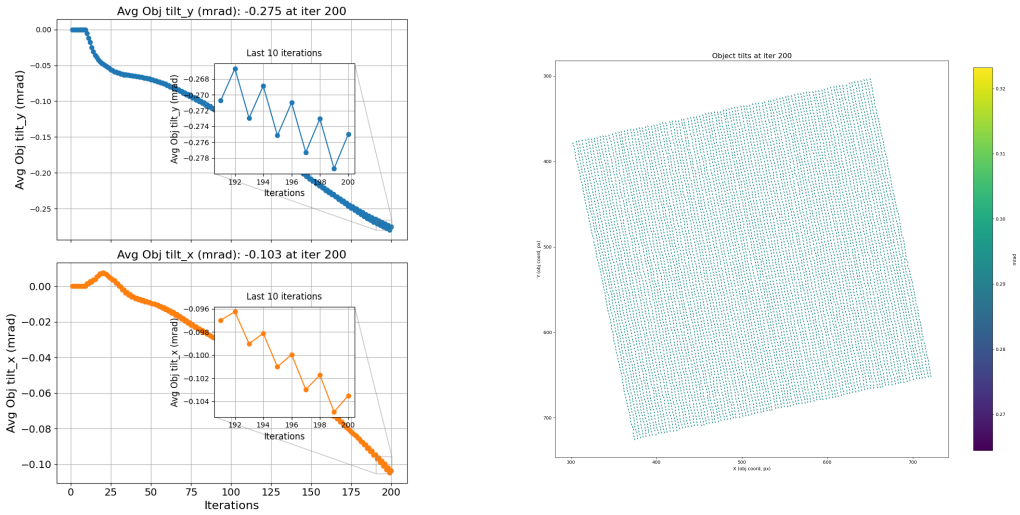
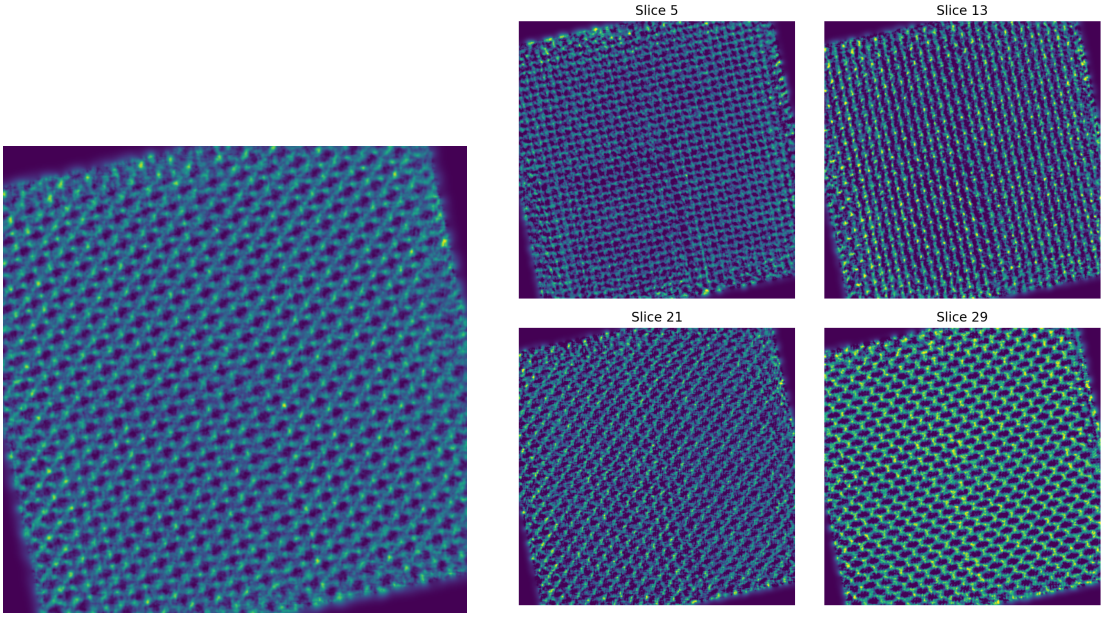


Figure 5-6. Initial and optimised scan positions are overlaid in object coordinates, and the arrow indicates the visual scale of the residual correction.



(a) Average tilt components versus iteration. **(b)** Spatial visualisation of the learned tilt field.

Figure 5-7. The left panel records the average tilt components as functions of iteration, while the right panel visualises the learned tilt field over the active scan set. Because the present baseline uses a global tilt parameter, the spatial field is uniform by construction.



(a) Cropped z-summed object phase.

(b) Representative depth slices extracted from the reconstructed volume.

Figure 5-8. The left panel provides the primary 2D summary of the reconstructed object, while the right panel shows representative depth slices for volumetric sanity checking.

5.4.6 Object outputs: z-sum and representative z-slices

This final diagnostic group is the only figure family intended primarily as a specimen-side output. The cropped z-summed phase serves as the main projected view used later for lateral image discussion, while the representative slices are included only as complementary volumetric evidence. Their role in this section is descriptive and diagnostic rather than conclusory.

5.5 Baseline Behaviour and Observed Limitations

This section interprets the standard diagnostics established in Section 5.4. The discussion therefore proceeds from trajectory-level behaviour to objective-level tension, then to the apparent allocation of modelling freedom across parameter blocks, and finally to the specimen-side outputs that the current baseline is able, and not yet able, to support.

5.5.1 From rapid early descent to later slowing-down

The baseline exhibits a clear two-regime optimisation trajectory. During the early stage, the total objective decreases rapidly, indicating that the reconstruction is able to capture coarse structure efficiently. By contrast, the later part of the trajectory is characterised by a much smaller but still persistent rate of improvement. In the present baseline, the total loss falls from 0.4472 at the first iteration to 0.2224 by iteration 20, and then continues to decrease more gradually to 0.1822 by iteration 200. The final ten iterations remain monotonically decreasing rather than fully flat. It is therefore more accurate to describe the late behaviour as clear slowing-down, or at most an incipient plateau, rather than as definitive optimisation failure.

This distinction matters for the interpretation of the baseline. The trajectory does not suggest that the reconstruction stops working after its early coarse-fitting stage. Rather, it suggests that once the dominant low-order structure has been fitted, the remaining error lies in directions that are weaker, more coupled, or less efficiently driven by the current optimisation setup. This observation already points toward later questions about objective design and late-stage refinement, but at the present stage it is retained simply as a diagnostic feature of the baseline itself.

5.5.2 Objective-level behaviour under the baseline

The trajectories of the active loss terms are not uniform across optimisation stages. The dominant contribution to the objective remains `loss_single`, which continues to decrease throughout the run. By contrast, `loss_pacbed` drops sharply during the early iterations and then rapidly approaches a near-stationary level, while `loss_sparse` increases gradually as optimisation proceeds.

The behaviour of `loss_pacbed` is particularly suggestive. Its rapid early decrease implies that it functions mainly as an early global-statistic term rather than as a late local-detail term. Once the broad position-averaged diffraction morphology is aligned, further improvement appears to be driven mainly by the single-pattern term. The gradual rise of `loss_sparse` points in the opposite direction: as detector-side fitting improves, the reconstructed object phase becomes less sparse under the present regularisation. This does not by itself imply that the sparsity term is inappropriate, but it does indicate a genuine objective-level tension within the current baseline.

Under the present baseline, improvement at later iterations is dominated by continued reduction of `loss_single`, whereas `loss_pacbed` has already become largely saturated and `loss_sparse` moves against the overall descent. This uneven behaviour is one of the clearest optimisation questions revealed by the baseline.

5.5.3 Allocation of modelling freedom: probe, position, and tilt

The baseline also reveals a non-uniform allocation of modelling freedom across parameter blocks. The most visible active mechanism is the mixed-state probe. By iteration 200, the dominant probe mode no longer carries nearly all of the intensity, and a substantial fraction of the modal weight has been redistributed into secondary modes.

That redistribution, however, is not uniquely interpretable. On the one hand, it is consistent with beneficial mixed-state modelling of partial coherence or illumination-side non-ideality. On the other hand, substantial probe-mode adaptation may also indicate that part of the residual mismatch is being absorbed into probe-side flexibility rather than being assigned uniquely to the specimen. The safest reading is therefore a dual one: mixed-state probe freedom is clearly useful under the present baseline, but it also remains a possible mismatch sink.

The geometry-side diagnostics are more differentiated. Scan-position refinement remains comparatively modest when viewed against the overall scan lattice, so position correction does not appear to be the dominant active mechanism in the present run. Global object tilt, by contrast, is no longer negligible by iteration 200. Because the baseline uses a single shared tilt vector rather than position-dependent local tilts, the learned tilt should be interpreted as a low-dimensional global geometric correction.

5.5.4 Forward–data mismatch and residual allocation

The forward-pass comparisons show that the baseline already reproduces the main bright-field morphology of the measured diffraction patterns at representative scan positions. The central disk size, the dominant angular segmentation, and the broad intensity distribution are all captured to a scientifically useful extent. This shows that the baseline has learned a forward model that captures the principal detector-space structure of the experiment, not only a lower scalar objective.

At the same time, the remaining residual is structured rather than random. Relative

to the experimental diffraction patterns, the simulated patterns remain smoother and more regular, especially in the diffuse background and peripheral regions. This suggests that the unresolved error cannot be understood simply as an overall shortage of optimisation iterations. Instead, it remains unclear which part of the remaining mismatch should be attributed to measurement formation, forward-model specification, or specimen-side representation.

5.5.5 Specimen-side readability and remaining limitations

On the specimen side, the baseline already yields a qualitatively readable but still diagnostically limited output. The depth-summed phase image recovers the lattice-scale periodic structure of the specimen clearly enough to support scientific inspection of the main projected morphology.

The projected output still contains visible graininess, non-uniform background texture, and crop-boundary artefacts. These features do not invalidate the reconstruction, but they show that the current baseline does not yet provide a fully clean projected representation. Their presence is also consistent with the slower late-stage regime described above, where improvement continues but becomes progressively less efficient.

The representative depth slices further indicate that the multislice volume is nontrivial rather than completely collapsed. Different slices emphasise different structural motifs and contrast patterns, which is sufficient for a basic multislice sanity check. However, the slice-wise outputs remain considerably harder to interpret than the depth-summed image. The present baseline therefore supports only limited volumetric interpretability: it demonstrates that a depth-dependent volume is being formed, but it does not by itself justify a strong claim of validated depth resolution.

5.5.6 Interim conclusion

Taken together, the baseline is best understood as a workable but diagnostically revealing reconstruction state. It already produces a readable specimen-side output and reproduces the principal detector-space morphology, yet it also exposes three concrete optimisation tensions: later-stage slowing-down after early coarse fitting, uneven behaviour among the active objective terms, and an unresolved allocation of mismatch between probe-side flexibility and structured detector-space residual.

5.6 Working Hypotheses for Chapter 6

The baseline analysis above does not by itself determine a unique optimisation strategy. It does, however, identify a small set of sufficiently concrete and recurrent behaviours from which controlled experimental questions can be formulated. The purpose of this section is therefore not to claim final causal explanations, but to convert the Chapter 5 observations into a concise set of working hypotheses that can be tested in Chapter 6 under fixed and comparable conditions.

5.6.1 H1: Objective terms contribute differently across optimisation stages

The first working hypothesis concerns the behaviour of the active loss terms under the baseline. As shown in Section 5.5, the optimisation trajectory is not characterised simply by uniform descent of a single objective, but by a clear change of regime: the early iterations reduce the objective rapidly, whereas later iterations enter a much slower improvement regime. At the same time, the trajectories of the active loss components are not uniform. The detector-side single-pattern term continues to decrease, the PACBED term stabilises comparatively early, and the sparsity term does not evolve in the same way as the dominant fitting term. This suggests that the present static objective may not provide equally useful optimisation signals across all reconstruction stages.

5.6.2 H2: Update scheduling may influence parameter coupling during fitting

The second hypothesis concerns parameter coupling. Even under the present baseline, the reconstruction is already staged rather than fully simultaneous: object amplitude and phase are active from the beginning, probe updates are enabled later, and geometry-related corrections are enabled later still. The baseline behaviour suggests that this scheduling is not incidental, since different parameter blocks do not appear to participate equally throughout the run. This leads to the working hypothesis that update schedule is not merely an implementation convenience, but a factor that may materially affect how strongly object, probe, and geometry variables compete or cooperate during fitting.

5.6.3 H3: Mixed-state probe freedom is useful, but its role remains ambiguous

The third hypothesis concerns the role of mixed-state probe modelling. The baseline clearly shows that probe-mode redistribution is substantial, which indicates that mixed-state probe freedom is not dormant but actively used by the reconstruction. This behaviour is consistent with the intended role of mixed-state modelling as a more flexible forward description of partial coherence or illumination instability. At the same time, strong modal redistribution does not by itself establish that all probe-side adaptation is physically desirable. As discussed earlier in the thesis, additional modes can also absorb structured residual mismatch that may not belong uniquely to the probe description.

5.6.4 H4: Batch geometry may alter how overlap information is used per update

The fourth hypothesis concerns batch geometry. Under the present baseline, the effective scan subset is fixed and the grouping mode is random. This provides a practical and stable reference configuration, but it also defines how the optimiser sees the data during each update. In ptychographic reconstruction, real-space overlap is a property of the acquisition itself, whereas grouping determines how much of that overlap structure is exposed jointly within a single optimisation step. This motivates the working hypothesis that different grouping geometries may alter the balance between local overlap exploitation and global statistical coverage, and may therefore change both optimisation behaviour and final reconstruction character.

5.6.5 H5: A limited second-stage refinement may be more effective than continuing unchanged first-order optimisation

The final hypothesis concerns late-stage refinement. The baseline does not terminate in a hard plateau, but it does show clear slowing-down and an incipient plateau-like regime. This suggests that simply extending the same first-order optimisation may not be the most effective way to use additional computation once coarse structure has already been established. The corresponding working hypothesis is that a limited second-stage refinement, introduced after the reconstruction has entered the slow regime, may be more effective than continuing with the unchanged first-order baseline alone.

Chapter 6

Controlled Optimisation Studies in PtyRAD

This chapter develops a small set of controlled optimisation studies motivated directly by the working hypotheses formulated at the end of Chapter 5.It is structured around a limited number of experimentally actionable hypothesis families: objective design, stage-dependent loss scheduling, parameter-release schedule, batch geometry, and late-stage refinement. A fifth hypothesis, concerning the ambiguity of mixed-state probe freedom, is retained as a cross-family interpretive axis rather than as an independent intervention family.

6.1 Controlled Study Design and Common Protocol

The shared reference throughout this chapter is the practical reduced baseline already established in Chapter 5: the central 100×100 scan subset extracted from the original 200×200 acquisition, on-the-fly detector-plane resampling by 0.5×0.5, `random` grouping with effective batch size 100, a 32-slice multislice forward model with four probe modes, staged Adam-based optimisation, and the active objective

$$\mathcal{J}_{\text{base}} = 1 \times \texttt{loss\_single} + 0.5 \times \texttt{loss\_pacbed} + 0.1 \times \texttt{loss\_sparse}. \tag{6.1}$$

Unless explicitly stated otherwise, all controlled studies reported in this chapter inherit that baseline and modify only one family of choices at a time.

6.2 Objective Design and Loss Form

This section tests the first working hypothesis of Chapter 5: the idea that the present static objective may not provide equally useful optimisation signals across all reconstruction stages. The aim is not to rank loss choices by total loss alone, but to identify which parts of the optimisation behaviour actually change when the objective is altered.

6.2.1 L_1-type versus L_2-like sparse regularisation

Between the two completed runs, the only intended objective-side change is the norm order used by `loss_sparse`: the baseline keeps the present L1-type choice, whereas the comparison run replaces it with an L2-like alternative.

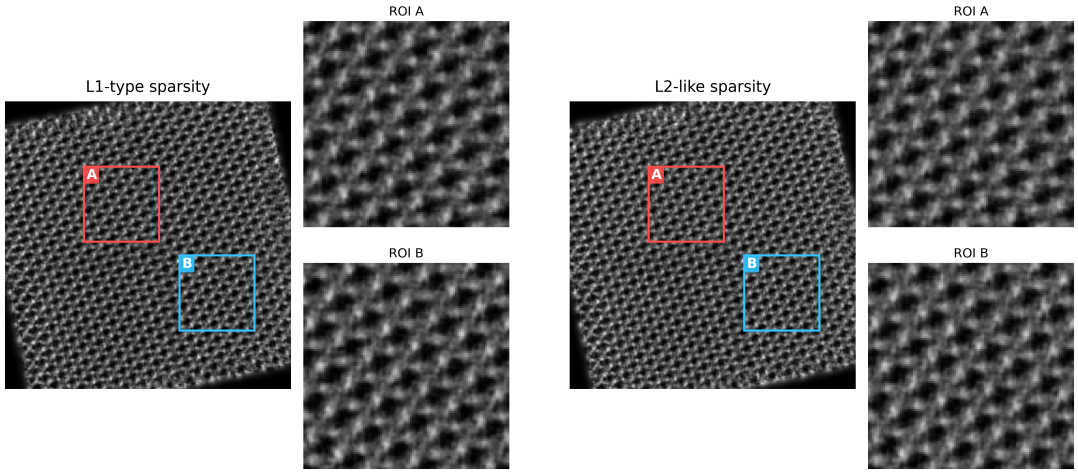


Figure 6-1. Direct comparison between the present L1-type sparsity setting and an L2-like alternative under otherwise identical reconstruction conditions. The left half shows the L1-type result and the right half the L2-like result. In each half, the full projected phase is shown together with two matched regions of interest (A and B) enlarged below, so that local lattice readability can be compared under a common spatial reference.

Figure 6-1 shows that the difference between the two runs is not dramatic at the scale of the whole projected image, but becomes more visible under matched local enlargements. In the selected regions, the L1-type result tends to preserve slightly more isolated lattice motifs and somewhat clearer separation between neighbouring bright features, whereas the L2-like result appears mildly broader and more diffuse.

6.2.2 With versus without PACBED participation

Between the two completed runs, the only intended objective-side change is whether `loss_pacbed` participates in the active objective. The baseline retains the present `loss_single + loss_pacbed + loss_sparse` setting, whereas the comparison run removes `loss_pacbed` while keeping the same random grouping and L1-type sparse regularisation. Because active objective is changed, the present comparison is interpreted through detector-side and image-side evidence rather than through total loss alone.

Figure 6-2 compares the representative forward pass at scan index `idx20075`.

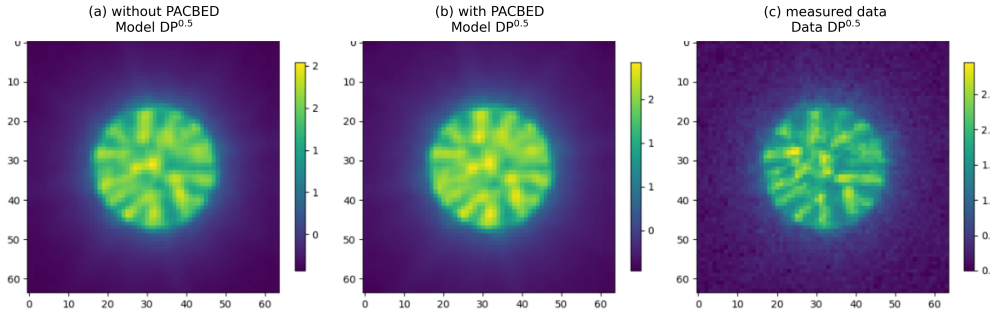


Figure 6-2. Forward-pass comparison at the representative scan position `idx20075` for the PACBED-participation study. The left panel shows the model diffraction pattern from the run without PACBED, the middle panel shows the model diffraction pattern from the baseline run with PACBED, and the right panel shows the measured diffraction pattern as a shared reference.

At this particular scan position, the run without PACBED appears slightly closer to the measured diffraction pattern in its local bright-field segmentation and intensity allocation. By contrast, the PACBED-on baseline is somewhat lighter and more broadly regulated. This is consistent with the structure of the PACBED term itself: it acts on a batch-mean diffraction statistic, and therefore redistributes optimisation pressure toward position-averaged detector behaviour rather than the best possible fit at one specific scan position. Under the present representative comparison, PACBED is therefore better read as a global diffraction-statistics term than as a term that improves local per-pattern fidelity.

Figure 6-3 shows a matched local region extracted from the projected phase. In this local comparison, the run without PACBED appears slightly cleaner, with somewhat sharper separation between neighbouring bright lattice features. In the current static objective, PACBED acts mainly as a global detector-statistics term: it improves position-averaged diffraction consistency, while the PACBED-off run preserves slightly cleaner local lattice separation in the projected image.

6.2.3 Stage-dependent loss scheduling

Figure 6-4 shows that replacing the fixed-loss baseline by the tested stage-dependent schedule gives a stronger final result under the shared baseline. Panels (a) and (b) summarise the two objective schedules, while panels (c) and (d) show their reconstruction outcomes at the common iteration-200 horizon. Relative to the fixed-loss baseline,

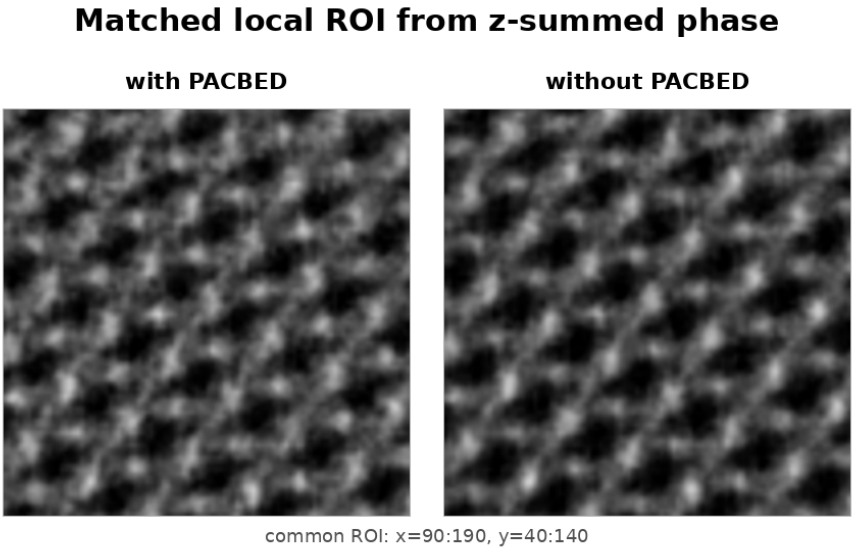


Figure 6-3. Matched local region extracted from the depth-summed object phase for the PACBED-on and PACBED-off runs. The same ROI is used in both panels so that local lattice readability can be compared under a common spatial reference.

the scheduled run finishes with a lower single-pattern term, decreasing from 0.1721 to 0.1663, and also a lower reported total loss, decreasing from 0.1822 to 0.1744.[1] The representative forward-pass comparison at scan index idx20075 is also visibly closer to the measured pattern in the scheduled run, and the corresponding absolute-difference view is reduced, indicating that the gain is reflected not only in scalar diagnostics but also in local detector-plane agreement.

This result is consistent with a stage-dependent reading of the objective terms. In the present framework, `loss_pacbed` acts mainly as a global diffraction-statistics term: it regulates position-averaged detector behaviour, but when applied too early or too strongly it can also smooth away some local structure that is more cleanly captured by scan-resolved fitting. It is therefore more naturally used as a late regulation term, once the dominant local structure has already been learned under `loss_single`. By contrast, the L_1-type sparsity term acts on object-side phase structure and is the more natural

[1] Because the scheduled run uses a reduced late-stage PACBED weight (0.3 rather than 0.5), the raw total loss is not directly comparable without adjustment. At iter 200, the scheduled run reports $\text{loss}_{\text{pacbed}} = 0.0030$ under weight 0.3, which corresponds to an implied contribution of 0.0050 under the baseline PACBED weight 0.5. The corresponding weight-restored total is therefore 0.1764, which remains below the fixed-loss baseline total 0.1822. The single-pattern term remains the cleanest like-for-like detector-side comparison because its weight stays at 1.0 in both runs.

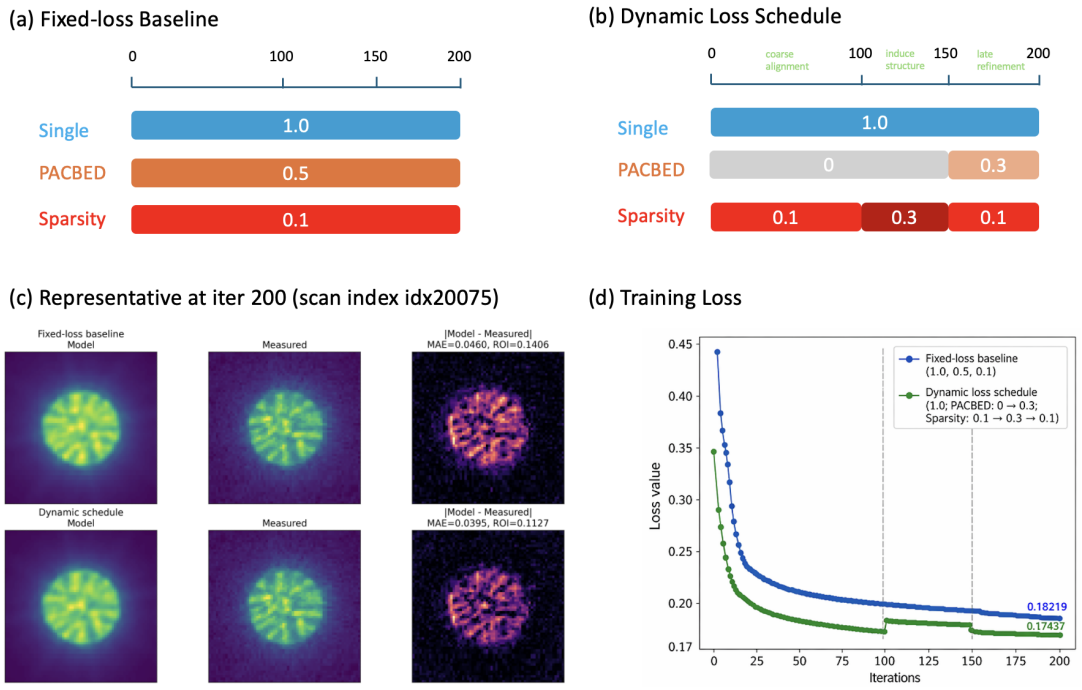


Figure 6-4. Comparison between the fixed-loss baseline and the tested stage-dependent loss schedule under the shared baseline. Panels (a) and (b) show the fixed and scheduled objective weights as functions of iteration. Panel (c) gives a representative forward-pass comparison at iter 200 and scan index idx20075. Panel (d) shows the training-loss trajectories for the two runs.

term for encouraging isolated lattice recovery, but under the fixed baseline its constant weight 0.1 appears too weak to shape the developing object strongly. Temporarily increasing that weight in the middle stage is therefore interpretable as a deliberate structural pulse: once the single-pattern term has established the main detector-side fit, stronger sparsity pressure can be used to promote cleaner specimen-side recovery before a weaker PACBED term is reintroduced to regulate global consistency.

6.3 Staged Activation and Parameter Coupling

Across the three completed runs, the only intended intervention-side change is the start iteration assigned to the trainable parameter blocks. The baseline uses the present schedule in which object amplitude and phase are active from iter 1, probe updates begin at iter 5, and tilt together with sub-pixel position correction begin at iter 10. The first comparison run, referred to here as the early-probe / late-geometry variant (`ltst`), moves probe updates forward to iter 1 while delaying tilt and position to iter 20. The

second comparison run, referred to as the object-only warm-up variant (`ltstProb`), delays probe, tilt, and position together to iter 20.

Table 6-1. Comparison of the three staged-activation schedules. Because the objective family is unchanged across this experiment group, the total loss remains directly comparable.

Run	Start-iter pattern	Iter 5	Iter 10	Iter 20	Iter 200	Avg. iter (s)
Baseline	probe@5; tilt,pos@10	0.3444	0.2601	0.2224	0.1822	10.265
`ltst`	probe@1; tilt,pos@20	0.2949	0.2380	0.2217	0.1821	10.232
`ltstProb`	probe,tilt,pos@20	0.3475	0.3256	0.3125	0.1828	10.241

The detailed loss trajectories, probe-mode summaries, and average-tilt diagnostics are retained as the main evidence for this section and are shown in Figs. D-1, D-3, and D-2. The comparison can be read most directly in four points:

(a) **Early-stage behaviour is most sensitive to probe timing.** Between the two delayed-geometry runs, the `ltst` variant reaches 0.2949 at iter 5, 0.2380 at iter 10, and 0.2217 at iter 20, whereas `ltstProb` remains at 0.3475, 0.3256, and 0.3125 at the same checkpoints. Since these two runs share the same delayed opening of tilt and position, this early gap is best understood as a schedule effect of probe timing rather than of geometry timing.

(b) **Object-only warm-up does not improve the implemented sparsity behaviour.** The `ltstProb` run carries the largest early `loss_sparse` values within this family, reaching 0.0033 at iter 10 and 0.0041 at iter 20, compared with 0.0028 / 0.0029 for the baseline and 0.0024 / 0.0028 for `ltst`. Once probe, tilt, and position are released at iter 20, `ltstProb` enters a second fast descent, with the PACBED term dropping rapidly from its earlier elevated level. Under the present objective, delaying probe therefore does not behave as an implicit regulariser; instead, it temporarily leaves more detector-side mismatch concentrated on the object.

(c) **Probe and tilt diagnostics show schedule-sensitive freedom allocation.** Figures D-3 and D-2 show that the three schedules do not end in the same internal state, even when their final total losses are very close. Relative to the baseline and `ltstProb`, the `ltst` run redistributes more power away from the dominant probe mode and settles to a different average-tilt trajectory. The present evidence does not by itself decide whether the extra mixed-state probe activity represents more

useful physical modelling or more residual absorption. It does, however, establish that mixed-state probe participation is strongly schedule-sensitive.

(d) **The main effect of schedule is on path rather than on final detector-side fit.** By the common 200-iteration horizon, the three runs are already close in total loss and runtime: the baseline reaches 0.1822, `ltst` 0.1821, and `ltstProb` 0.1828, while the average iteration times remain within a few hundredths of a second. The schedule intervention is therefore better read as a change in optimisation path and modelling-freedom allocation than as a source of dramatic final quality separation.

Taken together, this study suggests that early probe release matters more than early geometry release in first twenty iterations of present baseline, whereas delaying geometry to iter 20 appears largely benign at common horizon. The strongest evidence lies in the trajectory and parameter-side diagnostics rather than in a large specimen-side separation.

6.4 Batch Geometry and PACBED Estimation Quality

Across the three completed runs, the only intended intervention-side change is the grouping rule used to form each batch: the baseline keeps the present `random` grouping, whereas the two comparison runs replace it with `compact` and `sparse` alternatives.

Figure D-4 shows that the two alternatives expose scan positions to each update in clearly different ways. Relative to the random baseline of Fig. 5-2, compact grouping keeps each batch spatially local, whereas sparse grouping distributes the same batch budget more broadly across the active field of view. Under the present implementation, this difference is not purely geometric:

(a) `loss_pacbed` is evaluated from the batch-mean diffraction pattern, so a more spatially distributed batch is better placed to provide a representative PACBED estimate for each update.

(b) `loss_sparse` is evaluated on batch-local object phase patches, so a more compact batch is naturally better aligned with this local regulariser.

The grouping choice therefore changes the quality of the batch-wise PACBED estimate seen by each update, while also changing how easily the local sparsity term can be satisfied. Table 6-2 summarises the comparison at the common iteration-200 horizon, and the detailed trajectories and object-side outputs are shown in Figs. D-5 and D-6.

Table 6-2. Comparison of the three batch geometries at the common iteration-200 horizon. The detector-side terms provide the main evidence for this section, while `loss_sparse` is reported as a secondary diagnostic because it acts on batch-local phase patches.

Grouping	Total loss	`loss_single`	`loss_pacbed`	`loss_sparse`	Avg. iter time (s)
Random	0.1822	0.1721	0.0053	0.0048	10.265
Compact	0.2319	0.2134	0.0137	0.0049	9.517
Sparse	0.1816	0.1716	0.0052	0.0049	10.602

The comparison can be read most directly in four points:

(a) The compact run gives the weakest detector-side fit. Its final total loss rises to 0.2319, with `loss_single`=0.2134 and `loss_pacbed`=0.0137, both much higher than the random baseline and the sparse variant. Its loss trace is also visibly more oscillatory in Fig. D-5.

(b) In the selected zoom region of Fig. D-6, the sparse result appears slightly cleaner than the random baseline, whereas the compact reconstruction is visibly more degraded. This local comparison is consistent with the idea that sparse grouping provides a slightly better batch-wise PACBED estimate.

(c) A secondary but important observation is that compact grouping is better aligned with the local L1-type sparsity term, but this advantage should not be overstated. Across all three runs, `loss_sparse` is comparatively low at the beginning and then rises as optimisation proceeds, so the sparsity term is not the dominant source of late-stage improvement in this family. Compact grouping therefore does show a local regularisation advantage, but that advantage is secondary: `loss_sparse` carries a much smaller weight than the detector-side terms, and its behaviour is partly counterposed to detector-side fitting rather than indicative of globally better reconstruction. The present evidence is therefore more consistent with a detector-dominated trade than with a sparsity-dominated ranking.

(d) Runtime shows the trade-off directly. Compact is also the fastest run, with an average iteration time of 9.517 s, compared with 10.265 s for the random baseline and 10.602 s for the sparse variant. The present comparison is therefore better read as a runtime–quality trade-off than as a simple ranking: compact exchanges detector-side fidelity for speed, whereas sparse preserves slightly cleaner local output at a modest runtime cost.

Under the present objective, batch geometry changes reconstruction behaviour through PACBED estimation quality. The main effect appears on detector-side fitting, while the local sparsity term provides only a weaker, secondary counter-tendency in favour of compact grouping.

6.5 Late-stage LBFGS refinement

Across the completed runs in this family, the common reconstruction horizon remains fixed at iter 200, and the only intended intervention-side change is whether the final twenty iterations are left on the unchanged Adam trajectory or replaced by a late-stage LBFGS refinement stage introduced at iter 181. The baseline therefore continues with Adam throughout, whereas the main comparison switches all variables already active under the baseline schedule to LBFGS over iter 181–200. A secondary geometry-only variant is also shown in Fig. 6-5 for reference.[2] Because the active objective is unchanged across this comparison family, the total loss remains directly comparable.

Figure 6-5 shows that the late-stage LBFGS switch does not provide additional last-mile descent under the present implementation. Relative to the baseline, the all-active refinement leaves the existing descent trajectory immediately at the switch point and then continues upward throughout the final window. The jump at iter 181 is therefore not a transient fluctuation but the start of a qualitatively different late-stage behaviour. The geometry-only variant follows the same pattern, so the negative result cannot be attributed solely to the larger freedom of the all-active refinement.

A cautious theoretical reading is more informative here than a simple optimizer ranking. Quasi-Newton refinement relies on locally coherent curvature information, whereas the present LBFGS stage is inserted into a mini-batch reconstruction loop and the iterates remain subject to persistent post-step constraints. In that setting, the local curvature approximation need not remain consistent from one update to the next, so the present deterioration should not be read as evidence that curvature-aware refinement is useless in principle. The safer interpretation is narrower: terminal LBFGS switching appears poorly matched to the projected, mini-batch implementation used here, even

[2]The geometry-only variant exhibits the same qualitative deterioration, rising from 0.1857 at iter 180 to 0.1954 at iter 181 and ending at 0.2230 by iter 200, so restricting the refinement to geometry variables does not recover a useful late-stage descent. A non-geometry-only variant was also explored, but it is not retained here because operational feasibility under the available GPU-memory budget was not established.

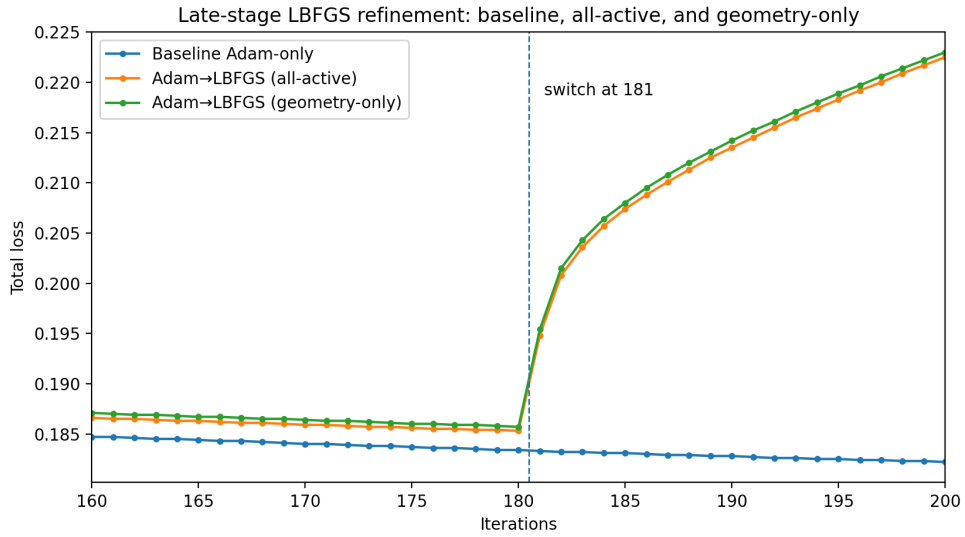


Figure 6-5. Zoomed total-loss trajectories over iterations 160–200 for the baseline Adam reconstruction, the all-active late-stage LBFGS switch, and the geometry-only late-stage LBFGS switch. The baseline continues its late-stage monotone descent to 0.1822 at iter 200. By contrast, the all-active LBFGS run rises immediately from 0.1853 at iter 180 to 0.1948 at iter 181 and then continues upward to 0.2225 by iter 200. The geometry-only variant shows the same qualitative behaviour, rising from 0.1857 to 0.1954 at the switch and ending at 0.2230.

after the baseline has already entered its slow regime.

For the purposes of the present thesis, the resulting conclusion is negative but limited. As implemented here, late-stage LBFGS refinement is not worthwhile as a default continuation strategy for the shared baseline. A more plausible next test than terminal switching would be a limited curvature pulse introduced earlier in the descent, but that possibility is deferred to Chapter 7.

Chapter 7

Discussion

This chapter draws the controlled studies together around four recurring points already foreshadowed by the working hypotheses of Chapter 5 and tested in Chapter 6. The first is that objective terms regulate different scales of agreement, and that their usefulness may change across optimisation stages. The second is that PtyRAD is path-dependent, so staged activation changes how modelling freedom is allocated during fitting. The third is that overlap is not only acquired by the scan, but also exposed to the optimiser through batch geometry. The fourth is that mixed-state flexibility is useful, but residual allocation and model mismatch remain unresolved. These four points are not all supported at the same evidential level. Some follow directly from the present comparisons, whereas others remain physically or numerically plausible but still require more targeted validation. The discussion therefore re-reads the Chapter 6 results through an evidence-tiered lens, so that what is established, what is only suggested, and what remains a clear limit of the present work are kept distinct.

7.1 What Is Well Supported by Evidence

Objective terms regulate different scales of agreement. The first conclusion already well supported by the present work is that objective variation should not be read as a search for a single universally better loss. At the level of formulation, the active terms in Eq. (4.12) do not enforce the same notion of fit. The single-pattern term of Eq. (4.16) acts on scan-resolved detector mismatch, the PACBED term of Eq. (4.18) acts on a position-averaged diffraction statistic, and the sparsity term of Eq. (4.19a) acts directly on object-side phase structure [14, 16, 30, 32]. This distinction is also visible empirically. Section 5.5.2 already showed that the three active components do not evolve uniformly under the practical reduced baseline, and Section 6.2.2 together with Section 6.2.1 then made the trade-offs explicit: the L_1-type sparse setting preserves local lattice isolation somewhat more clearly than the L_2-like alternative, whereas PACBED participation improves global diffraction consistency while slightly relaxing local feature separation.

The present reconstruction is path-dependent, and staged activation changes modelling-freedom allocation. The second conclusion already well supported by the present work is that the reconstruction behaviour is visibly path-dependent under staged activation. This follows first from the structure of the baseline itself, which is already staged rather than fully simultaneous, and then more directly from the schedule comparison of Section 6.3. There, early probe release changes the first part of the descent much more strongly than delayed geometry release, whereas the three runs remain close in total loss and runtime by the common iteration-200 horizon. At the same time, the supplementary probe-mode and average-tilt diagnostics show that similar terminal detector-side loss does not imply the same internal state: probe-mode redistribution and tilt trajectories remain schedule-sensitive. Under the present objective, object-only warm-up also does not behave as an implicit regulariser; it temporarily leaves more detector-side mismatch concentrated on the object. The strongest evidence therefore concerns optimisation path and freedom allocation, not a dramatic final-quality separation.

Batch geometry mediates how overlap information is exposed to each update. A further conclusion well supported by the present work is that batch geometry should not be treated as a neutral implementation detail. Under the present baseline, real-space overlap is fixed by the acquisition itself, but grouping determines how much of that overlap structure is presented jointly within a single optimisation step. This matters because the active objective is not uniformly local: Section 4.3 already showed that `loss_pacbed` is evaluated from a batch-mean diffraction statistic, whereas `loss_sparse` acts on batch-local object phase patches. Section 6.4 then makes the practical consequence clear. Compact grouping gives the weakest detector-side fit, especially in `loss_single` and `loss_pacbed`, even though it does not perform badly on the sparsity term itself and is also the fastest configuration. Sparse grouping, by contrast, remains close to the random baseline in detector-side loss while appearing slightly cleaner in the selected local comparison. Under the present objective, batch geometry therefore changes reconstruction behaviour primarily through detector-side consistency, not through a large direct advantage in local sparsity satisfaction.

7.2 What Is Plausible but Not Yet Fully Validated

PACBED may also stabilise system-sensitive variables, and the current static weighting may be stage-mismatched. A stronger interpretation of the objective-design results is physically plausible, but not yet fully demonstrated by the present experiments. Because the PACBED term collapses local scan dependence while preserving a position-averaged diffraction statistic, Chapter 4 already identified it as especially relevant to variables whose effects are diffuse or strongly coupled, such as thickness and tilt [6, 16, 32]. The current PACBED-on/PACBED-off comparison is consistent with that role, but it does not isolate thickness or tilt recovery directly. The earlier objective-design results suggested that fixed weighting may be stage-mismatched, and the dynamic-schedule test now provides preliminary support for that reading: delaying PACBED and pulsing sparsity changes the optimisation path and improves the dominant single-pattern term under the shared baseline.

Schedule may do more than delay updates; it may route mismatch and partially pre-shape the later basin. A second plausible reading is that schedule changes not only when a parameter block is updated, but also where unresolved mismatch is provisionally stored. The Section 6.3 comparison is consistent with a mismatch-routing picture in which the block released earlier carries more of the initial explanatory burden, while later blocks inherit a basin already partially shaped by the earlier fit. This interpretation also helps explain why similar terminal losses can coexist with different probe-mode occupancies and different tilt trajectories. Even so, the present thesis has not yet isolated the causal mechanism strongly enough to decide how much of this behaviour reflects useful physical modelling and how much reflects residual absorption. For that reason, mismatch routing and basin pre-shaping are retained here as working interpretations rather than validated conclusions.

Batch geometry may influence more than PACBED estimation quality alone. A stronger interpretation is plausible, but not yet fully validated by the present comparison. The immediate mechanism visible in Section 6.4 is that different grouping geometries change the quality of the batch-wise PACBED estimate seen by each update. Even so, the role of grouping may be somewhat broader than PACBED estimation alone. Because

compact and sparse batches expose different spatial distributions of probe positions within the same effective batch budget, they may also alter which overlap couplings become operational during a given update. This reading is consistent with the 5.6.4 formulation of Chapter 5, but the present experiments do not yet isolate that mechanism directly. For that reason, the safer interpretation is retained here: the current evidence strongly supports a detector-side effect and is compatible with a wider overlap-exposure effect, but does not yet prove it.

Mixed-state ambiguity is tied to structured residual and unresolved mismatch allocation. A stronger interpretation becomes plausible once the mixed-state behaviour is read together with the residual structure of the baseline. The remaining detector-side mismatch is not random; it remains structured, especially in the diffuse background and peripheral parts of the diffraction patterns. This makes it difficult to assign the unresolved error to one source alone. The current baseline does not yet disentangle measurement formation, illumination-side non-ideality, and forward-model adequacy cleanly enough for probe-mode redistribution to be read as a uniquely physical signature of partial coherence. A plausible interpretation is therefore that part of the additional modal activity reflects unresolved mismatch absorption rather than purely improved physical identification, although the present experiments do not isolate that mechanism directly.

7.3 What This Reveals About PtyRAD as an Optimisation Problem

Objective design changes the reconstruction itself. Taken together, these results show that the objective actively shapes reconstruction. Different terms emphasise different comparison variables, aggregation levels, and optimisation stages, so changing the objective changes which discrepancies are penalised and which corrections are favoured. In PtyRAD, objective design sets the balance between detector-side consistency, position-averaged diffraction structure, and object-side admissibility.

Staged activation is part of the solver. The schedule study shows that staged activation determines which parameter blocks can respond to mismatch at each stage of the run. It therefore changes optimisation path, internal freedom allocation, and the final internal

state, even when terminal losses remain close. PtyRAD is a coupled nonconvex inverse problem whose practical behaviour is shaped by when different parameter freedoms are released.

Batch geometry helps define the optimisation signal. Under the present AD-based reconstruction loop, batching determines more than computational workload. Because some active objective terms are evaluated on batch-level summaries, the grouping rule changes which overlap relations and detector statistics are exposed to each update. Batch geometry therefore belongs to the practical solver description of PtyRAD.

Additional modelling freedom serves two roles. Additional modelling freedom increases the model's ability to fit the data, and it can also absorb unresolved mismatch. The present baseline already shows that mixed-state modelling is useful, but modal redistribution cannot yet be read as a uniquely physical signature. Within the current reconstruction system, increased flexibility can improve data response while also blurring the boundary between physically meaningful adaptation and numerically convenient residual absorption.

Late-stage refinement remains implementation-dependent. The negative LBFGS result should be read narrowly but clearly. In the present mini-batch reconstruction loop with persistent post-step constraints, terminal switching to LBFGS does not improve the baseline and is not supported here as a default continuation strategy. The result does not rule out curvature-based refinement in general, but it does show that the insertion point and implementation details matter.

7.4 Limits of the Present Work

All claims above remain conditional on one reduced experimental baseline and one common horizon. All conclusions above are drawn from the practical reduced $SrTiO_3$ baseline defined in Chapter 5: a central 100×100 scan subset, on-the-fly detector resampling to an effective $N_{\mathrm{pix}} = 64$, a 32-slice multislice model, four probe modes, a fixed objective family, and a common 200-iteration horizon. The results are therefore evidence-based within this setting, but they are not intended as context-free prescriptions

for all PtyRAD reconstructions. In particular, the present schedule comparison supports only the restrained statement that delaying geometry to iter 20 appears largely benign within this baseline and horizon, not that geometry timing is generically unimportant.

Several of the key interpretations remain proxy-based rather than directly validated. A second limitation is that the present evidence remains proxy-based at several important points. The dataset is experimental, so no exact ground-truth object is available; cleaner local phase structure and lower detector-side loss cannot by themselves certify a physically better internal allocation of freedom. For the same reason, the thesis cannot yet decide whether schedule-sensitive probe redistribution represents more appropriate physical modelling or more residual absorption, nor can it directly validate the PACBED-as-stabiliser interpretation or the mismatch-routing conjecture. These are informative boundaries rather than failures, and they define the most natural targets for the next round of controlled study.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

This thesis analysed PtyRAD as a concrete example of modern automatic-differentiation-based electron ptychography. On the theoretical side, Chapters 2–4 developed a layered account of the problem, moving from the phase problem and overlap-based recoverability to the forward model and the optimisation formulation. Within this framework, object, probe, scan geometry, multislice propagation, mixed-state components, and detector-side measurement formation enter a single reconstruction model. The reconstruction objective combines detector-side fidelity terms, regularisation, and iteration-wise constraint enforcement. This framework provides a clearer basis for interpreting how physical modelling and optimisation interact in PtyRAD.

On the practical side, the thesis defined a practical reduced baseline on the experimental $SrTiO_3$ dataset and used that baseline to construct a small set of controlled studies. By fixing the reconstruction setting and modifying only one intervention family at a time, the work was able to separate the effects of objective design, staged activation, and batch geometry without losing contact with an experimentally meaningful pipeline. The baseline itself already proved diagnostically useful. It reproduced the main detector-space morphology and a readable projected object, yet it also exposed three recurrent tensions: clear late-stage slowing-down after early coarse fitting, uneven behaviour across the active objective terms, and an unresolved allocation of residual mismatch between probe-side flexibility and structured detector-side error.

Within that controlled setting, several preliminary but coherent findings emerged:

(a) First, objective terms regulate different scales of agreement: single-pattern fidelity, PACBED participation, and sparsity regularisation do not ask the reconstruction to match the same aspect of the data, and altering them changes the balance between scan-resolved fitting, position-averaged detector statistics, and object-side structural selectivity.

(b) Second, PtyRAD is path-dependent: staged activation changes when modelling freedom is released, and this changes the optimisation path and internal freedom allocation even when the final detector-side losses remain close.

(c) Third, overlap is not only acquired by the scan; it is also exposed to the optimiser through batch geometry. The compact, random, and sparse grouping comparison shows that how measurements are partitioned per update can materially alter detector-side fitting behaviour.

(d) Fourth, mixed-state flexibility is clearly useful under the present baseline, but its interpretation remains non-unique because probe redistribution cannot yet be disentangled cleanly from residual absorption and model mismatch.

(e) Fifth, late-stage LBFGS switching did not improve the shared baseline under the present implementation and is therefore not supported here as a default continuation strategy.

(f) Sixth, a preliminary stage-dependent loss schedule improves the dominant scan-resolved detector term under the shared baseline, which supports the view that objective design in PtyRAD should be treated temporally as well as structurally.

These conclusions are intentionally conditional. They do not establish a universally optimal reconstruction recipe, and they do not yet fully resolve the role of PACBED, the mechanism of mismatch routing, or the physical meaning of mixed-state redistribution. Even with these limits, the thesis establishes a clear framework for interpreting PtyRAD as a nonconvex reconstruction problem shaped jointly by the forward model, the objective, the update schedule, and the batching scheme. Its main contribution is a more explicit account of how these elements interact in modern electron ptychography.

8.2 Future Work

Several next steps follow directly from the present limits. First, semismooth Newton methods remain a promising but as yet unrealised direction. As argued in Chapter 4, their real value would emerge only after positivity-like constraints, clipping operations, sparsity terms, or similar nonsmooth components are reformulated as explicit nonsmooth subproblems rather than handled mainly by post-step enforcement. In that setting, the semismooth Newton and semismooth Newton–augmented-Lagrangian line associated with Defeng Sun and collaborators offers a plausible route to stronger local refinement

on structured nonsmooth blocks, especially when sparsity-promoting penalties and low-dimensional system variables must be treated together [44].

Second, the present regularisation strategy could be strengthened beyond hand-crafted penalties. Recent work on MEP-Diffusion shows that a generative prior can be integrated with multislice electron ptychography through diffusion posterior sampling, thereby augmenting an iterative solver with a learned structural prior for 3D reconstruction [45]. In the present context, such a direction is attractive not as a replacement for the physical forward model, but as a stronger regularisation layer for weakly informed volumetric directions.

Finally, broader validation is still needed. Simulated ground-truth studies and additional experimental datasets would allow future work to distinguish improved physical modelling from residual absorption more directly, and would provide the natural setting in which to test stage-dependent weighting, stage-dependent batching, more explicit detector/background modelling, and targeted second-stage refinement on low-dimensional system variables.

Bibliography

[1] Laura Clark and Peter D. Nellist. Electron ptychography, 2025. URL https://arxiv.org/abs/2503.10917.

[2] Hao Yang, Timothy J. Pennycook, and Peter D. Nellist. Efficient phase contrast imaging in STEM using a pixelated detector. part II: Optimisation of imaging conditions. *Ultramicroscopy*, 151:232–239, 2015. doi: 10.1016/j.ultramic.2014.10.013.

[3] Yi Jiang, Zhen Chen, Yimo Han, Pratiti Deb, Hui Gao, Saien Xie, Prafull Purohit, Mark W. Tate, Jiwoong Park, Sol M. Gruner, Veit Elser, and David A. Muller. Electron ptychography of 2D materials to deep sub-ångstr"om resolution. *Nature*, 559(7714):343–349, 2018. doi: 10.1038/s41586-018-0298-5.

[4] P. A. Midgley. An introduction to off-axis electron holography. *Micron*, 32(2):167–184, 2001. doi: 10.1016/S0968-4328(99)00105-5.

[5] Marco Beleggia, Mark A. Schofield, Vadim V. Volkov, and Yimei Zhu. On the transport of intensity technique for phase retrieval. *Ultramicroscopy*, 102(1):37–49, 2004. doi: 10.1016/j.ultramic.2004.08.004.

[6] Colin Ophus. Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microscopy and Microanalysis*, 25(3):563–582, 2019. doi: 10.1017/S1431927619000497.

[7] H. M. L. Faulkner and J. M. Rodenburg. Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm. *Physical Review Letters*, 93(2):023903, 2004. doi: 10.1103/PhysRevLett.93.023903.

[8] Pierre Thibault, Martin Dierolf, Oliver Bunk, Andreas Menzel, and Franz Pfeiffer. Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*, 109(4):338–343, 2009. doi: 10.1016/j.ultramic.2008.12.011.

[9] Oliver Bunk, Martin Dierolf, Søren Kynde, Ian Johnson, Othmar Marti, and Franz Pfeiffer. Influence of the overlap parameter on the convergence of the

ptychographical iterative engine. *Ultramicroscopy*, 108(5):481–487, 2008. doi: 10.1016/j.ultramic.2007.08.003.

[10] Andrew M. Maiden and John M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10):1256–1262, 2009. doi: 10.1016/j.ultramic.2009.05.012.

[11] F. H"ue, J. M. Rodenburg, A. M. Maiden, and P. A. Midgley. Extended ptychography in the transmission electron microscope: Possibilities and limitations. *Ultramicroscopy*, 111(8):1117–1123, 2011. doi: 10.1016/j.ultramic.2011.02.005.

[12] Andrew Maiden, Daniel Johnson, and Peng Li. Further improvements to the ptychographical iterative engine. *Optica*, 4(7):736–745, 2017. doi: 10.1364/OPTICA.4.000736.

[13] Amirafshar Moshtaghpour, Abner Velazco-Torrejon, Alex W. Robinson, Nigel D. Browning, and Angus I. Kirkland. LoRePIE: ℓ_0 regularized extended ptychographical iterative engine for low-dose and fast electron ptychography. *Optics Express*, 33(5):9357–9368, 2025. doi: 10.1364/OE.551986.

[14] Michal Odstrcil, Andreas Menzel, and Manuel Guizar-Sicairos. Iterative least-squares solver for generalized maximum-likelihood ptychography. *Optics Express*, 26(3):3108–3123, 2018. doi: 10.1364/OE.26.003108.

[15] Saugat Kandel, Siddharth Maddali, Marc Allain, Stephan O. Hruszkewycz, Chris Jacobsen, and Youssef S. G. Nashed. Using automatic differentiation as a general framework for ptychographic reconstruction. *Optics Express*, 27(13):18653–18672, 2019. doi: 10.1364/OE.27.018653.

[16] Chia-Hao Lee, Steven E. Zeltmann, Dasol Yoon, Desheng Ma, and David A. Muller. PtyRAD: A high-performance and flexible ptychographic reconstruction framework with automatic differentiation. *Microscopy and Microanalysis*, 31(4): ozaf070, 2025. doi: 10.1093/mam/ozaf070.

[17] Colin Gilgenbach, Menglin Zhu, and James M. LeBeau. phaser: a unified and extensible framework for fast electron ptychography. *npj Computational Materials*, 12:89, 2026. doi: 10.1038/s41524-026-01956-8.

[18] J. M. Rodenburg and A. M. Maiden. Ptychography. In Peter W. Hawkes and John C. H. Spence, editors, *Springer Handbook of Microscopy*, pages 819–904. Springer, Cham, 2019. ISBN 978-3-030-00068-4. doi: 10.1007/978-3-030-00069-1_17.

[19] J. M. Rodenburg and H. M. L. Faulkner. A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20):4795–4797, 2004. doi: 10.1063/1.1823034.

[20] Si Gao, Peng Wang, Fucai Zhang, Gerardo T. Martinez, Peter D. Nellist, Xiaoqing Pan, and Angus I. Kirkland. Electron ptychographic microscopy for three-dimensional imaging. *Nature Communications*, 8:163, 2017. doi: 10.1038/s41467-017-00150-1.

[21] J. M. Cowley and A. F. Moodie. The scattering of electrons by atoms and crystals. i. a new theoretical approach. *Acta Crystallographica*, 10(10):609–619, 1957. doi: 10.1107/S0365110X57002194.

[22] K. Ishizuka and N. Uyeda. A new theoretical and practical approach to the multislice method. *Acta Crystallographica Section A*, 33(5):740–749, 1977. doi: 10.1107/S0567739477001879.

[23] Laura Clark, Gerardo T. Martinez, Colum M. O'Leary, Hao Yang, Zhiyuan Ding, Timothy C. Petersen, Scott D. Findlay, and Peter D. Nellist. The effect of dynamical scattering on single-plane phase retrieval in electron ptychography. *Microscopy and Microanalysis*, 29(1):384–394, 2023. doi: 10.1093/micmic/ozac022.

[24] Pierre Thibault and Andreas Menzel. Reconstructing state mixtures from diffraction measurements. *Nature*, 494(7435):68–71, 2013. doi: 10.1038/nature11806.

[25] Zhen Chen, Michal Odstrcil, Yi Jiang, Yimo Han, Ming-Hui Chiu, Lain-Jong Li, and David A. Muller. Mixed-state electron ptychography enables sub-angstrom resolution imaging with picometer precision at low dose. *Nature Communications*, 11:2994, 2020. doi: 10.1038/s41467-020-16688-6.

[26] G. McMullan, A. R. Faruqi, D. Clare, and R. Henderson. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy*, 109 (9):1126–1143, 2009. doi: 10.1016/j.ultramic.2009.04.002.

[27] R. S. Ruskin, Z. Yu, and N. Grigorieff. Quantitative characterization of electron detectors for transmission electron microscopy. *Journal of Structural Biology*, 184 (3):385–393, 2013. doi: 10.1016/j.jsb.2013.10.016.

[28] PyTorch Developers. Autograd mechanics, 2025. URL https://docs.pytorch.org/docs/stable/notes/autograd.html. PyTorch 2.10 documentation, last updated Jun 16, 2025.

[29] Atilim Gunes Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018.

[30] Chia-Hao Lee. PtyRAD: Ptychographic reconstruction with automatic differentiation. GitHub repository, 2026. URL https://github.com/chiahao3/ptyrad. Accessed 2026-03-27.

[31] Xukang Wei, H. Paul Urbach, and Wim M. J. Coene. Crámer–rao lower bound and maximum-likelihood estimation in ptychography with Poisson noise. *Physical Review A*, 102(4):043516, 2020. doi: 10.1103/PhysRevA.102.043516.

[32] James M. LeBeau, Scott D. Findlay, Leslie J. Allen, and Susanne Stemmer. Position averaged convergent beam electron diffraction: Theory and applications. *Ultramicroscopy*, 110(2):118–125, 2010. doi: 10.1016/j.ultramic.2009.10.001.

[33] Vladimir Katkovnik and Jaakko Astola. Sparse ptychographical coherent diffractive imaging from noisy measurements. *Journal of the Optical Society of America A*, 30(3):367–379, 2013. doi: 10.1364/JOSAA.30.000367.

[34] Michal Odstrcil, P. Baksh, S. A. Boden, R. Card, J. E. Chad, J. G. Frey, and W. S. Brocklesby. Ptychographic coherent diffractive imaging with orthogonal probe relaxation. *Optics Express*, 24(8):8360–8369, 2016. doi: 10.1364/OE.24.008360.

[35] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus, and J. M. Rodenburg. An annealing algorithm to correct positioning errors in ptychography. *Ultramicroscopy*, 120:64–72, 2012. doi: 10.1016/j.ultramic.2012.06.001.

[36] Ashish Tripathi, Ian McNulty, and Oleg G. Shpyrko. Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient

descent methods. *Optics Express*, 22(2):1452–1466, 2014. doi: 10.1364/OE.22.001452.

[37] Peng Li, Tega Edo, Darren Batey, John Rodenburg, and Andrew Maiden. Breaking ambiguities in mixed state ptychography. *Optics Express*, 24(8):9038–9052, 2016. doi: 10.1364/OE.24.009038.

[38] Saugat Kandel, Siddharth Maddali, Youssef S. G. Nashed, Stephan O. Hruszkewycz, Chris Jacobsen, and Marc Allain. Efficient ptychographic phase retrieval via a matrix-free Levenberg–Marquardt algorithm. *Optics Express*, 29(15):23019–23055, 2021. doi: 10.1364/OE.422768.

[39] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, 2 edition, 2006. doi: 10.1007/978-0-387-40065-5.

[40] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015. URL https://arxiv.org/abs/1412.6980.

[41] PyTorch Developers. Adam — PyTorch documentation, 2026. URL https://docs.pytorch.org/docs/stable/generated/torch.optim.Adam.html. Accessed 2026-03-28.

[42] Michael Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. SIAM, Philadelphia, 2011. doi: 10.1137/1.9781611970692.

[43] Florian Mannel and Armin Rund. A hybrid semismooth quasi-newton method for nonsmooth optimal control with PDEs. *Optimization and Engineering*, 22: 2087–2125, 2021. doi: 10.1007/s11081-020-09523-w.

[44] Xudong Li, Defeng Sun, and Kim-Chuan Toh. A highly efficient semismooth newton augmented lagrangian method for solving Lasso problems. *SIAM Journal on Optimization*, 28(1):433–458, 2018. doi: 10.1137/16M1097572.

[45] Christian K. Belardi, Chia-Hao Lee, Yingheng Wang, Justin Lovelace, Kilian Q. Weinberger, David A. Muller, and Carla P. Gomes. Improving multislice electron

ptychography with a generative prior, 2025. URL https://arxiv.org/abs/2507.17800.

[46] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13:281–305, 2012. URL https://www.jmlr.org/papers/v13/bergstra12a.html.

[47] Kevin Jamieson and Ameet Talwalkar. Non-stochastic best arm identification and hyperparameter optimization. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 240–248. PMLR, 2016. URL https://proceedings.mlr.press/v51/jamieson16.html.

[48] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185):1–52, 2018. URL https://www.jmlr.org/papers/v18/16-558.html.

[49] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2623–2631. ACM, 2019. doi: 10.1145/3292500.3330701.

[50] James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In *Advances in Neural Information Processing Systems 24*, pages 2546–2554, 2011.

[51] Optuna Developers. optuna.samplers.tpesampler — optuna documentation. Online documentation, 2026. URL https://optuna.readthedocs.io/en/stable/reference/generated/optuna.samplers.TPESampler.html. Accessed 11 April 2026.

[52] Optuna Developers. optuna.samplers.basesampler — optuna documentation. Online documentation, 2026. URL https://optuna.readthedocs.io/en/stable/reference/generated/optuna.samplers.BaseSampler.html. Accessed 11 April 2026.

[53] Optuna Developers. optuna.trial.trial — optuna documentation. Online documentation, 2026. URL https://optuna.readthedocs.io/en/stable/reference/generated/optuna.trial.Trial.html. Accessed 11 April 2026.

[54] Optuna Developers. optuna.pruners.hyperbandpruner — optuna documentation. Online documentation, 2026. URL https://optuna.readthedocs.io/en/stable/reference/generated/optuna.pruners.HyperbandPruner.html. Accessed 11 April 2026.

[55] Optuna Developers. optuna.study.create_study — optuna documentation. Online documentation, 2026. URL https://optuna.readthedocs.io/en/stable/reference/generated/optuna.study.create_study.html. Accessed 11 April 2026.

Chapter A

Supplementary Notes for Multislice Model

A.1 Notation and Fourier-domain conventions

This appendix supports Section 3.3 by making the multislice propagation operator explicit and by recording the minimal derivational bridge from a paraxial electron-wave picture to the recursion. Let $\mathbf{r} = (x, y)$ denote the transverse real-space coordinate, let z denote the beam direction, and let $\mathbf{q} = (q_x, q_y)$ denote the reciprocal-space coordinate used in Chapter 3. We write the two-dimensional Fourier transform as $\mathcal{F}\{\cdot\}$ with inverse $\mathcal{F}^{-1}[\cdot]$. When needed, the corresponding angular-frequency variables are

$$k_x = 2\pi q_x, \qquad k_y = 2\pi q_y, \qquad k = \frac{2\pi}{\lambda}, \tag{A.1}$$

where λ is the electron wavelength. These conventions make the reciprocal-space propagator unambiguous and allow the small-angle Fresnel form to be related directly to the more exact angular-spectrum form [20, 22].

A.2 From the electron wave equation to the multislice propagator

This section makes the propagation operator $\mathcal{P}_{\Delta z}$ mathematically explicit. For a fast electron travelling predominantly along the optical axis, the total wave may be written in envelope form as

$$\Psi(\mathbf{r}, z) = \psi(\mathbf{r}, z)\, e^{ikz}, \tag{A.2}$$

where the rapidly oscillating carrier e^{ikz} is separated from the slowly varying envelope $\psi(\mathbf{r}, z)$. Under the standard forward-scattering and paraxial-envelope approximation, the envelope obeys

$$\frac{\partial \psi}{\partial z} = \frac{i}{2k}\nabla_\perp^2 \psi + i\sigma V(\mathbf{r}, z)\,\psi = \frac{i\lambda}{4\pi}\nabla_\perp^2 \psi + i\sigma V(\mathbf{r}, z)\,\psi, \tag{A.3}$$

where $V(\mathbf{r}, z)$ is the electrostatic potential, $\nabla_\perp^2$ is the transverse Laplacian, and σ is the electron–matter interaction constant [21, 22]. Eq. (A.3) already displays the two physical processes that the multislice method separates: a local interaction term and a free-propagation term.

For a sufficiently thin slice extending from z_n to $z_n + \Delta z$, the potential term is approximated by a multiplicative transmission operator. Writing the slice-projected potential as

$$V_n^{\text{proj}}(\mathbf{r}) = \int_{z_n}^{z_n+\Delta z} V(\mathbf{r}, z)\, dz, \tag{A.4}$$

one obtains the standard elastic phase-grating form

$$O_n(\mathbf{r}) = \exp\left(i\sigma V_n^{\text{proj}}(\mathbf{r})\right). \tag{A.5}$$

In a purely elastic derivation, Eq. (A.5) is phase-only. In the present thesis, however, it is convenient to retain the more general transmission notation

$$O_n(\mathbf{r}) = A_n(\mathbf{r})e^{i\phi_n(\mathbf{r})}, \tag{A.6}$$

so that slice-wise amplitude and phase remain consistent with the object parameterisation used in Chapter 3 and the optimisation formulation developed later in Chapter 4. The phase-grating model is then recovered as the special case $A_n \equiv 1$.

If the potential term is momentarily dropped, Eq. (A.3) reduces to free propagation. In reciprocal space this propagation is diagonal, so the inter-slice propagator can be written as

$$\mathcal{P}_{\Delta z}[\psi](\mathbf{r}) = \mathcal{F}^{-1}\left[H(\mathbf{q}; \Delta z)\, \mathcal{F}\{\psi\}(\mathbf{q})\right](\mathbf{r}). \tag{A.7}$$

A natural exact choice is the angular-spectrum transfer function

$$H_{\text{AS}}(\mathbf{q}; \Delta z) = \exp\big(i\Delta z\, k_z\big), \qquad k_z = \sqrt{k^2 - k_x^2 - k_y^2}, \tag{A.8}$$

with $k_x = 2\pi q_x$ and $k_y = 2\pi q_y$. This form is especially useful because it makes the propagation physics explicit: each plane-wave component acquires the phase appropriate to its longitudinal wave-number [22].

Under the paraxial or small-angle approximation,

$$k_z \approx k - \frac{k_x^2 + k_y^2}{2k}, \tag{A.9}$$

so Eq. (A.8) reduces to the Fresnel form

$$H_{\mathrm{F}}(\mathbf{q}; \Delta z) \approx e^{ik\Delta z} \exp\left[-i\pi\lambda\Delta z \left(q_x^2 + q_y^2\right)\right]. \tag{A.10}$$

The prefactor $e^{ik\Delta z}$ is global phase and therefore has no effect on intensities, although it is harmless to keep. Eqs. (A.8) and (A.10) together explain operator notation, $\mathcal{P}_{\Delta z}$.

A.3 From single-slice interaction to multislice recursion

This section explains why the multislice recursion has the form stated in Eq. (3.8). For a probe centred at scan position $\mathbf{R}_j$, the incident wave is

$$\psi_j^{(0)}(\mathbf{r}) = P(\mathbf{r} - \mathbf{R}_j). \tag{A.11}$$

Within each slice, the interaction is represented by multiplication with the slice transmission, and between slices the wave is propagated by $\mathcal{P}_{\Delta z}$. It is therefore useful to separate one update into two substeps:

$$\tilde{\psi}_j^{(n)}(\mathbf{r}) = O_n(\mathbf{r})\, \psi_j^{(n)}(\mathbf{r}), \tag{A.12}$$

$$\psi_j^{(n+1)}(\mathbf{r}) = \mathcal{P}_{\Delta z}\left[\tilde{\psi}_j^{(n)}\right](\mathbf{r}). \tag{A.13}$$

Combining Eqs. (A.12)–(A.13) immediately yields the recursion:

$$\psi_j^{(n+1)}(\mathbf{r}) = \mathcal{P}_{\Delta z}\left[O_n(\mathbf{r})\, \psi_j^{(n)}(\mathbf{r})\right]. \tag{A.14}$$

This makes the physical content of the multislice model clearer than the single-slice picture. A single-slice description compresses the entire specimen into one effective transmission. Multislice, by contrast, asserts that the specimen acts as an ordered depth composition: the wave arriving at slice $n + 1$ already contains the cumulative effect of all earlier slice interactions and propagations[21, 22].

That depth ordering can be written explicitly by introducing the multiplication operator $\mathcal{M}_{O_n}[\psi] = O_n\psi$. If no additional free propagation is inserted after the final slice, then the exit wave after N_z slices is

$$\psi_j^{(N_z)} = \mathcal{M}_{O_{N_z-1}} \circ \left(\mathcal{P}_{\Delta z} \circ \mathcal{M}_{O_{N_z-2}}\right) \circ \cdots \circ \left(\mathcal{P}_{\Delta z} \circ \mathcal{M}_{O_0}\right)\psi_j^{(0)}. \tag{A.15}$$

The detector-plane wave and intensity are then formed as in Section 3.2,

$$\Psi_j(\mathbf{q}) = \mathcal{F}\{\psi_j^{(N_z)}\}(\mathbf{q}), \qquad I_j(\mathbf{q}) = \left|\Psi_j(\mathbf{q})\right|^2. \tag{A.16}$$

Eq. (A.15) is particularly useful for the argument of the thesis, because it shows explicitly that slice thickness, tilt-modified propagation, and slice-wise transmission all enter the forward model as operator ingredients rather than as post hoc metadata.

A.4 Slice thickness, sampling, and physical credibility

This section records the main numerical constraints that make the multislice approximation physically credible in practice. First, the slice thickness Δz cannot be chosen arbitrarily large. The transmission description inside each slice assumes that the potential can be compressed into a projected interaction over that slice; if Δz is too large, that approximation becomes less faithful and the separation between "interaction" and "propagation" becomes harder to justify [22]. Thus, slice thickness is not merely a discretisation parameter: it controls how plausible the local projection approximation remains.

Second, the propagation step of Eq. (A.7) is implemented through FFT-based convolution in reciprocal space. If reciprocal-space support is too narrow, or if the real-space / reciprocal-space grids are too coarsely matched to the wavefield structure, then the FFT-based propagation can wrap energy around the periodic computational window and generate aliasing artefacts. Ishizuka and Uyeda emphasised precisely this point in their practical multislice analysis, noting that adequate reciprocal-space support (or, equivalently, sufficiently many beams) is required for stable multislice computation [22]. For the present thesis, the main consequence is simple: increasing the number of slices and choosing adequate sampling are not cosmetic numerical choices, but part of making the multislice model both physically and computationally credible.

Chapter B

Supplementary Notes for Hyperparameter Optimisation in PtyRAD

This appendix supplements Chapter 3-4 by considering the optimisation problem that arises when the reconstruction configuration itself is treated as an object of inference. The goal is not to replace the inner ptychographic reconstruction analysed in the main text, but to clarify how the current PtyRAD implementation supports an outer hyperparameter-optimisation layer through Optuna. The discussion begins from notation and optimisation structure, then moves to the relevant algorithmic theory, and only afterwards turns to the implemented form used in the present code base.

B.1 Notation and optimisation levels

Chapter 4 formulates PtyRAD reconstruction as a physics-constrained nonconvex optimisation problem over the trainable parameter blocks Θ_{train}, together with the forward-participating but fixed quantities collected in Ξ_{forward}. In the notation of Eqs. (4.2)–(4.12), the inner reconstruction task is therefore to update physically meaningful variables such as object amplitude and phase, probe, position corrections, tilts, and slice thickness so that the model prediction becomes more consistent with the measured data [16, 30].

For the purposes of the present appendix, it is useful to introduce a second collection of variables,

$$\boldsymbol{\xi} \in \mathcal{X}, \tag{B.1}$$

where $\boldsymbol{\xi}$ denotes the hyperparameter or configuration vector of the reconstruction. The members of $\boldsymbol{\xi}$ are not updated by the inner-loop gradient calculation itself. Instead, they govern how the reconstruction problem is instantiated and solved. In the current PtyRAD hypertune pathway, representative members of $\boldsymbol{\xi}$ include optimiser choice, batch size, parameter-block learning rates, selected calibration variables, selected probe-model settings, selected geometry settings, and selected depth-discretisation settings [16, 30].

This distinction creates two optimisation levels. At the inner level, one fixes $\boldsymbol{\xi}$ and runs the iterative reconstruction analysed in Chapter 4. At the outer level, one compares different choices of $\boldsymbol{\xi}$. To make that distinction explicit, let

$$\widehat{\Theta}_{\mathrm{train},T}(\boldsymbol{\xi}) := \mathcal{A}_T(D;\boldsymbol{\xi}) \tag{B.2}$$

denote the reconstruction returned after T inner iterations on dataset D under configuration $\boldsymbol{\xi}$. Equation (B.2) should be read as a truncated reconstruction map: for each fixed configuration, the underlying iterative solve remains the problem over Θ_{train}, while the map $\mathcal{A}_T$ records the finite-budget output of that solve.

The present appendix therefore concerns not a new specimen-side inverse problem, but a second optimisation layer built around repeated evaluations of Eq. (B.2). This distinction is conceptually important. The inner loop updates physical variables by differentiating through the forward model; the outer loop compares reconstruction configurations by observing how different settings alter the output of the inner loop.

B.2 From reconstruction objective to hyperparameter objective

For fixed $\boldsymbol{\xi}$, the inner reconstruction remains the objective-based problem described in Section 4.2. In principle, one might imagine an ideal outer target

$$\mathscr{F}^{\star}(\boldsymbol{\xi}) := Q\Big(\widehat{\Theta}_{\mathrm{train},\infty}(\boldsymbol{\xi}), D\Big), \tag{B.3}$$

where Q denotes some scientifically satisfactory notion of final reconstruction quality. In experimental ptychography, however, Eq. (B.3) is generally not observable in any exact sense. Ground-truth object fields are unavailable, full convergence may be computationally unaffordable, and even a visually plausible reconstruction may remain ambiguous in its internal allocation of modelling freedom. The exact outer target is therefore not directly accessible.

What is accessible is an implemented finite-budget surrogate,

$$\widetilde{\mathscr{F}}_T(\boldsymbol{\xi}) := M_T\Big(\widehat{\Theta}_{\mathrm{train},T}(\boldsymbol{\xi}), D\Big), \tag{B.4}$$

where M_T is a scalar trial-level metric evaluated after T reconstruction iterations. In the

current PtyRAD hypertune implementation, this scalar is chosen from a small set of proxy metrics rather than from the layered evaluation framework used in Chapter 5. The outer search problem may therefore be written schematically as

$$\boldsymbol{\xi}_T^\star \in \arg\min_{\boldsymbol{\xi}\in\mathcal{X}} \widetilde{\mathscr{F}}_T(\boldsymbol{\xi}). \tag{B.5}$$

Equation (B.5) already explains why the present setting is not naturally handled by the same tools as the inner problem. The map $\boldsymbol{\xi} \mapsto \widetilde{\mathscr{F}}_T(\boldsymbol{\xi})$ is expensive to evaluate because each function call requires a substantial reconstruction run involving forward simulation, backpropagation, optimiser updates, and iteration-wise constraints. It is also naturally black-box from the viewpoint of the outer loop: even though the inner reconstruction is differentiable with respect to Θ_{train}, the outer target is not presented to the user as a conveniently differentiable function of all configuration choices. Finally, Eq. (B.5) is budget-dependent: two configurations may be ranked differently at iteration 20 and iteration 200. These three properties, costly evaluation, black-box access, and finite-budget dependence, are precisely what motivate sequential hyperparameter optimisation methods rather than naive exhaustive search or direct calculus on $\boldsymbol{\xi}$ [46–49].

B.3 Sampler theory

A useful conceptual baseline for outer-loop search is random search. Bergstra and Bengio argue that in many hyperparameter problems only a subset of the variables materially affects performance, and that the identity of this subset may differ from one dataset to another [46]. In such settings, grid search spends too much of its budget resolving coordinates that may not matter, whereas random search can cover more relevant directions under the same trial budget. This point is directly relevant to PtyRAD, because the search space may contain optimiser-family choices, multiple learning rates, batch-geometry-related scales, calibration variables, and model-complexity controls whose relative importance need not remain fixed across reconstruction regimes.

Random search is nevertheless only a baseline. In a sequential setting, one would like later trials to depend on information obtained from earlier ones. Let

$$\mathcal{H}_n := \left\{\left(\boldsymbol{\xi}^{(i)}, y^{(i)}\right)\right\}_{i=1}^{n}, \qquad y^{(i)} = \widetilde{\mathscr{F}}_T\left(\boldsymbol{\xi}^{(i)}\right), \tag{B.6}$$

denote the trial history after n completed evaluations. Tree-structured Parzen estimation (TPE) is one way of converting $\mathcal{H}_n$ into an adaptive proposal rule [50]. Rather than modelling the response surface directly as $\mathsf{P}(y \mid x)$ for each parameter coordinate x, the TPE reformulation models the inverse conditionals

$$\ell(x) = \mathsf{P}\big(x \mid y < y^\star\big), \qquad g(x) = \mathsf{P}\big(x \mid y \geq y^\star\big), \tag{B.7}$$

where $y^\star$ is a threshold induced by a chosen quantile over the observed trial values. The first density collects parameter values associated with comparatively good trials; the second collects the remainder.

The intuition of Eq. (B.7) is straightforward. If a candidate value x is far more likely under ℓ than under g, then the observed trial history suggests that configurations containing that value tend to lie disproportionately in the better-performing region of the search. The expected-improvement objective can be rewritten, up to factors that depend on the chosen quantile but not on the candidate ranking, in a form whose maximiser is monotone in $\ell(x)/g(x)$ [50]. Thus, the familiar density-ratio rule is not merely heuristic shorthand; it is a re-expression of an acquisition principle in which promising candidates are those that are comparatively concentrated in the good-trial density.

This theory also clarifies what Optuna's `TPESampler` is doing at implementation level. In the library's own description, TPE fits one density model to the set of parameter values associated with the best objective values and another to the remaining values, then selects candidates that maximise the ratio between the two [51]. Optuna further distinguishes between *independent* sampling and *relative* sampling: the former samples each parameter without modelling inter-parameter dependence, whereas the latter samples multiple parameters jointly so that correlations can be used [52].

B.4 Budgeted trials, Successive Halving, Hyperband-style pruning

The second theoretical ingredient is budget allocation. In the present setting, each trial is not an instantaneous scalar evaluation but an iterative reconstruction. One may therefore regard the trial budget as a resource variable

$$r \in \{1, 2, \ldots, R\}, \tag{B.8}$$

where r counts the amount of inner-loop computation already spent on the current configuration and R denotes the maximum budget allowed to a full trial. In the current PtyRAD implementation, the natural resource is simply the reconstruction iteration count, because intermediate metrics are reported once per iteration [30].

This perspective leads naturally to early-stopping methods. Jamieson and Talwalkar show that, when candidate configurations can be evaluated progressively, hyperparameter optimisation may be cast as a resource-allocation problem rather than as repeated full-budget evaluation [47]. Under a reduction factor $\eta > 1$, a Successive-Halving-style schedule may be written schematically as

$$r_s = \eta^s r_0, \qquad n_s \approx n_0 \eta^{-s}, \tag{B.9}$$

where n_0 configurations begin at a small initial budget r_0, and only a shrinking subset survives to later stages. Equation (B.9) captures the essential trade-off: poor configurations are stopped early so that more budget can be concentrated on promising ones.

Hyperband generalises this idea by running several such schedules with different initial (n_0, r_0) trade-offs under a common overall budget [48]. The point is not that one universal early-stopping schedule is best, but that expensive black-box optimisation often benefits from allocating resources across both breadth and depth. Some brackets favour many shallow trials; others favour fewer, deeper trials.

For PtyRAD, this logic is particularly natural because a reconstruction trial already exposes an intermediate scalar after each iteration. Pruning is therefore not an external add-on layered onto a static objective; it is an operation made possible by the iterative structure of the reconstruction itself. In Optuna's implementation, this is expressed through `Trial.report()` and `Trial.should_prune()`, which use intermediate values to decide whether a trial should continue [53, 54]. In a ptychographic setting, this means that the outer search does not need to spend the full budget on every configuration before identifying evidently weak candidates.

B.5 Implemented Optuna logic in PtyRAD

The current PtyRAD code base instantiates the preceding theory in a concrete but intentionally restricted form. At the highest level, `PtyRADSolver.run()` chooses between or-

dinary reconstruction mode and hypertune mode through the boolean flag `if_hypertune`. When hypertune mode is active, the solver enters `PtyRADSolver.hypertune()`, where the sampler, pruner, study name, storage path, trial budget, and chosen scalar error metric are read from `hypertune_params`. A study is then created in minimisation mode through `optuna.create_study(...)` [55].

Within that study, one trial corresponds to one reconstruction run under one proposed configuration. The core logic resides in `optuna_objective()`, which first copies the parameter dictionary and then injects the current trial suggestions into selected members of the reconstruction configuration. In the current explicit injection path, these trial-controlled families include batch size, optimiser choice, multiple learning rates, selected calibration terms such as `dx`, selected probe-side quantities such as `pmode_max`, `conv_angle`, and `z_shift`, selected aberration coefficients, selected depth variables such as `Nlayer` and `dz`, scan-affine corrections, and global tilts [30]. After the necessary re-initialisation steps, the code rebuilds the `PtychoModel`, constructs the optimiser, and prepares the indices and batches exactly as in an ordinary reconstruction run.

The inner loop that follows is therefore not a distinct solver. It reuses the reconstruction machinery already analysed in the main text: forward simulation, loss evaluation, backpropagation, optimiser update, and iteration-wise constraints are executed through the same reconstruction pathway. This point is conceptually important. The trial does not replace reconstruction; it repeatedly evaluates differently configured reconstructions.

The metric semantics are also worth stating carefully. The helper `compute_optuna_error()` currently supports two trial-level objectives. When `error_metric = loss`, the objective returned to Optuna is the latest scalar loss stored in `model.loss_iters`. When `error_metric = contrast`, the code evaluates a contrast proxy derived from the cropped, z-summed object phase and returns its negative so that the study remains a minimisation problem [30]. The outer objective is therefore not a full replacement for layered evidence framework of Chapter 5; it is a deliberately compressed scalar proxy.

Pruning enters directly through the iterative nature of the trial. After each reconstruction iteration, the current trial metric is reported to Optuna. If the active pruner judges the trial unpromising, the code raises `TrialPruned` and terminates that configuration early; otherwise the trial continues until the iteration budget is exhausted. Results may then be stored both in the Optuna study database and, when enabled, in a collated filesystem view built around the hypertune output directory [30, 53].

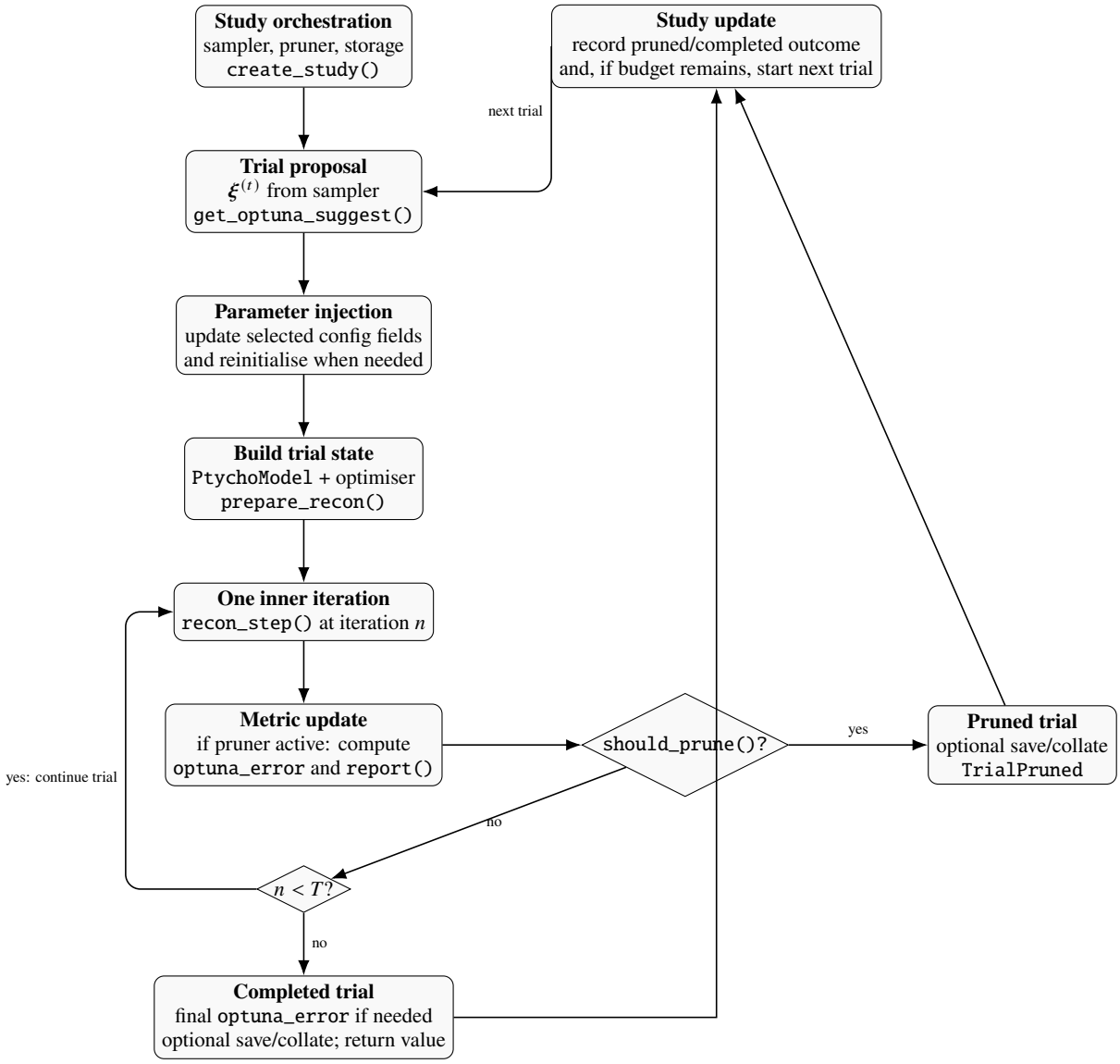


Figure B-1. Implemented Optuna pathway in PtyRAD. A study first creates and maintains the sampler, pruner, storage, and trial history. Each trial proposes a configuration vector $\boldsymbol{\xi}^{(t)}$, injects the selected values into the reconstruction configuration, rebuilds the trial state, and then enters the inner reconstruction loop. After each inner iteration, an intermediate scalar is reported when pruning is active. A pruned trial terminates immediately and returns control to the study; a non-pruned trial either continues to the next inner iteration or, when the iteration budget is exhausted, returns a completed trial value.

Chapter C

Supplementary Notes for Data-Fidelity Terms

This appendix collects the derivational remarks deferred from Section 4.3.

C.1 Notation and averaging conventions

For each scan index $j = 1, \dots, N$ and detector pixel $k = 1, \dots, K$, let

$$I_{j,k}^{\text{meas}} \in \mathbb{N}_0, \qquad I_{j,k}^{\text{model}}(\Theta) \in \mathbb{R}_+, \tag{C.1}$$

where $I_{j,k}^{\text{meas}}$ denotes the measured detector count (or detector intensity after the chosen preprocessing) at scan position j and detector pixel k, and $I_{j,k}^{\text{model}}(\Theta)$ denotes the corresponding forward prediction.

C.2 Gaussian-like term: least-squares interpretation

C.2.1 Probabilistic model on the comparison variable

Define the transformed comparison variables

$$Y_{j,k}^{\text{meas}} := \mathcal{T}_{p_G}\left(I_{j,k}^{\text{meas}}\right), \qquad Y_{j,k}^{\text{model}}(\Theta) := \mathcal{T}_{p_G}\left(I_{j,k}^{\text{model}}(\Theta)\right). \tag{C.2}$$

A statistical anchor for Gaussian-like term is the transformed-domain observation model

$$Y_{j,k}^{\text{meas}} = Y_{j,k}^{\text{model}}(\Theta) + \varepsilon_{j,k}, \qquad \varepsilon_{j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \sigma_G^2\right). \tag{C.3}$$

Conditional independence across (j, k) then gives the likelihood

$$p := \mathsf{P}\left(\left\{Y_{j,k}^{\text{meas}}\right\} \mid \Theta, \sigma_G^2\right) = \prod_{j=1}^{N} \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi\sigma_G^2}} \exp\left[-\frac{\left(Y_{j,k}^{\text{meas}} - Y_{j,k}^{\text{model}}(\Theta)\right)^2}{2\sigma_G^2}\right]. \tag{C.4}$$

Taking the negative logarithm yields

$$-\log p = \frac{NK}{2}\log(2\pi\sigma_G^2) + \frac{1}{2\sigma_G^2}\sum_{j=1}^{N}\sum_{k=1}^{K}\left(Y_{j,k}^{\text{meas}} - Y_{j,k}^{\text{model}}(\Theta)\right)^2. \tag{C.5}$$

The first term is constant with respect to Θ, and the prefactor $(2\sigma_G^2)^{-1}$ is a positive constant. Therefore,

$$\arg\min_{\Theta}\left[-\log p\right] = \arg\min_{\Theta}\sum_{j=1}^{N}\sum_{k=1}^{K}\left|\mathcal{T}_{p_G}\left(I_{j,k}^{\text{model}}(\Theta)\right) - \mathcal{T}_{p_G}\left(I_{j,k}^{\text{meas}}\right)\right|^2. \tag{C.6}$$

This is precisely least-squares matching on the transformed comparison variable. Using the averaging notation, the same minimiser is obtained from

$$\left\langle\left|\mathcal{T}_{p_G}\left(I_{j,k}^{\text{model}}(\Theta)\right) - \mathcal{T}_{p_G}\left(I_{j,k}^{\text{meas}}\right)\right|^2\right\rangle_{j,k}. \tag{C.7}$$

C.2.2 From least-squares to the implemented Gaussian-like objective

The Gaussian-like term used in the main text is written as

$$L_G(\Theta) := \frac{\left\langle\left|\mathcal{T}_{p_G}\left(I_{j,k}^{\text{model}}(\Theta)\right) - \mathcal{T}_{p_G}\left(I_{j,k}^{\text{meas}}\right)\right|^2\right\rangle_{j,k}^{1/2}}{\left\langle\mathcal{T}_{p_G}\left(I_{j,k}^{\text{meas}}\right)\right\rangle_{j,k}}. \tag{C.8}$$

Two simple observations justify this form.

(a) The square root does not change the minimiser, because $x \mapsto x^{1/2}$ is strictly increasing for $x \geq 0$.

(b) The denominator depends only on the measured data, not on Θ, so it is a positive scale factor that normalises the numerical value of the loss but does not alter the minimising argument.

Hence Eq. (C.8) has the same minimiser as Eqs. (C.6)–(C.7). In this precise sense, the implemented Gaussian-like term is a normalised transformed-domain RMSE whose optimisation target is still transformed-domain least-squares [14, 16].

C.2.3 Special cases and statistical interpretation

Two special choices of p_G are particularly useful.

(a) $p_G = 1$**.** Then $\mathcal{T}_{p_G}$ is the identity, so Eq. (C.8) reduces to a normalised intensity-domain least-squares objective. In this case the exact statistical model is additive Gaussian noise on the raw intensities themselves.

(b) $p_G = \frac{1}{2}$**.** Then the comparison is carried out in an amplitude-like or square-root domain. This is not the exact Poisson likelihood, but it admits an approximate variance-stabilising interpretation. Indeed, if $X \sim \text{Poisson}(\lambda)$ and $g(x) = x^{1/2}$, then the delta method gives

$$\text{Var}\big(g(X)\big) \approx \big(g'(\lambda)\big)^2 \text{Var}(X) = \left(\frac{1}{2\sqrt{\lambda}}\right)^2 \lambda = \frac{1}{4}. \tag{C.9}$$

Thus, for sufficiently large counts, the square-root transformed Poisson variable has approximately constant variance. This provides a statistical rationale for square-root-domain least-squares, even though it should still be described as an approximation rather than as an exact Poisson model.

The main honesty point is therefore the following: the Gaussian-like term is exact if one explicitly posits a Gaussian observation model on the transformed quantity $\mathcal{T}_{p_G}(I)$; when $p_G = \frac{1}{2}$, it may also be read as an approximate variance-stabilised comparison for Poisson data. What it is *not* is an exact Poisson likelihood on the raw detector counts.

C.3 Poisson-like term: exact likelihood and surrogate interpretation

C.3.1 Exact Poisson likelihood on raw counts

A standard counting-statistics model assumes conditional independence and

$$I_{j,k}^{\text{meas}} \mid \Theta \sim \text{Poisson}\big(\lambda_{j,k}(\Theta)\big)\,, \qquad \lambda_{j,k}(\Theta) := I_{j,k}^{\text{model}}(\Theta). \tag{C.10}$$

Under this model,

$$p^* := \mathsf{P}\Big(\Big\{I_{j,k}^{\text{meas}}\Big\} \mid \Theta\Big) = \prod_{j=1}^{N}\prod_{k=1}^{K} \frac{e^{-\lambda_{j,k}(\Theta)}\lambda_{j,k}(\Theta)^{I_{j,k}^{\text{meas}}}}{I_{j,k}^{\text{meas}}!}. \tag{C.11}$$

Taking the negative logarithm gives

$$-\log p^* = \sum_{j=1}^{N}\sum_{k=1}^{K}\left[\lambda_{j,k}(\Theta) - I_{j,k}^{\text{meas}}\log\lambda_{j,k}(\Theta) + \log\left(I_{j,k}^{\text{meas}}!\right)\right]. \tag{C.12}$$

Because the factorial term depends only on the data, it may be discarded without changing the optimiser. The exact Poisson negative log-likelihood is therefore, up to additive constants,

$$\mathcal{L}_{\text{Pois}}^{\text{exact}}(\Theta) := \sum_{j=1}^{N}\sum_{k=1}^{K}\left[I_{j,k}^{\text{model}}(\Theta) - I_{j,k}^{\text{meas}}\log I_{j,k}^{\text{model}}(\Theta)\right]. \tag{C.13}$$

This is the strict statistical anchor for a Poisson-based ptychographic objective [14, 31].

A useful derivative identity is

$$\frac{\partial\mathcal{L}_{\text{Pois}}^{\text{exact}}}{\partial I_{j,k}^{\text{model}}} = 1 - \frac{I_{j,k}^{\text{meas}}}{I_{j,k}^{\text{model}}}. \tag{C.14}$$

This makes the asymmetry of the Poisson model explicit: the gradient is not a simple residual, and under-predicted measured counts are penalised differently from over-predicted ones.

C.3.2 Relation to the implemented Poisson-like objective

The Poisson-like term used in the main text is

$$L_P(\Theta) := -\frac{\left\langle \mathcal{T}_{p_P}\left(I_{j,k}^{\text{meas}}\right)\log\left(\mathcal{T}_{p_P}\left(I_{j,k}^{\text{model}}(\Theta)\right)+\varepsilon\right) - \mathcal{T}_{p_P}\left(I_{j,k}^{\text{model}}(\Theta)\right)\right\rangle_{j,k}}{\left\langle \mathcal{T}_{p_P}\left(I_{j,k}^{\text{meas}}\right)\right\rangle_{j,k}}, \qquad \varepsilon > 0. \tag{C.15}$$

This expression should be compared carefully with Eq. (C.13).

(a) If $p_P = 1$, if $\varepsilon \to 0$, and if the denominator is ignored, then Eq. (C.15) reduces to the exact Poisson negative log-likelihood up to the harmless factor $(NK)^{-1}$ coming from the average in $\langle\cdot\rangle_{j,k}$.

(b) The denominator $\langle\mathcal{T}_{p_P}(I_{j,k}^{\text{meas}})\rangle_{j,k}$ depends only on the data, so it rescales the loss but does not change the minimiser.

(c) The stabilising constant ε is a numerical device that prevents the logarithm from

diverging when the predicted intensity approaches zero. It is useful computationally, but it is not part of the ideal Poisson model itself.

(d) If $p_P \neq 1$, the loss is no longer an exact likelihood on the raw counts. In that case it is best described as a Poisson-motivated surrogate objective defined on the transformed intensities.

The mathematically honest summary is therefore: Eq. (C.15) inherits its statistical anchor from Eq. (C.13), but it becomes exact only in the untransformed limit $p_P = 1$ with vanishing stabilisation and data-only normalisation removed.

C.3.3 Local weighted least-squares approximation

The exact Poisson objective is often compared with least-squares because the Poisson negative log-likelihood admits a local quadratic expansion near a good fit. Fix a measured count $y \geq 0$ and define

$$f_y(\lambda) := \lambda - y \log \lambda, \qquad \lambda > 0. \tag{C.16}$$

Then

$$f_y'(\lambda) = 1 - \frac{y}{\lambda}, \qquad f_y''(\lambda) = \frac{y}{\lambda^2}. \tag{C.17}$$

When λ is expanded around the data-consistent point $\lambda = y$ (for $y > 0$), one has $f_y'(y) = 0$ and $f_y''(y) = 1/y$, so Taylor expansion gives

$$f_y(\lambda) = f_y(y) + \frac{1}{2y}(\lambda - y)^2 + O\Big((\lambda - y)^3\Big). \tag{C.18}$$

Summing over (j, k) shows that, near a good fit,

$$\mathcal{L}_{\text{Pois}}^{\text{exact}}(\Theta) \approx \text{const} + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{K} \frac{\left(I_{j,k}^{\text{model}}(\Theta) - I_{j,k}^{\text{meas}}\right)^2}{I_{j,k}^{\text{meas}}}. \tag{C.19}$$

Thus Poisson likelihood is locally related to a *weighted* least-squares model whose weights are inversely proportional to the counts. This clarifies both the connection and the distinction: the Gaussian one is exact least-squares under a Gaussian model on the chosen comparison variable, whereas the Poisson one is an exact counting-statistics likelihood whose local quadratic approximation resembles weighted least-squares [31].

Chapter D

Additional Figures for Chapter 6

D.1 Supplementary Figures for Section 6.3

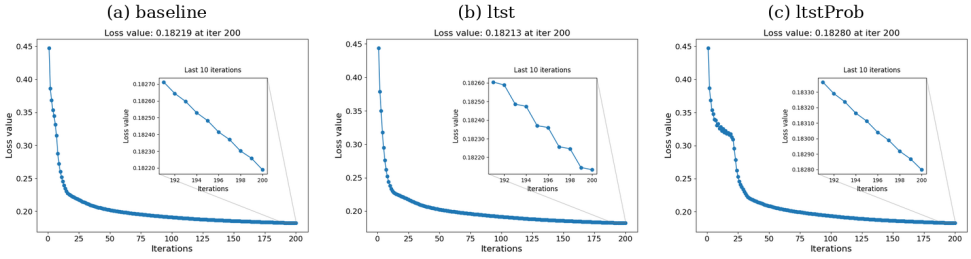


Figure D-1. Loss trajectories under three staged-activation schedules at the common iteration-200 horizon. Panels (a)–(c) correspond to the baseline, the early-probe / late-geometry variant (`ltst`), and the object-only warm-up variant (`ltstProb`), respectively. The inset in each panel enlarges the final ten iterations to make the late-stage behaviour more visible.

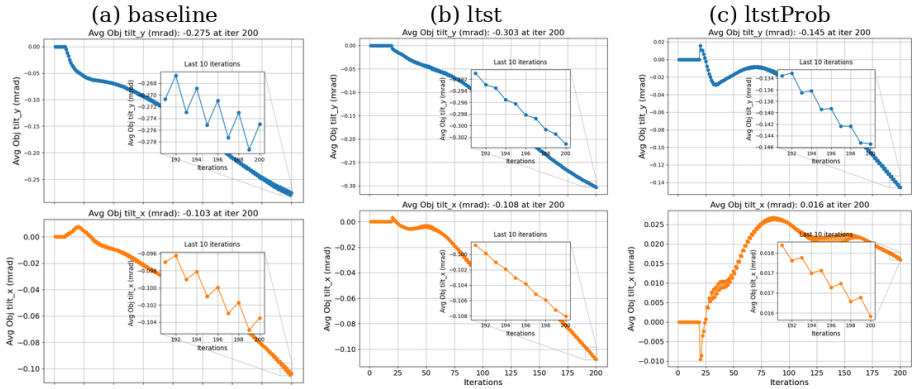


Figure D-2. Average object-tilt trajectories under the three staged-activation schedules. Panels (a)–(c) correspond to the baseline, `ltst`, and `ltstProb`, respectively. Panel (a) reproduces the baseline average-tilt diagnostic, while panels (b) and (c) show the two schedule variants. In each panel, the upper curve reports the average θ_y component and the lower curve the average θ_x component, with an inset enlarging the final ten iterations.

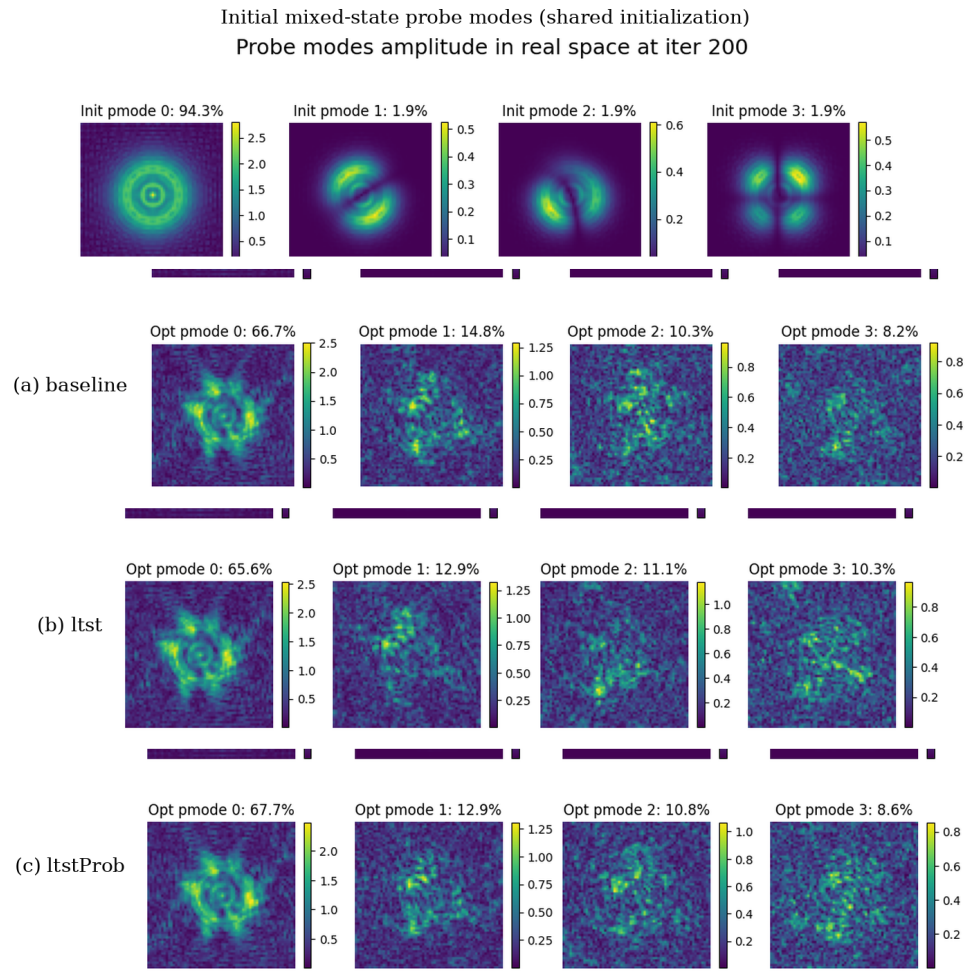


Figure D-3. Real-space probe-mode amplitudes after optimisation under the three staged-activation schedules. Panels (a)–(c) correspond to the baseline, `ltst`, and `ltstProb`, respectively. In each panel, the initial mixed-state probe is shown above and the optimised probe modes at iter 200 below, with modal occupancies reported above each mode. The comparison is intended to show that mixed-state probe redistribution is sensitive to activation timing, even when the final detector-side losses remain close.

D.2 Supplementary Figures for Section 6.4

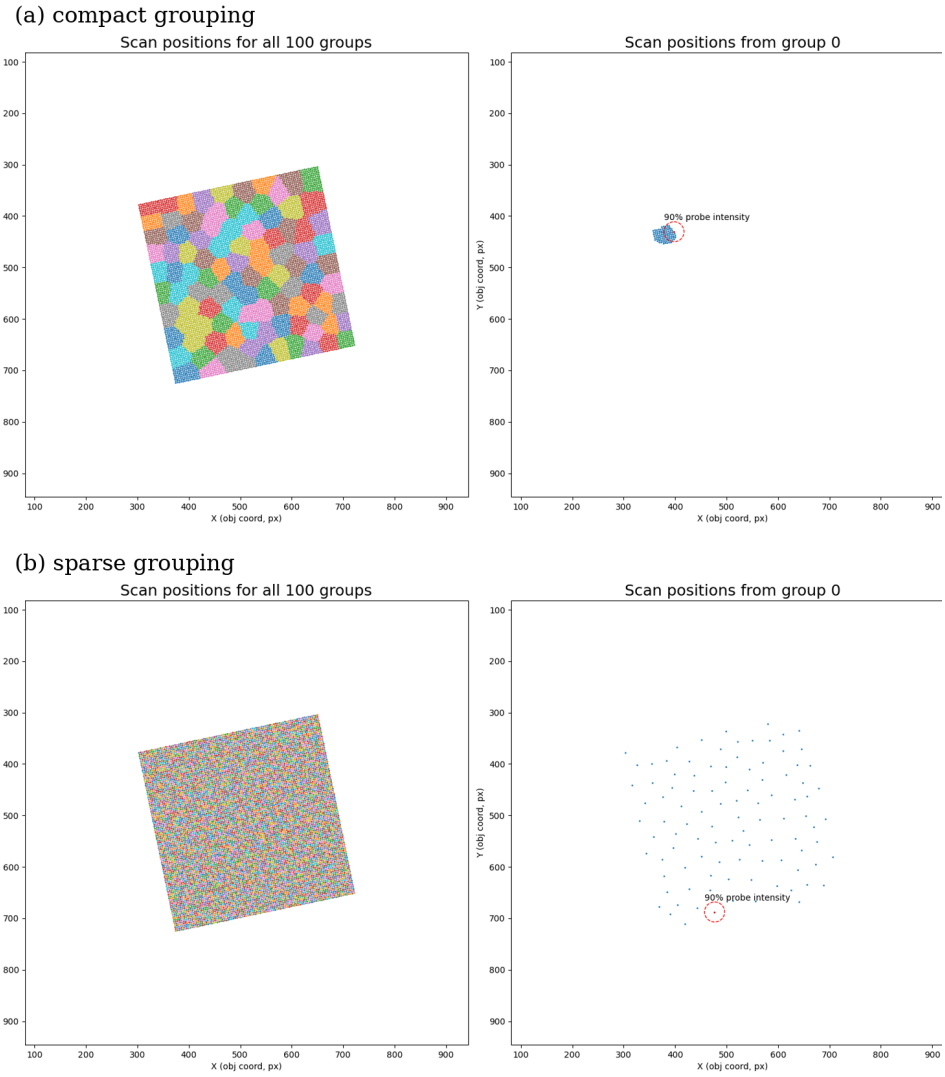


Figure D-4. Alternative grouping geometries tested in the H4 comparison. The random baseline grouping is the same as Fig. 5-2. The upper row shows the compact grouping, while the lower row shows the sparse grouping. In each row, the left panel shows all groups over the active scan subset, and the right panel shows one representative group together with the same 90% probe-intensity footprint marker used in the baseline diagnostic.

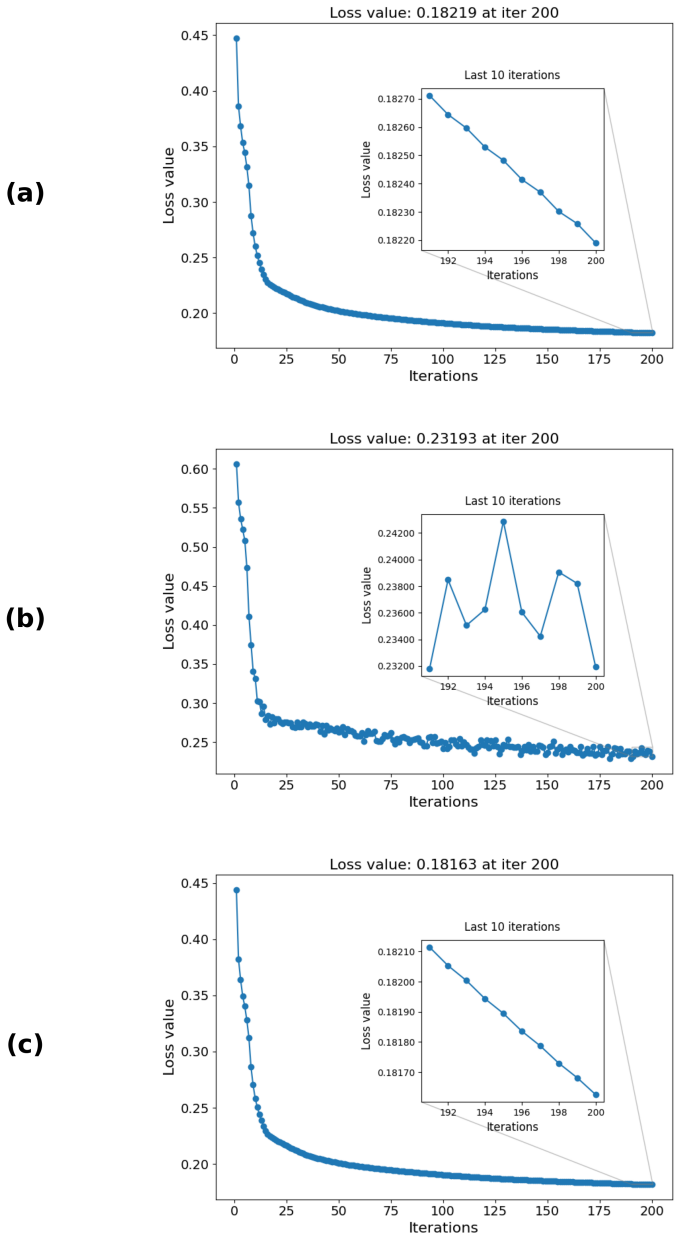


Figure D-5. Loss trajectories under different batch geometries at the common iteration-200 horizon. Panels (a)–(c) correspond to the random baseline, compact grouping, and sparse grouping, respectively. The inset in each panel enlarges the final ten iterations in order to make the late-stage behaviour more visible.

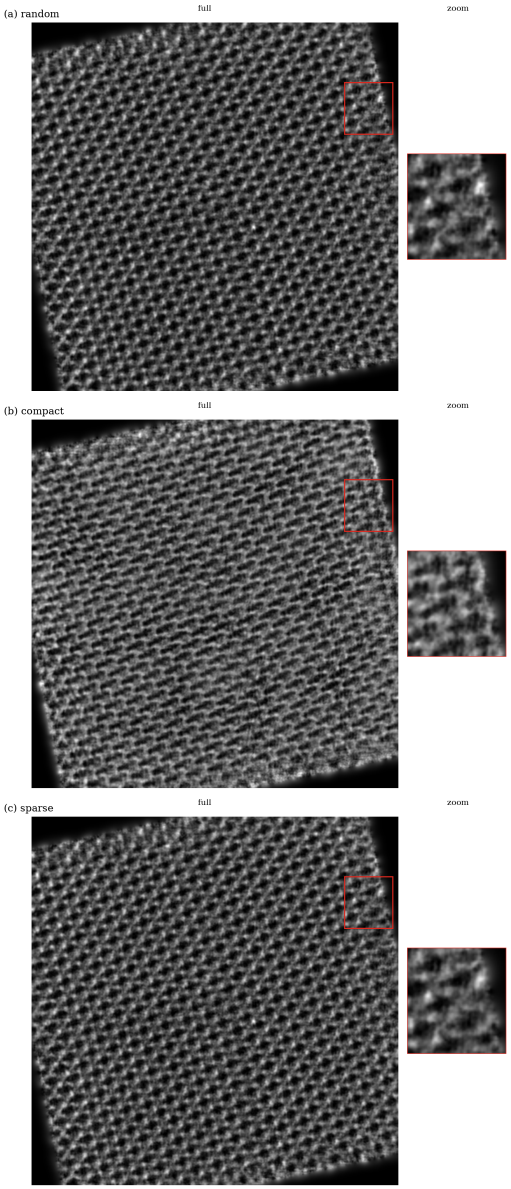


Figure D-6. Cropped z-summed object phase under different batch geometries, with a shared zoom region in the upper-right part of the reconstructed field. Panels (a)–(c) correspond to the random baseline, compact grouping, and sparse grouping, respectively. The same ROI is used in all three panels to support direct local comparison.