Development and Validation of a Physics-Informed Neural Network for the Reconstruction of Background-Oriented Schlieren Measurements

Tesi di Laurea Magistrale in
Mechanical Engineering - Ingegneria Meccanica

Author: **Giovanni Brugnola**

Student ID: 244776
Advisor: Prof. Paolo Gaetani
Co-advisor: Prof. Fabrizio Fontaneto
Academic Year: 2025-2026

Abstract

The push for more efficient, compact aero-engines has renewed interest in transonic compressors, where intrusive probes distort the very shock structure they are meant to measure. Background-Oriented Schlieren (BOS) offers a non-intrusive alternative, but its signal responds to the density *gradient*, not the density itself: any uniform shift leaves it unchanged, so the absolute density level remains undetermined, a null-space indeterminacy inherent to the technique.

This thesis develops and validates a Unified BOS/Physics-Informed Neural Network framework (UBOS/PINN) that embeds the BOS measurement operator and the compressible Euler equations in a single differentiable optimisation problem, inferring the density field from a single BOS image pair together with a minimal set of supplemental measurements. It asks which minimal subset of measurable information lifts the indeterminacy when the full inlet boundary condition is unavailable. Validation uses a synthetic compressible Rankine vortex with known closed-form solution, through three models of decreasing reliance on synthetic boundary information: Model 1, with ideal ground-truth anchoring; Model 2, the conceptual core, constrained only by a total-temperature field and by sparse static-pressure taps; and Model 3, a gauge-invariant formulation isolating the optical gradient alone.

Model 2 suggests that the information content of the measurable set $\{T_0, p_s\}$ (an inlet total temperature and a few static-pressure taps), combined with the isentropic EOS closure, is in principle sufficient to lift this indeterminacy. On the validation case an adaptive weighting heuristic (AWH) recovers the absolute density level to a relative L_2 error of order 10^{-3}, whereas the purely measurable fixed-weight configuration leaves a substantially larger residual offset. A fully unsupervised, measurable-only pipeline therefore remains to be demonstrated.

By contrast, the optical channel alone (Model 3) constrains only the gradient structure of the reconstructed field, reaching a relative L_2 error of order 10^{-2} on the density gradient and thereby locating the boundary between optical measurement and thermodynamic inference. A dedicated ablation on the shared training strategy further shows that the BOS-weighted importance sampler improves the gradient-channel accuracy by more than an order of magnitude relative to uniform sampling. Because the synthetic Rankine vortex is smooth, subsonic and shock-free, it does not exercise the flow discontinuities of the target transonic regime; verification on a case containing shocks is deferred to future work.

Keywords: Background-Oriented Schlieren; Physics-Informed Neural Networks; compressible flow; inverse problem; density reconstruction.

Abstract in lingua italiana

La spinta verso motori aeronautici più efficienti e compatti ha rinnovato l'interesse per i compressori transonici, dove le sonde intrusive distorcono la stessa struttura d'urto che dovrebbero misurare. La Schlieren a Sfondo Strutturato (BOS, Background-Oriented Schlieren) offre un'alternativa non intrusiva, ma il suo segnale risponde al *gradiente* di densità, non alla densità stessa: qualsiasi traslazione uniforme la lascia invariata, cosicché il livello assoluto di densità rimane indeterminato, un'indeterminazione di spazio nullo intrinseca alla tecnica.

Questa tesi sviluppa e valida un framework unificato BOS/Rete Neurale Informata dalla Fisica (UBOS/PINN) che incorpora l'operatore di misura BOS e le equazioni di Eulero comprimibili in un unico problema di ottimizzazione differenziabile, inferendo il campo di densità da una singola coppia di immagini BOS unitamente a un insieme minimo di misure supplementari. Essa si interroga su quale sottoinsieme minimo di informazione misurabile rimuova l'indeterminazione quando la condizione al contorno completa in ingresso non è disponibile. La validazione impiega un vortice di Rankine comprimibile sintetico con soluzione nota in forma chiusa, attraverso tre modelli con dipendenza decrescente dall'informazione di contorno sintetica: il Modello 1, con ancoraggio ideale al dato di riferimento (ground-truth); il Modello 2, il nucleo concettuale, vincolato unicamente da un campo di temperatura totale e da prese di pressione statica sparse; e il Modello 3, una formulazione invariante di gauge che isola il solo gradiente ottico.

Il Modello 2 suggerisce che il contenuto informativo dell'insieme misurabile $\{T_0, p_{s,k}\}$ sia in linea di principio sufficiente a rimuovere questa indeterminazione. Nel caso di validazione un'euristica di pesatura adattiva (AWH, adaptive weighting heuristic) recupera il livello assoluto di densità con un errore L_2 relativo dell'ordine di 10^{-3}, mentre la configurazione a peso fisso puramente misurabile lascia un offset residuo sostanzialmente più ampio. Una pipeline completamente non supervisionata e basata solo su grandezze misurabili resta dunque da dimostrare.

Per contro, il solo canale ottico (Modello 3) vincola unicamente la struttura del gradiente del campo ricostruito, raggiungendo un errore L_2 relativo dell'ordine di 10^{-2} sul gradiente di densità; uno studio di ablazione dedicato mostra che la strategia di campionamento per importanza pesato sulla BOS migliora l'accuratezza del canale del gradiente di quasi un ordine di grandezza rispetto al campionamento uniforme, quantificando il confine tra misura ottica e inferenza termodinamica. Poiché il vortice di Rankine sintetico è regolare, subsonico e privo di urti, esso non sollecita le discontinuità di flusso del regime transonico obiettivo; la verifica su un caso contenente urti è rimandata a lavori futuri.

Parole chiave: Background-Oriented Schlieren; Physics-Informed Neural Networks; flussi comprimibili; problema inverso; ricostruzione della densità.

Contents

1 | Introduction

Section 1.1 traces the argument from the decarbonisation mandate reshaping civil aviation to the measurement challenge that motivates Background-Oriented Schlieren (BOS). Section 1.2 states the central research question, recalling the BOS campaign carried out at the von Karman Institute for Fluid Dynamics (VKI) on a supersonic compressor cascade, and Section 1.3 outlines how the remaining chapters answer it.

1.1. Motivation

The growth of air-transport demand now collides with tightening environmental expectations: emissions and fuel-consumption limits have become the primary criteria against which a new engine is judged. Manufacturers respond by attacking thermodynamic efficiency on two fronts. The first raises Brayton-cycle efficiency by increasing the overall pressure ratio at the compressor outlet; achieving this without enlarging the engine shrinks the core flow channel and drives the compressor blade height down. The second raises propulsive efficiency by increasing the bypass ratio: a slowly turning geared fan produces thrust by accelerating a large air mass to modest velocity, yielding in its most aggressive form open-fan architectures such as the CFM RISE demonstrator, with bypass ratios beyond 70 [12]. These two strategies are coupled: a large slow fan must be driven by a core compact enough not to negate the bypass gain, and a compact core delivers the required work only by rotating faster, since by the Euler turbomachinery relation the specific work per stage grows with blade speed. The relative tip speeds of the inner compressor stages are thereby pushed beyond the speed of sound. The transonic and supersonic compressor, once a niche object, has consequently returned to the centre of engine development, and with it the need to validate aggressive new blade designs experimentally.

Validating such machines is where conventional testing practice begins to fail. As the core channel narrows, any physical probe inserted into it occupies a non-negligible fraction of the passage: the resulting blockage throttles the stage, displaces its operating point, and distorts the flow field the probe was meant to characterise. In a supersonic passage the intrusion is worse still: the probe generates its own bow shock, contaminating the shock pattern that is the principal object of the measurement. Making probes ever smaller encounters a hard limit: below a certain size, bandwidth and spatial resolution can no longer be preserved together, and the harsh temperature, pressure and contamination of a representative test environment erode what robustness remains. Non-intrusive diagnostics therefore become a necessity, and optical techniques, reading the flow without touching it, are the natural candidates.

A quantitative description of a compressible flow requires, beyond velocity, at least one independent state variable. Density is an especially convenient choice because it can be inferred optically: in a transparent gas the refractive index depends linearly on density through the Gladstone–Dale relation, so any density gradient deflects a light ray by a measurable amount [31, 37]. Among the diagnostics exploiting this principle, BOS is increasingly adopted in industrial facilities: it replaces the precision optics of classical Schlieren with a single camera and a structured background, so the field of view is no longer bounded by any optical component [31], and each pixel of the deformed (wind-on) image yields a local, quantitative estimate of the line-of-sight-integrated density gradient.

These advantages come at the cost of three structural limitations, formalised in Chapter 2, that make density reconstruction ill-posed: the line-of-sight integration, the sensitivity to the density *gradient* alone (any uniform shift $\rho \to \rho + c$ leaves the data unchanged), and the sensitivity–resolution trade-off imposed by the finite aperture. Only the second, the *absolute-level indeterminacy*, is the central concern of this thesis, since it cannot be removed by any purely image-based inversion; the other two are handled but remain secondary. The line-of-sight integration is neutralised by the two-dimensional setting adopted throughout: when the flow is uniform along the line of sight, the integration collapses to a multiplication by the known depth, leaving an indeterminacy of gradient rather than of integration. This is justified by the linear compressor cascade geometry, whose high-aspect-ratio blading sustains a flow approximately uniform along the span at midspan [7]. The sensitivity–resolution trade-off, by contrast, is set by the optical and background configuration rather than by the inversion; it is taken up only where Chapter 3 quantifies how the background pattern conditions the measurement signal.

Classical BOS pipelines confront these difficulties through a sequential chain: displacement extraction followed by Poisson integration of the gradient field under a boundary condition that fixes the integration constant. Each stage is sound in isolation, but the chain accumulates error: regularisation choices made during displacement extraction propagate unchecked into the density field, and the Dirichlet boundary condition required by the Poisson step must be imported from an independent source, typically a computational fluid dynamics (CFD) simulation. The advent of Physics-Informed Neural Networks (PINNs) [32], and more specifically of unified BOS (UBOS)/PINN formulations [26], offers a fundamentally different route: the optical measurement model and the governing equations of the flow are embedded in a single optimisation problem, so that data fidelity and physical consistency are enforced together and the fragmented post-processing chain is replaced by a coherent, end-to-end inference.

1.2. Problem Statement

The concrete point of departure for this work is the BOS campaign carried out at VKI on a supersonic linear compressor cascade representative of the Mid-Fan section of the CFM RISE engine, operating at supersonic inlet relative Mach number [7]. That study treated BOS as a complete end-to-end measurement chain, running from aerodynamic CFD simulation and optical design through image processing and density integration to a first exploratory PINN extension, providing the natural baseline for the present work. It established the feasibility of BOS for this configuration, delivered an optical-design methodology grounded in a synthetic-to-experimental workflow, demonstrated the advantage of optical-flow over cross-correlation in resolving fine shock structures, and produced a first proof of principle that BOS data and the compressible Euler equations can be fused inside a PINN. It also left three problems unresolved, which together delimit the scope of this thesis:

1. *Data-side ambiguity from optical-flow regularisation.* The reconstruction depends on the optical-flow regularisation parameter α_{OF}, so that, absent a ground-truth reference, the quality of the input displacement field is structurally ambiguous and the resulting bias propagates into every downstream stage.

2. *Absolute density offset.* The Poisson integration of classical pipelines introduces a systematic absolute-level indeterminacy. This is not a pipeline-specific defect but an intrinsic property of any gradient-based measurement, captured exactly by the null-space of the unified BOS operator; removing it requires identifying which minimal set of physical observables suffices to pin the absolute density level.

3. *Single-scalar PINN loss weighting.* The first VKI PINN attempt balanced the BOS data term and the four Euler residuals through a single scalar weight, a formulation too rigid for stable multi-field reconstruction and one that motivates a component-wise, possibly adaptive, weighting strategy.

Of these open problems, the present work directly confronts the second and the third, targeting the absolute density offset while addressing the rigidity of the loss balancing through an *adaptive weighting heuristic* (AWH). The AWH adopted here is *ground-truth-aware*: it tunes the component weights using the known closed-form solution, and therefore probes sufficiency *in principle* rather than operational sufficiency under a measurable-only weighting. Problem (1) is set aside by construction rather than ignored: because the investigation uses synthetic data whose measurement signal $\mathbf{b}$ is prescribed directly (Chapter 3), the optical-flow regularisation parameter α_{OF} never enters the reconstruction, so the data-side ambiguity it would otherwise introduce does not arise. The UBOS/PINN formulation of Molnar et al. [26] addresses the second and third open problems within this single framework, introducing in particular a dedicated inlet loss $\mathcal{L}_{\mathrm{in}}$ that pins the absolute density level but only insofar as a full Dirichlet condition on the four-component state $(\rho_{\mathrm{in}}, u_{\mathrm{in}}, v_{\mathrm{in}}, E_{\mathrm{in}})$ is available on the inflow boundary. Their supersonic wind-tunnel experiment supplies this complete inflow state through well-characterised inlet conditions. In the VKI compressor cascade, by contrast, a realistic campaign accesses far more restrictive inflow information: the total temperature T_0 of the entering flow and a small number of static pressure taps at discrete wall locations. This contrast defines the central research question of the thesis:

> *Which minimal subset of supplemental information, compatible with the measurements actually available in the VKI compressor cascade, is sufficient to lift the null-space indeterminacy of the unified BOS operator and recover the absolute density level, when the full inlet boundary condition of the original UBOS/PINN framework is not accessible?*

The thesis answers it by developing and validating three reconstruction models of decreasing reliance on supplemental information, each formalising a distinct hypothesis on how the absolute level can be anchored or isolated. Synthetic data are used throughout because only a closed-form test case permits attributing a reconstruction error to a specific deficiency of supplemental information rather than to experimental noise, while still exercising a non-trivial density-gradient field. The chosen case is smooth, deliberately excluding the shock structures of the transonic and supersonic target regime, whose treatment is deferred to Chapter 8.

The governing physics is the compressible Euler system: exact for the smooth synthetic vortex and a deliberate model choice for the shock-dominated, inviscid-driven target machine. Its limitations for the real cascade (boundary-layer behaviour, shock/boundary-layer interaction) and the transferability of the models to the VKI dataset are deferred to the concluding chapter.

1.3. Thesis Outline

The remaining chapters follow the causal chain laid out above, from foundations to a controlled experiment and its interpretation.

Chapter 2 provides the theoretical and methodological foundation: the physical principles of Schlieren imaging and BOS, the PINN framework, the previous VKI work as experimental baseline, and the unified BOS/PINN formulation.

Chapter 3 introduces a compressible Rankine vortex test case whose full primitive state is known in closed form, synthesises the reference and deformed BOS image pair through a controlled pipeline, defines the training signal $\mathbf{b}$ driving the measurement loss, maps the joint feasibility region of the discrete UBOS operator through a sensitivity campaign and verifies the operator against the analytical ground truth at that baseline before reconstruction.

Chapter 4 introduces the common UBOS/PINN framework shared by the three reconstruction models, together with the spectrum of supplemental information along which the three models are placed, and benchmarks the BOS-weighted importance sampler against a uniform-sampling baseline, establishing it as a quantitatively motivated shared component of the training strategy.

Chapters 5–7 present those three models as a single, intentionally staged experiment: a *controlled erosion* of the supplemental information available to the reconstruction, in which each chapter answers a question left open by the one before.

- **Model 1**: strongly anchored baseline with ground-truth Dirichlet conditions on all domain boundaries.
- **Model 2**: realistic thermodynamic configuration anchored by a single inlet total temperature and a small number of static-pressure taps.
- **Model 3**: reduced UBOS formulation driven by a gauge-invariant data loss.

Model 1 thus deliberately exceeds the supervision of the original UBOS/PINN framework, which anchors only the inflow boundary: its four-boundary Dirichlet data provide an over-supervised upper bound against which the erosion is measured.

Chapter 8 draws the three experiments together, summarises the main findings, discusses the methodological implications and limitations of the proposed framework, and outlines directions for future work, with particular emphasis on the transition from the synthetic validation environment to the experimental VKI dataset.

2 | Theoretical Background and Literature Review

This chapter provides the theoretical and methodological background: the physical principles of Schlieren imaging and BOS, the minimal background on neural networks for inverse problems and on PINNs, the previous VKI work as experimental baseline, and the literature situating its residual limitations. It converges on the unified BOS/PINN (UBOS/PINN) formulation of Molnar et al., the immediate precursor of the synthetic validation of Chapter 3. The three open problems of the VKI baseline recalled from Chapter 1, optical-flow regularisation ambiguity, absolute-density offset of gradient-based integration, and single-scalar weighting rigidity, act as the connecting thread of the chapter.

2.1. Schlieren Imaging and Background-Oriented Schlieren

As anticipated in Chapter 1, density completes the thermodynamic description of a compressible flow [13] and is inferable non-intrusively because the refractive index of a gas depends linearly on it (Eq. (2.1)). The techniques exploiting this principle (shadowgraphy, classical Schlieren, interferometry, BOS) are sensitive to the same phenomenon but differ in which spatial derivative of the refractive index they encode and in the optical complexity required to capture it [31, 37].

Refractive-Index Variations and Light Propagation in Gases

For gaseous media, the relationship between the refractive index n and the density ρ is well approximated by the Gladstone–Dale relation

$$n - 1 = G_{\mathrm{GD}}\, \rho, \tag{2.1}$$

where G_{GD} is the Gladstone–Dale constant; for air at visible wavelengths, $G_{\mathrm{GD}} \approx 2.26 \times 10^{-4}$ m^3/kg [31, 37]. Since the relation is linear, any spatial variation of ρ produces a proportional variation of n; moreover $G_{\mathrm{GD}}\rho \ll 1$, so n remains close to unity, which underpins the paraxial treatment below. In the geometric-optics limit, light propagation in such a medium is governed by the eikonal equation, which reduces, away from caustics, to the ray equation [37]

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\right) = \nabla n, \tag{2.2}$$

where $\mathbf{x}(s)$ denotes the spatial position along a ray parameterised by arclength s. Under the canonical paraxial conditions relevant to Schlieren and BOS, weak refractive-index variations and a test region much shorter than the distance to the detector [42], the ray remains close to its undisturbed straight path. For a line of sight aligned with the z-axis and a deflection in the transverse direction $\alpha \in \{x, y\}$, integration of Eq. (2.2), with n_0 the ambient refractive index ($n_0 \approx 1$ for air), gives

$$\varepsilon_\alpha \;=\; \frac{1}{n_0}\int \frac{\partial n}{\partial \alpha}\,\mathrm{d}z, \tag{2.3}$$

where ε_α is the angular deflection of the emerging ray [13, 37]. Substituting Eq. (2.1) into Eq. (2.3) expresses the deflection directly as an integrated density gradient,

$$\varepsilon_\alpha \;=\; \frac{G_{\mathrm{GD}}}{n_0}\int \frac{\partial \rho}{\partial \alpha}\,\mathrm{d}z, \qquad \alpha \in \{x, y\}. \tag{2.4}$$

From Classical Schlieren to BOS

The classical Schlieren technique introduced by Toepler converts the angular deflection of Eq. (2.4) into an image-plane intensity by means of a knife-edge or graded filter placed at the focal plane of a collimated optical chain (lens-type or z-type, Fig. 2.1(a)–(b)) [15, 37]. Although capable of high accuracy in laboratory conditions, the technique imposes severe practical constraints: the field of view is limited by the diameter of the matched collimating optics, the alignment of source, lenses (or parabolic mirrors), and cutoff requires a dedicated optical bench, and the technique remains intrinsically a parallel-light diagnostic. A structural sensitivity–range trade-off, controlled by the cutoff geometry, further constrains its operation [15]. These limitations provide the practical motivation for the BOS configuration of Fig. 2.1(d), in which the entire optical chain is replaced by a single camera and a structured background [31].

Background-Oriented Schlieren

BOS retains the physical principle of Eq. (2.4) but eliminates the precision optics that characterise the classical setup. The canonical configuration is shown in Fig. 2.2: a single camera, focused on a structured background pattern (typically a random-dot or wavelet-noise texture), images the background through the test region, so that the only optical elements between the flow and the sensor are the camera lens itself and the patterned plane [31]. Two images are acquired: a reference (wind-off) image, with the flow at rest, and a wind-on image, under flow conditions. Density gradients in the test region deflect each ray by an angle ε_α given by Eq. (2.4), so that the apparent position of every background point is shifted by a small displacement (δ_x, δ_y) in the image plane [31, 42]. The displacement field is recovered by digital comparison of the two images through a pattern-matching algorithm (cross-correlation inherited from particle image velocimetry (PIV), or optical-flow (OF) methods that exploit brightness constancy) [31, 36]; the trade-offs between the two are reviewed in the VKI context in Section 2.4. The output of either algorithm is a discrete two-component displacement field (δ_x, δ_y) that estimates, in pixel units, the apparent shift of the background pattern caused by the flow.

The connection between the angular deflection and the image-plane displacement follows from elementary geometric optics. In the canonical paraxial configuration, the displacement in physical units is

$$\Delta_\alpha \;=\; Z_D\, M\, \varepsilon_\alpha, \qquad M = \frac{f_{\text{lens}}}{Z_A + Z_D - f_{\text{lens}}}, \tag{2.5}$$

where Z_D is the distance from the phase object to the background, Z_A the distance from the phase object to the lens, f_{lens} the focal length, and M the optical magnification [31, 36]. Conversion to pixel units introduces the pixel pitch ℓ_{px}, so that $\delta_\alpha = \Delta_\alpha/\ell_{\text{px}}$. Combining Eqs. (2.4) and (2.5) and assuming that the density gradients are confined to a region of finite extent ΔZ_{grad} along the line of sight, the BOS forward model takes the compact form

$$\delta_\alpha \;=\; C_{\text{sys}} \int_0^{\Delta Z_{\text{grad}}} \frac{\partial \rho}{\partial \alpha}\, \mathrm{d}s, \qquad C_{\text{sys}} \;=\; \frac{G_{\text{GD}}\, Z_D\, M}{n_0\, \ell_{\text{px}}}, \tag{2.6}$$

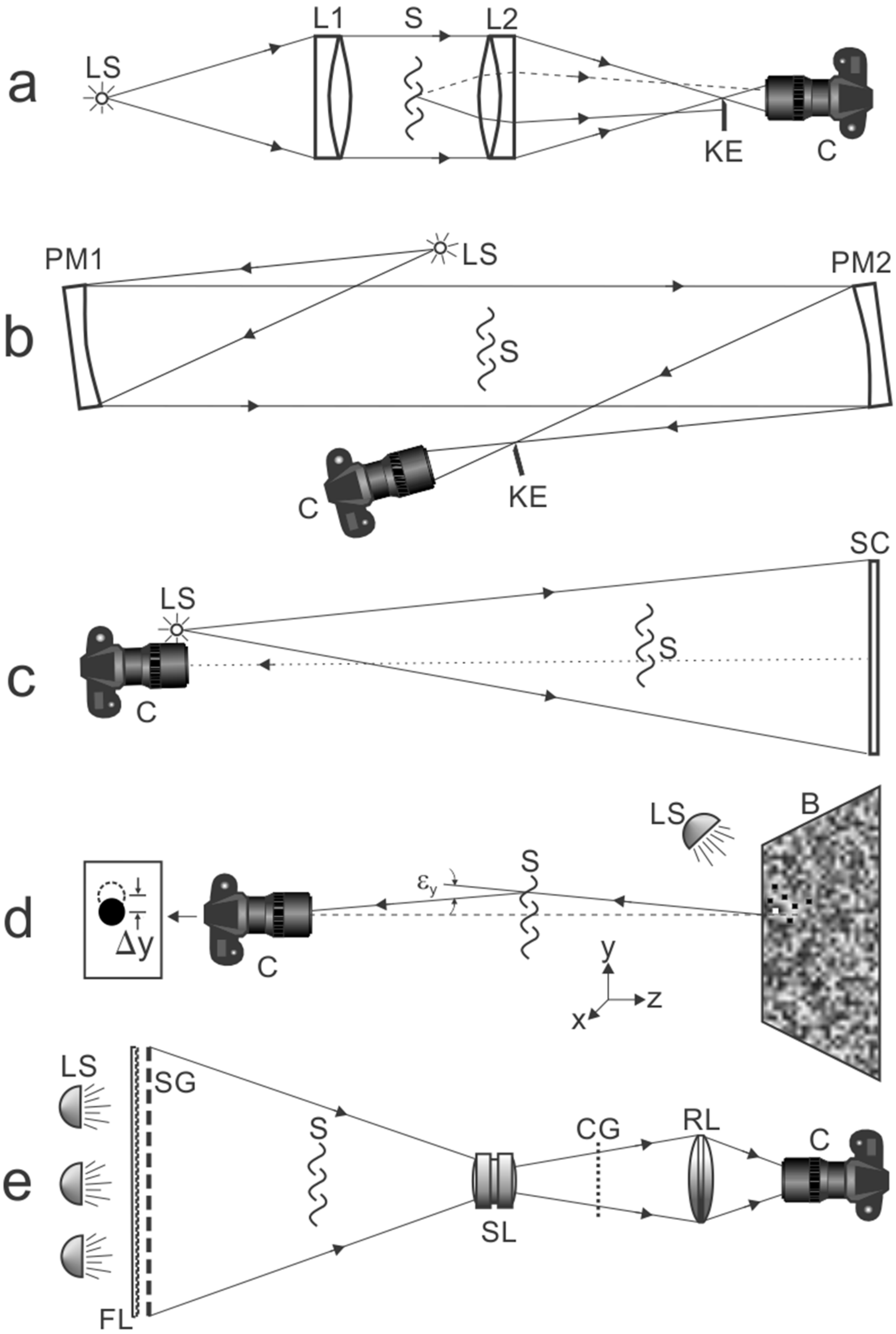


Figure 2.1: Schematic of five families of schlieren/shadowgraph instruments: (a) Toepler's lens-type schlieren system, with light source LS, collimating lens L_1, schlieren object S, focusing lens L_2, knife-edge KE at the focal plane, and camera C; (b) z-type system in which L_1 and L_2 are replaced by parabolic mirrors PM_1 and PM_2; (c) direct shadowgraphy; (d) BOS, discussed in Section Background-Oriented Schlieren; (e) focusing (lens-and-grid) schlieren. Reproduced from [37].

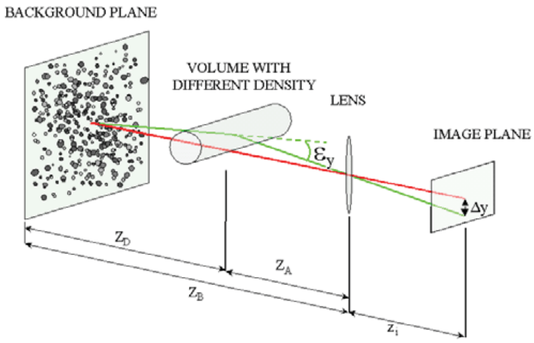


Figure 2.2: Canonical BOS setup: a camera focused on a background plane with a phase object (P) placed between them. The distances Z_D (object-background), Z_A (object-lens), and focal length f_{lens} define the magnification M (Eq. (2.5)) and system constant C_{sys} (Eq. (2.6)). Reproduced from [31].

where the integration runs along the (straight) reference ray, with $\mathrm{d}s \approx \mathrm{d}z$ and a non-zero contribution only within the gradient-bearing thickness ΔZ_{grad}, and C_{sys} is the system constant lumping together the Gladstone–Dale constant and the geometric parameters of the setup [26, 27]. Eq. (2.6) thus states that the sensor displacement is a linear functional of the density-gradient field along each line of sight, modulated by C_{sys}.

Comparative Assessment

Replacing the precision Schlieren optics by a structured background and a single camera carries three advantages: commodity hardware; a field of view limited by the background size rather than by any optical element; and intrinsic quantitativeness, each wind-on pixel yielding a local (δ_x, δ_y) mapped to the integrated density gradient through Eq. (2.6) without a calibration object [15, 31, 37].

These advantages come with three structural limitations intrinsic to the gradient-based, line-of-sight formulation of Eq. (2.6). (i) The *projection nature* of a single view: a fully three-dimensional field can be recovered only through tomography or through two-dimensional/axisymmetric closure assumptions [2, 31]. (ii) The *absolute-level indeterminacy*: any shift $\rho \to \rho + c$ leaves the measurement unchanged, so recovery of ρ requires an independent condition fixing the integration constant [42]. (iii) The *sensitivity–resolution trade-off*, by which the finite aperture couples C_{sys} to depth of field and spatial resolution [27, 36]. Limitations (ii) and (iii) act on the inversion itself, leaving the discrete operator non-injective (ii) and ill-conditioned (iii); BOS density reconstruction is thus classically ill-posed in the sense of Section 2.2, while the projection nature (i) is a separate recoverability question, neutralised here by the two-dimensional closure.

2.2. Neural Networks for Inverse Problems

Inverse problems arise whenever the quantity of interest cannot be observed directly and must instead be inferred from indirect measurements. In compact form, the forward model may be written as

$$\mathbf{y} = \mathcal{H}(\mathbf{u}) + \boldsymbol{\nu}, \tag{2.7}$$

where $\mathbf{u}$ denotes the unknown state, $\mathbf{y}$ the measured data, $\mathcal{H}$ the observation operator, and $\boldsymbol{\nu}$ the measurement noise [38]. In many engineering applications the inverse task is *ill-posed* in the sense of Hadamard: existence or uniqueness may fail, and, decisively, the solution may not depend continuously on the data, so small perturbations $\delta\mathbf{y}$ produce arbitrarily large variations in the reconstructed state, turning naive inversion into a noise amplifier [1]. The standard analytical response to ill-posedness is to embed prior knowledge directly into the cost functional,

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \left[\|\mathcal{H}(\mathbf{u}) - \mathbf{y}\|_2^2 + \lambda\,\mathcal{R}(\mathbf{u}) \right], \tag{2.8}$$

where $\mathcal{R}(\mathbf{u})$ encodes prior structure on the solution and λ controls the balance between data fidelity and regularisation [39]. The choice of $\mathcal{R}$ embodies whatever structural assumption is made about the unknown, total variation for piecewise-constant solutions, ℓ^1 penalties for sparsity, Sobolev norms for smoothness.

Neural Networks as Parametric Inverse Operators

Rather than solving Eq. (2.8) independently for each realisation, one may seek a parametric approximation $\hat{\mathbf{u}}_\theta = f_\theta(\mathbf{y})$ of the inverse map itself [9]. The function f_θ is realised by a feedforward neural network, i.e. a composition of affine transformations and pointwise nonlinear activations,

$$\mathbf{h}^{(\ell+1)} = \sigma^{(\ell)}\big(W^{(\ell)}\mathbf{h}^{(\ell)} + \boldsymbol{\beta}^{(\ell)}\big), \qquad \ell = 0, \ldots, L-1, \tag{2.9}$$

where $W^{(\ell)}$ and $\boldsymbol{\beta}^{(\ell)}$ are the weight matrices and bias vectors of layer ℓ and $\sigma^{(\ell)}$ the pointwise activation. The network f_θ is trained by gradient-based minimisation of a data-fidelity loss; modern implementations rely on backpropagation [34] to compute gradients through the chain rule and on stochastic optimisers such as Adam. Universal approximation results [9] guarantee that a sufficiently expressive network can represent any *continuous* target function on a compact set.

This is a necessary but not a sufficient condition for using neural architectures as surrogate inverse operators: the inverse map itself may be discontinuous when the problem is ill-posed in the sense of Hadamard introduced above, and universal approximation says nothing about optimisability, that is, whether such a network can actually be trained to the target by gradient-based minimisation.

A Taxonomy of Learning-Based Reconstruction

Following Ongie et al. [28], learning-based methods for inverse problems can be classified along two axes: the role of the forward operator $\mathcal{H}$ (fully known, partially known, unknown) and the supervision regime (paired, unpaired, or measurement-only data).

The four resulting families (post-processing of a classical reconstruction [18]; end-to-end supervised learning; model-unrolled iterative schemes [14]; generative priors [4]), differ primarily in *where* the regularising prior is injected, but share a common feature: none of them removes the ill-posedness of the underlying inverse problem. They merely *redistribute* it across architectural choices, training data, and optimisation procedures, replacing the explicit penalty $\lambda\mathcal{R}(\mathbf{u})$ of Eq. (2.8) with implicit inductive biases [1, 28].

For the BOS reconstruction problem of this thesis, the forward operator is known analytically but matched *experimental* ground-truth/measurement pairs are unavailable: supervised approaches are not viable for the single-instance reconstruction pursued here, although synthetic or CFD-generated pairs can re-enable them at the level of operator learning, a route taken up in Chapter 8. In the per-instance setting, the regularising prior must instead be carried by the governing equations themselves, the regime addressed next.

2.3. Physics-Informed Neural Networks

The discussion of Section 2.2 established that, across all four taxonomic strategies, the regularising information is ultimately carried by data, either labelled pairs or the statistical structure of a training corpus. In many problems of practical interest, however, a different and far more structured source of prior information is available: the governing equations of the physical system. The modern PINN formulation, building on early collocation-based neural partial differential equation (PDE) solvers [10], was consolidated by Raissi, Perdikaris, and Karniadakis [32] and integrates this source directly into the learning process.

The PINN Formulation

Consider a generic PDE problem

$$\mathcal{F}(u(\mathbf{z}); \boldsymbol{\xi}) = 0, \qquad \mathbf{z} \in \Omega, \tag{2.10}$$

supplemented by boundary and/or initial conditions

$$\mathcal{B}(u(\mathbf{z})) = g(\mathbf{z}), \qquad \mathbf{z} \in \partial\Omega, \tag{2.11}$$

where u denotes the unknown field, $\mathbf{z} = (x_1, \ldots, x_d, t)$ collects the spatio-temporal coordinates, $\boldsymbol{\xi}$ collects the physical parameters of the model (which may themselves be unknown in the inverse setting), $\mathcal{F}$ is a possibly nonlinear differential operator, and $\mathcal{B}$ is the boundary/initial-condition operator. In a PINN, the unknown solution is approximated by a fully connected neural network $\hat{u}_\theta(\mathbf{z})$ with the layered structure of Eq. (2.9), taking the coordinates as input and returning one or more physical fields as output. The central idea is to define the PDE residual

$$r_\theta(\mathbf{z}) = \mathcal{F}(\hat{u}_\theta(\mathbf{z}); \boldsymbol{\xi}) \tag{2.12}$$

and to penalise its violation during training [32]. The partial derivatives in Eq. (2.12) are evaluated through *automatic differentiation* (AD) [3]: AD propagates exact derivative information through the computational graph at the cost of a single backward pass, allowing the governing equations to be imposed at arbitrarily distributed collocation points without constructing a mesh.

The training of a PINN is therefore formulated as the minimisation of a composite loss function,

$$\mathcal{L}(\theta, \boldsymbol{\xi}) = w_{\mathrm{PDE}}\, \mathcal{L}_{\mathrm{PDE}} + w_{\mathrm{data}}\, \mathcal{L}_{\mathrm{data}} + w_{\mathrm{IC}}\, \mathcal{L}_{\mathrm{IC}} + w_{\mathrm{BC}}\, \mathcal{L}_{\mathrm{BC}}, \tag{2.13}$$

whose four terms penalise the PDE residual at interior collocation points and the mismatch with measurements, initial and boundary conditions. In the standard collocation-based formulation these are mean-squared errors over training-point sets [8, 32]:

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| r_\theta(\mathbf{z}_f^i) \right|^2, \tag{2.14}$$

$$\mathcal{L}_{\text{data}} = \frac{1}{N_d} \sum_{i=1}^{N_d} \left| \hat{u}_\theta(\mathbf{z}_d^i) - u_d^i \right|^2, \tag{2.15}$$

$$\mathcal{L}_{\text{IC}} = \frac{1}{N_{ic}} \sum_{i=1}^{N_{ic}} \left| \hat{u}_\theta(\mathbf{z}_{ic}^i) - u_{ic}^i \right|^2, \tag{2.16}$$

$$\mathcal{L}_{\text{BC}} = \frac{1}{N_{bc}} \sum_{i=1}^{N_{bc}} \left| \mathcal{B}\big(\hat{u}_\theta(\mathbf{z}_{bc}^i)\big) - g(\mathbf{z}_{bc}^i) \right|^2. \tag{2.17}$$

In forward problems $\boldsymbol{\xi}$ is known and the goal is to recover u; in inverse problems, the unknown physical parameters $\boldsymbol{\xi}$ are learned jointly with the network weights by minimising the same composite loss. This joint estimation of fields and parameters is the distinctive feature of PINNs and a primary motivation for their adoption here [32].

Training Pathologies and Mitigation Strategies

The composite loss in Eq. (2.13) is the source of two recurring difficulties in PINN training. The first is a *gradient imbalance*: when the PDE residual and the boundary terms operate on fields of very different magnitudes, the back-propagated gradients become unbalanced and the optimiser effectively ignores one or more components, with reported accuracy losses of one to two orders of magnitude [43]. The second is the inherent *spectral bias* of fully connected networks [24, 45]: high-frequency components, including sharp gradients near shocks, are learned far more slowly than smooth ones, concentrating residuals in convection-dominated and shock-containing flows. Spectral bias is one of the reasons the synthetic testbed of Chapter 3 is chosen deliberately smooth, isolating the inverse and gauge problem from the additional burden of shock capture; the implications of this choice are detailed in Section Application to Fluid Mechanics and Relevance to BOS and at the end of the chapter.

Proposed remedies, including adaptive loss-weighting from gradient statistics [43] or the Neural Tangent Kernel [45], self-adaptive point-wise weights [24], variational/Galerkin formulations [20], domain decompositions [17] and Bayesian PINNs [46], redistribute the multi-scale structure of the loss rather than removing it.

Application to Fluid Mechanics and Relevance to BOS

Fluid mechanics is the most natural application domain for PINNs in the present context. For incompressible Navier–Stokes, Jin et al. showed that physics constraints propagate information across spatial dimensions even from sparse cross-sectional observations [19], a capability directly relevant to projection-based diagnostics such as BOS. The compressible extension is harder: hyperbolic equations, near-discontinuous shocks incompatible with the smooth function class of fully connected networks, and residuals spanning many orders of magnitude across the four Euler equations. Cai et al. [6] confirm on a compressible-flow case study that dynamic loss weighting is essential for stable high-Mach training, and Cuomo et al. [8] identify shock-containing regimes as an open problem of the field. This reinforces the choice of a smooth, shock-free synthetic validation, noted in Section Training Pathologies and Mitigation Strategies.

PINNs are thus a natural candidate for the BOS inverse problem, with the failure modes above setting the constraints for any practical implementation.

2.4. Experimental Context

The experimental and methodological baseline of this thesis is the BOS study carried out at VKI on a supersonic compressor linear cascade [7], introduced in Chapter 1. The present section retraces that study only as far as needed to expose the three residual limitations from which the UBOS/PINN formulation, and the synthetic validation of the following chapters, departs.

Facility, Aerodynamic Scenario and Reference Field

The campaign was conducted in the CT2-SS supersonic wind tunnel at VKI, whose test section hosts a supersonic linear compressor cascade representative of the Mid-Fan section of the CFM RISE engine: a highly loaded stage designed to operate at a low-supersonic inlet relative Mach number. A convergent–divergent nozzle upstream sets the inlet Mach number, while sidewall boundary-layer suction and motorised exit tailboards enforce two-dimensional periodicity. The aerodynamic field was characterised beforehand by a steady three-dimensional Reynolds-Averaged Navier–Stokes (RANS) simulation of the cascade at the design point, whose density field was span-averaged into a quasi-two-dimensional reference map, used both to assess the reconstruction error and to supply the Dirichlet condition of the density-integration step.

Optical Design and Image Processing

Before the campaign, a synthetic assessment coupled the CFD solution with a numerical ray-tracing model of the full BOS imaging chain, generating reference/wind-on image pairs from which the displacement field could be extracted and compared against the known CFD input. This assessment confirmed that BOS design is governed by a fundamental sensitivity–resolution trade-off, whose quantitative form is discussed in Section 2.5; the same synthetic-to-experimental logic is the direct precedent of the validation strategy adopted in this thesis.

Displacement extraction was performed with the Horn–Schunck OF algorithm [16], adopted as the primary method because it resolves the leading-edge shock peak more sharply than cross-correlation, which suffers from limited dynamic range where density gradients are highest [7]. The OF energy balances a brightness-constancy term against a smoothness penalty weighted by α_{OF}: low values preserve sharp gradients but amplify noise, high values suppress noise but smear physically meaningful structures. On synthetic data this parameter can be tuned against the known CFD ground truth, but in the experiment no such reference exists, so its selection is necessarily heuristic and the resulting bias propagates unchecked into the integration step. This data-side ambiguity is the first of the three open problems left by the VKI baseline.

Density Integration and First PINN Extension

Once the displacement field (δ_x, δ_y) is known, density is recovered by inverting the BOS measurement model. Under the quasi-two-dimensional reduction of the line-of-sight integral of Eq. (2.6), the displacement components are proportional to the integrated density gradient through the system constant C_{sys} of Eq. (2.6); taking the divergence of this relation yields the Poisson equation [42]:

$$\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} = \frac{1}{C_{\text{sys}}} \left(\frac{\partial \delta_x}{\partial x} + \frac{\partial \delta_y}{\partial y} \right), \tag{2.18}$$

discretised with a second-order finite-difference scheme on the pixel grid. Dirichlet conditions were imposed at the inlet and outlet boundaries using the CFD span-averaged density, Neumann conditions near the blades, and periodicity on the upper and lower boundaries; at least one Dirichlet condition is required to fix the integration constant of the Laplacian system, a manifestation of the absolute-level indeterminacy formalised in Section 2.6. With noiseless displacements and exact CFD-based boundary conditions, the mean relative density error reached a best-case value of about 4% in the synthetic case [7]; the experimental error is necessarily larger, owing to OF noise propagation and to the mismatch between the realised operating point and the CFD-based boundary condition. This systematic absolute-density offset is the second open problem, and the one the present work isolates and confronts directly.

The VKI study also included a first, exploratory PINN extension: a network of 10 hidden layers with 64 neurons and tanh activations, mapping (x, y) to (p, u, v, E), trained by combining a data term (the mismatch between the network density output, recovered from the equation of state Eq. (2.26), and the BOS-reconstructed field) with a physics term built from the residuals of the steady, two-dimensional, compressible Euler system formalised in Section 2.6, the two being balanced through a single scalar weight w_{phys}. No value of w_{phys} proved uniformly satisfactory: small values reproduced the BOS-derived density but yielded physically meaningless pressure and energy fields, large values reduced the Euler residuals at the cost of departing from the measurements and re-introducing the absolute-density indeterminacy of the data term, and intermediate values gave only partial agreement [7]. Two mechanisms explain this: a single scalar cannot balance the data term against the four Euler residuals in a genuinely multi-field problem, and the gradient-derived data term is insensitive to a uniform offset $\rho \to \rho + c$, an indeterminacy no retuning of w_{phys} can remove. This rigidity of single-scalar weighting is the third open problem. It directly motivates the AWH developed in Section 6.3.

These three open problems, one on the measurement side and two on the data-assimilation side, organise the targeted review of Section 2.5 along its two axes, before Section 2.6 brings them together in the UBOS/PINN formulation [26].

2.5. Targeted Review Addressing the Baseline Limitations

The two families of work reviewed here follow the two axes anticipated at the end of Section 2.4. The first family, *Image-Processing and Optical-Model Improvements*, addresses the measurement-side ambiguity exposed by the VKI baseline. The second family, *Data-Assimilation and Learning-Based Improvements*, addresses the absolute-density indeterminacy of gradient-based measurements and the difficulty of balancing heterogeneous loss components in multi-field physics-informed reconstruction.

Image-Processing and Optical-Model Improvements

The recent BOS review by Schmidt et al. [36] situates the field's progress from basic displacement sensing toward a richer framework encompassing advanced image processing, tomography, and data-assimilation strategies. A central message of that review, directly relevant to the VKI baseline, is that the canonical sensitivity factor and the circle-of-confusion of any BOS system grow together with the geometric parameters Z_D, Z_A, f_{lens} and the aperture: any gain in sensitivity is purchased at the cost of a proportional loss of resolution. This is the quantitative form of the trade-off that Cospito encountered when closing the lens aperture to suppress blur [7].

A second strand of work has reconsidered the optical forward model itself. The canonical thin-ray approximation of Eq. (2.6) implicitly assumes a pinhole-like imaging configuration, but real cameras operate with finite apertures and collect a cone of rays whose averaging introduces systematic blur when the depth-of-field cannot be neglected [27]. Aperture-corrected forward operators and hardware-level solutions based on directional-reference backgrounds (DRBOS) [22] both demonstrate that some of the reconstruction error traditionally attributed to inversion in fact originates at the imaging stage. The general principle that emerges is that the measurement chain should be treated as a whole before any inversion strategy is selected, a principle adopted in the unified formulation of Section 2.6.

Data-Assimilation and Learning-Based Improvements

A common message of the recent learning-based BOS literature is that the measurement model should be embedded *inside* the inversion rather than applied as a preprocessing step. Molnar and Grauer [25] formalise this by replacing a post-processing data loss (which inherits the bias of an intermediate algebraic reconstruction) with a measurement loss $\|A\hat{\boldsymbol{\rho}}_\theta - \mathbf{b}\|_2^2$ that compares the network output to the original projection data through the forward operator A. The reported gains over post-processing are dramatic ($> 95\%$ reduction in concentration- and velocity-field errors, in a passive-scalar jet tomography case distinct from the BOS-density problem treated here), and demonstrate that the point at which the measurement model is applied matters more than the choice of architecture. This insight is the direct precursor of the unified BOS/PINN framework adopted in this thesis (Section 2.6).

Three further developments, less central to the present work but relevant for future extensions, deserve mention. Bayesian PINNs [25, 46] address overfitting under noisy measurements by sampling a posterior over network parameters; energy-based data losses [30] replace the implicit Gaussian noise assumption of the squared-error term with a data-driven likelihood, relevant when experimental BOS images exhibit non-Gaussian residuals; and nonlinear multifidelity Gaussian processes [29, 33] provide a principled way of fusing dense low-fidelity simulations with sparse high-fidelity measurements such as the static pressure taps available in the VKI cascade.

2.6. Physics-Informed BOS via UBOS/PINN

The UBOS/PINN framework of Molnar et al. addresses the second and third open problems of the VKI baseline (the absolute-density offset and the single-scalar weighting rigidity) by enforcing measurement consistency and the governing equations within a unified optimisation problem [7, 26]. Its mathematical structure also makes explicit why the absolute density level cannot be identified from gradient-based measurements alone, a property that conditions every modelling choice of the chapters that follow.

The Unified BOS Operator and Its Null Space

The starting point is the continuous BOS deflection model of Eq. (2.6), which under the paraxial approximation relates the image-plane deflection to the density gradient along the line of sight through the optical system constant C_{sys} [26]. Discretising the density on a basis $\Phi = \{\varphi_j\}_{j=1}^{N}$, so that $\rho(\mathbf{x}) \approx \sum_j \rho_j\, \varphi_j(\mathbf{x})$, turns this relation into

$$\boldsymbol{\delta}_\alpha = C_{\text{sys}}\, D_\alpha\, \boldsymbol{\rho}, \tag{2.19}$$

with $\boldsymbol{\delta}_\alpha \in \mathbb{R}^{N_p}$ evaluated at N_p pixel locations, D_α a sparse matrix integrating the discrete density gradient along each ray, and $\boldsymbol{\rho} \in \mathbb{R}^N$ the vector of density coefficients [26].

The unification step embeds the OF relation directly into the density reconstruction. For small displacements the image pair satisfies the linear OF equation $I_x\, \delta_x + I_y\, \delta_y = -I_t$, with I_x, I_y the spatial image gradients and I_t the difference between distorted and reference frames [26]. Substituting Eq. (2.19) and collecting the image gradients into diagonal matrices $X = \text{diag}(I_x)$, $Y = \text{diag}(I_y)$ yields the unified BOS operator

$$A\,\boldsymbol{\rho} = \mathbf{b}, \qquad A = C_{\text{sys}}(X D_x + Y D_y), \qquad \mathbf{b} = -\mathbf{I}_t, \tag{2.20}$$

which maps the density field directly to the image-difference data [26]. The decisive consequence is that regularisation no longer acts on an intermediate deflection field but on the density field itself, the physical quantity of interest.

By construction, A is sensitive only to the spatial gradient structure of ρ, not to its absolute level: since D_α integrates the density gradient, a constant shift $\rho \to \rho + c$ leaves $\partial\rho/\partial\alpha$, and hence $A\boldsymbol{\rho}$, unchanged. Formally, the constant vector $\boldsymbol{c} = c\,\mathbf{1}$ lies in the null space of A: since D_α discretises a spatial-derivative operator, it annihilates constant fields, $D_\alpha\,\mathbf{1} = \mathbf{0}$ (exactly, whenever the basis reproduces the constant through a partition of unity), so that

$$A(\boldsymbol{\rho} + c\,\mathbf{1}) = A\boldsymbol{\rho} + c\,A\mathbf{1} = A\boldsymbol{\rho} = \mathbf{b}. \tag{2.21}$$

The invariance is therefore *exact*, by linearity of A, and the data $\mathbf{b}$ do not determine the absolute density level. This is the same indeterminacy affecting the classical Poisson integration of Eq. (2.18), where a free constant is fixed by a boundary condition: an intrinsic property of gradient-based measurements, not a defect of UBOS, and the precise form of the absolute-level indeterminacy anticipated in Chapter 1. Any reconstruction strategy must introduce supplemental information to fix the gauge.

Physics-Informed Closure and Loss Function

The second pillar of the framework is a physics-informed representation of the flow variables: a PINN maps the spatial coordinates (x, r) or (x, y) to the unknown fields (density ρ, velocity components u and v, specific total energy E), and the governing equations are enforced through the loss. For high-speed, approximately inviscid flows the chosen physics is the compressible Euler system supplemented by an irrotationality condition [26]. In compact form, for a two-dimensional or axisymmetric setting with $\eta = 1$ (axisymmetric) or $\eta = 0$ (planar, with r reducing to the Cartesian coordinate y),

$$\frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial r} + \eta\frac{1}{r}\rho v = 0, \tag{2.22}$$

$$\frac{\partial(\rho u^2 + p)}{\partial x} + \frac{\partial(\rho uv)}{\partial r} + \eta\frac{1}{r}\rho uv = 0, \tag{2.23}$$

$$\frac{\partial(\rho uv)}{\partial x} + \frac{\partial(\rho v^2 + p)}{\partial r} + \eta\frac{1}{r}\rho v^2 = 0, \tag{2.24}$$

$$\frac{\partial[(\rho E + p)u]}{\partial x} + \frac{\partial[(\rho E + p)v]}{\partial r} + \eta\frac{1}{r}(\rho E + p)v = 0, \tag{2.25}$$

closed by the ideal-gas (calorically perfect) equation of state

$$p = (\gamma - 1)\rho\left[E - \frac{1}{2}\left(u^2 + v^2\right)\right], \tag{2.26}$$

and by the irrotationality condition $\omega = \partial v/\partial x - \partial u/\partial r = 0$. The synthetic methodology of the following chapter is restricted to the planar case ($\eta = 0$), a deliberate specialisation of a framework already formulated for both configurations [26].

A practical requirement concerns the scaling of Eqs. (2.22)–(2.25): training with the dimensional Euler equations is unstable because the residuals of the four equations differ by several orders of magnitude, whereas non-dimensionalising by the inlet state ($\rho_{\mathrm{ref}} = \rho_{\mathrm{in}}$, $u_{\mathrm{ref}} = |\mathbf{u}_{\mathrm{in}}|$, and L_{ref} the domain length) brings all residuals to a comparable $\mathcal{O}(1)$ scale [26]. This non-dimensional form is adopted throughout the present thesis.

The training objective combines three terms. The measurement term is built directly from the unified BOS model,

$$\mathcal{L}_{\mathrm{meas}} = \|A\boldsymbol{\rho} - \mathbf{b}\|_2^2\,, \tag{2.27}$$

where the PINN-predicted density is sampled on the support nodes of the basis to populate $\boldsymbol{\rho}$; the notation $\mathcal{L}_{\mathrm{meas}}$ is used deliberately to distinguish this UBOS-induced coupling term from the generic data losses $\mathcal{L}_{\mathrm{data}}$ introduced earlier in the chapter [26]. The physics term $\mathcal{L}_{\mathrm{phys}}$ integrates the squared residuals of the Euler and irrotationality equations over the domain, approximated in practice by Monte Carlo sampling of collocation points, while an inlet term $\mathcal{L}_{\mathrm{in}}$ enforces the known inflow state $(\rho_{\mathrm{in}}, u_{\mathrm{in}}, v_{\mathrm{in}}, E_{\mathrm{in}})$ as a Dirichlet condition. The total objective is the weighted sum

$$\mathcal{L}_{\mathrm{total}} = w_{\mathrm{meas}}\mathcal{L}_{\mathrm{meas}} + w_{\mathrm{phys}}\mathcal{L}_{\mathrm{phys}} + w_{\mathrm{in}}\mathcal{L}_{\mathrm{in}}, \tag{2.28}$$

which specialises the general form Eq. (2.13) with $w_{\mathrm{meas}} \leftrightarrow w_{\mathrm{data}}$, $w_{\mathrm{phys}} \leftrightarrow w_{\mathrm{PDE}}$ and $w_{\mathrm{in}} \leftrightarrow w_{\mathrm{BC}}$.

By Eq. (2.21), $\mathcal{L}_{\mathrm{meas}}$ is insensitive to constant density offsets, and $\mathcal{L}_{\mathrm{phys}}$ is likewise insensitive provided pressure and velocity adjust through the equation of state; $\mathcal{L}_{\mathrm{in}}$ is therefore the only term that pins the absolute density level. A related mechanism is the use of pressure taps: even when the Euler equations are jointly enforced, the PINN can produce locally inconsistent velocity and pressure fields where the optical signal is weak or geometrically singular, and Molnar et al. show that adding a single synthetic pressure tap at the relevant wall removes this inconsistency, with the benefit strongest precisely where the optical signal is weakest [26]. This motivates one of the central hypotheses of this thesis: that a small number of strategically placed pressure-tap measurements from the VKI wind tunnel may suffice to remove the density offset, provided they are included in the total loss rather than used as a post-processing correction.

3 | Synthetic Data Generation and Validation

The UBOS/PINN formulation of Molnar et al. [26] addresses the density-integration weakness of the VKI baseline [7], but leaves open which minimal subset of supplemental information suffices to lift the null-space indeterminacy of Eq. (2.21). Answering this requires first validating the numerical pipeline that implements the UBOS measurement model in a setting where every ingredient is known in closed form. The chapter develops this validation in five steps: the synthetic two-dimensional test case and its full ground-truth state; the reference/deformed image pair and the training signal **b**; the linearised optical-flow framework against which **b** is tested; a sensitivity campaign mapping the joint feasibility region of the discrete UBOS operator; and the numerical consistency check at the selected baseline.

3.1. Construction of the Synthetic Dataset

Applying the framework directly to the experimental VKI CT2 images would entangle the unknown density, optical imperfections, camera noise, background limitations and operator discretisation error. Generating the BOS image pair from an analytical field makes every quantity in Eq. (2.20) available in closed form, isolating three objectives: verifying the discrete UBOS operator, characterising the null-space of Eq. (2.21) on a known reference, and selectively enabling boundary channels to test whether T_0 and a few p_s substitute for the dense inlet supervision [26]. The synthetic flow chosen for the validation stage is a compressible variant of the classical Rankine vortex [35]. Three features match the requirements of a UBOS/PINN validation benchmark: planar two-dimensionality, compatible with the $\eta = 0$ form of the Euler system in Eqs. (2.22)–(2.25); a non-trivial density gradient field that remains smooth almost everywhere, bounding discretisation errors; and a semi-analytical construction that yields the complete primitive state (ρ, u, v, p), together with the derived total energy E, without a CFD solver. The chosen benchmark is deliberately benign with respect to the target application: steady, subsonic, isentropic, shock-free, with sub-pixel displacements. Its role is to isolate and verify the numerical pipeline; the resulting limits of representativeness, in particular the absence of the shocks and O(pixel) displacements of transonic cascade flows, are stated in the limitations paragraph of Section 3.4.

Velocity Profile and Kinematic Parameters

In polar coordinates (r, θ) centred on the vortex core at $(x_0, y_0) = (0, 0)$, the Rankine velocity profile consists of a solid-body rotation inside a core of radius r_c and an irrotational outer region:

$$v_\theta(r) = \begin{cases} \dfrac{\Gamma}{2\pi r_c^2}\, r, & r \leq r_c, \\ \dfrac{\Gamma}{2\pi r}, & r > r_c, \end{cases} \tag{3.1}$$

where Γ is the circulation associated with the outer potential vortex; introducing the peak tangential velocity $v_c = \Gamma/(2\pi r_c)$, attained at the core boundary $r = r_c$, the corresponding solid-body angular velocity inside the core is $\omega_0 = v_c/r_c$. The validation case adopts $r_c = 10$ mm and $v_c = 130$ m/s, which yields a peak core Mach number based on the far-field speed of sound, $M_c = v_c/a_\infty \approx 0.382$, for the ambient state specified below (a_∞ defined in the following subsection). Because the temperature, and hence the local speed of sound $a(r) = \sqrt{\gamma p(r)/\rho(r)}$, decreases inside the core, the peak *local* Mach number v_θ/a slightly exceeds M_c, reaching ≈ 0.39 at the core boundary; the flow stays subsonic everywhere, making the isentropic closure of the following subsection physically admissible. The velocity components in Cartesian coordinates follow from the kinematic decomposition

$$u(x, y) = -v_\theta(r)\,\sin\theta, \qquad v(x, y) = +v_\theta(r)\,\cos\theta, \tag{3.2}$$

with $\cos\theta = (x - x_0)/r$ and $\sin\theta = (y - y_0)/r$ computed from the Cartesian grid.

The resulting Cartesian velocity field $\mathbf{u}(x, y) = (u, v)$ is shown in Fig. 3.1.

Isentropic Closure and Radial Density Profile

To close the system, the flow is assumed isentropic. This is a modelling choice instrumental to the validation rather than an exact physical statement. Imposing a single isentropic relation with a constant K across both the rotational core and the irrotational outer region is an idealisation: a real vortex, made non-isentropic by viscosity or by shocks, would

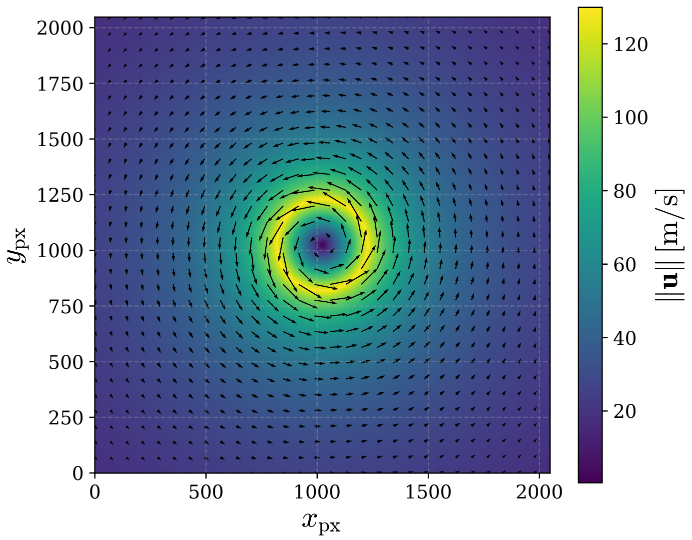


Figure 3.1: Cartesian velocity field $\mathbf{u}(x, y) = (u, v)$ for the Rankine vortex. Arrows indicate in-plane velocity vectors; background colour shows magnitude $\|\mathbf{u}\| = \sqrt{u^2 + v^2}$. The field displays the characteristic azimuthal pattern.

develop a radial entropy gradient. The inviscid field constructed here is nonetheless thermodynamically self-consistent. Because the flow is purely azimuthal and axisymmetric, uniform entropy is compatible with Crocco's relation $T\nabla s = \nabla h_0 - \mathbf{u} \times \boldsymbol{\omega}$, which for $\nabla s = 0$ reduces to $\mathrm{d}h_0/\mathrm{d}r = v_\theta\, \omega_z$ and is satisfied identically once the radial momentum balance of Eq. (3.4) holds. What matters for the present purpose is not the exact physical realisability of the state but its smoothness and reproducibility: enforcing isentropy everywhere yields a smooth, semi-analytical thermodynamic state suitable as ground truth for the optical pipeline.

The isentropic relation

$$p = K\, \rho^{\gamma}, \qquad K = \frac{p_\infty}{\rho_\infty^{\gamma}}, \tag{3.3}$$

is combined with the radial momentum balance of a steady circular flow,

$$\frac{1}{\rho}\frac{\partial p}{\partial r} = \frac{v_\theta^2}{r}, \tag{3.4}$$

to eliminate pressure. The isentropic constant $K = p_\infty/\rho_\infty^{\gamma}$ is fixed from the prescribed ambient state, so the synthetic ground truth and the physics loss of Eq. (2.26) share, by construction, the same closure $p = K\rho^{\gamma}$ with the same K. This is an idealisation of the validation testbed, not a property recovered by the reconstruction: equation of state and pressure information jointly fix the absolute scale of ρ, so K is a critical gauge parameter, and adopting its ground-truth value removes the indeterminacy by design rather than by measurement. In an experimental setting K would instead be estimated from the measurement channels, combining the mean tap pressure with the total-temperature reference; the explicit estimator and its use in the reconstruction are given in Chapter 6. This estimation, and the breakdown of global isentropy in shock-bearing flows, are discussed among the limitations of Section 3.4. Using $\partial p/\partial r = K\gamma\rho^{\gamma-1}\,\partial\rho/\partial r$ and the definition of the far-field speed of sound $a_\infty^2 = \gamma p_\infty/\rho_\infty$, one obtains the first-order ordinary differential equation governing the radial density profile,

$$\frac{d\rho}{dr} = \frac{v_\theta^2(r)}{a_\infty^2\, r}\,\rho^{\,2-\gamma}\,\rho_\infty^{\gamma-1}. \tag{3.5}$$

Eq. (3.5) is integrated from $r_{\max} = 1.1\,R_{\max}$, with $R_{\max}$ the largest radial coordinate on the computational grid, inward to $r = 0$ using a fourth-order Runge–Kutta scheme with Dirichlet condition $\rho(r_{\max}) = \rho_\infty$ [5]. The ambient state is set to $\rho_\infty = 1.225$ kg/m^3, $p_\infty = 101{,}325$ Pa, $R_{\text{gas}} = 287.05$ J/(kg K), and $\gamma = 1.4$. The implied temperature $T_\infty = p_\infty/(\rho_\infty R_{\text{gas}}) \approx 288.16$ K is consistent with ISA sea-level conditions (within 0.01 % of 288.15 K), as expected from the choice of ρ_∞, p_∞ and R_{gas}. The Dirichlet value ρ_∞ is the $r \to \infty$ density imposed at the finite radius $r_{\max} \approx 0.0778$ m ($R_{\max} \approx 0.0707$ m the grid-corner radius); since $r_{\max}/r_c \approx 7.8$ and the outer profile decays rapidly, the boundary mismatch is negligible at the scale of the density gradients used below. The constructed field is by design an exact solution of the steady, axisymmetric Euler system used as physical residual: with $\partial/\partial\theta = 0$ the continuity equation is satisfied identically, the radial momentum balance of Eq. (3.4) is imposed, and the isentropic relation closes the energy equation, so that the only region where the Euler residual is non-zero is the thin annulus at $r = r_c$; this localisation is recovered in the residual analysis of the model chapters.

Construction of the Full Ground-Truth State

The density profile $\rho(r)$ obtained from the integration of Eq. (3.5), together with the velocity profile of Eq. (3.1), is mapped onto a Cartesian grid of $H \times W = 2048 \times 2048$ pixels spanning $[-0.05, +0.05]$ m in both directions. The full state follows by combining Eqs. (3.2)–(3.3): the pressure field is obtained from $p = K\rho^\gamma$, the internal energy per unit mass from $e = p/\big((\gamma-1)\rho\big)$, and the derived total energy per unit mass from $E = e + \frac{1}{2}(u^2+v^2)$, in agreement with the closure adopted for the physics loss in Eq. (2.26).

In addition to the primitive variables, a total-temperature map

$$T_0(x,y) = \frac{p(x,y)}{\rho(x,y)\,R_{\text{gas}}} + \frac{u^2(x,y)+v^2(x,y)}{2\,c_p}, \qquad c_p = \frac{\gamma R_{\text{gas}}}{\gamma-1}, \tag{3.6}$$

is computed on the whole domain. The field is saved in the training package as a ground-truth reference for the T_0 soft Dirichlet constraint used in Chapter 6. The information content of this field is, however, strongly inhomogeneous: by constancy of total enthalpy in the steady, isentropic outer region $T_0 \approx T_\infty$ is essentially uniform, departing appreciably from T_∞ only inside the rotational core $r \leq r_c$. The T_0 soft-Dirichlet constraint therefore carries genuine information almost exclusively in the core and is nearly redundant in the far field, a useful but spatially concentrated supplement to the optical signal that underlies its role in the controlled erosion of supervision studied in the Model 2–3 chapters.

Finally, the density gradients $\partial_x \rho$, $\partial_y \rho$ required by the UBOS operator of Eq. (2.19) are computed on the Cartesian grid by the fourth-order, five-point, first-derivative stencil

$$\left.\frac{\partial \rho}{\partial x}\right|_{i,j} \approx \frac{\rho_{i,j-2} - 8\rho_{i,j-1} + 8\rho_{i,j+1} - \rho_{i,j+2}}{12\,\Delta x}, \tag{3.7}$$

where the index pair (i, j) follows the image-processing convention with i denoting the row (i.e. the y direction) and j the column (i.e. the x direction), so that j increases with x; the analogous formula for $\partial_y \rho$ operates on the index i with spacing Δy. The choice of a five-point stencil rather than a lower-order one is methodologically motivated: the same scheme is subsequently applied to the image gradients I_x, I_y derived from I_{ref} and entering the UBOS consistency check of Section 3.4. This shared differentiation rule lets the stencil truncation errors cancel insofar as ρ and I_{ref} share grid and sampling, and minimises the risk of interpreting discretisation artefacts as modelling errors [11]; this is relied upon in the consistency analysis of Section 3.2.

Fig. 3.2 summarises the ground-truth fields obtained from this construction: density, pressure, density-gradient components, total temperature, and total energy per unit mass.

Image Synthesis Pipeline

The ground-truth state provides the physical inputs but not yet the optical observables. The synthetic BOS image pair $(I_{\text{ref}}, I_{\text{def}})$ is constructed in three stages: assembly of the reference background I_{ref}; conversion of the density gradients into pixel-space deflections through C_{sys}; and a backward warp producing I_{def} from I_{ref} and the deflection field.

The Reference Background Image

The reference image I_{ref} is the background pattern recorded in the undisturbed configuration. It plays a dual role: it is displaced by the flow to produce I_{def}, and it carries the spatial gradients I_x, I_y populating the UBOS operator A of Eq. (2.20). For this role

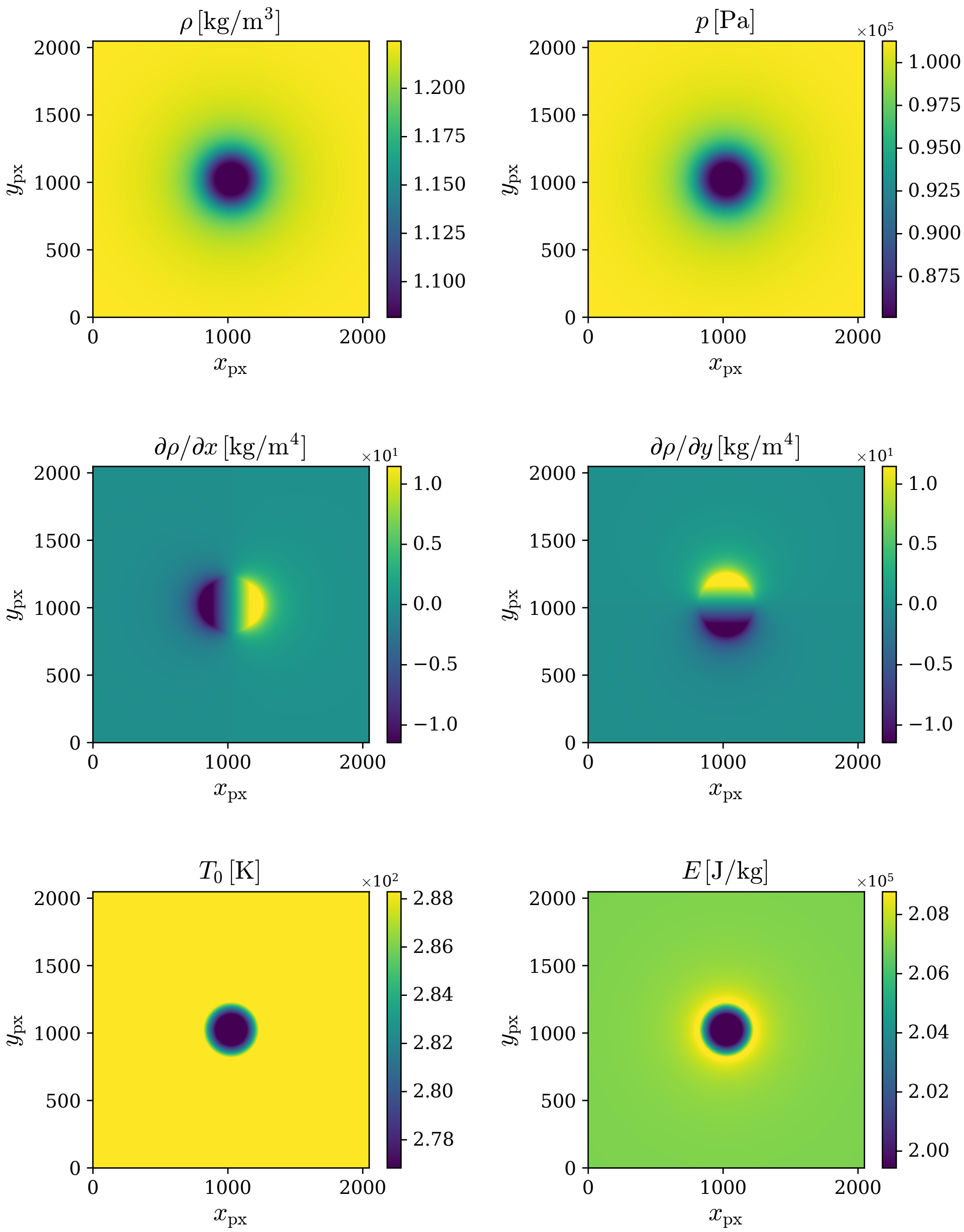


Figure 3.2: Ground-truth fields generated from the compressible Rankine vortex model. All fields are derived from the analytical velocity profile of Eq. (3.1) and from the isentropic thermodynamic reconstruction based on Eq. (3.5).

to be well-conditioned, ∇I_{ref} must be dense and isotropic, so that no extended row of A is silenced, and its dominant spatial scale must exceed the expected displacement $|\boldsymbol{\delta}|$ (a condition formalised in Section 3.2). I_{ref} is generated through the following four-stage procedure; the resulting 2048×2048, 16-bit image is the fixed background template of all validation experiments.

Stage 1 — Blue-noise dot placement. Dot centres are distributed across the image canvas by a grid-accelerated rejection-sampling algorithm that enforces a strict minimum inter-dot separation d_{min}. The canvas is partitioned into cells of side d_{min}; each candidate (x, y) drawn uniformly is accepted only if its Euclidean distance from every previously accepted dot in the surrounding 3×3 block exceeds d_{min}. The algorithm terminates when either N_{dot} dots have been placed or a saturation criterion (2000 consecutive rejections) is reached. The resulting spatial distribution is a *blue-noise* pattern in the sense of Ulichney [41]: low-spatial-frequency power is depleted up to a characteristic exclusion scale set by d_{min}, while high-frequency power remains approximately flat, ensuring the isotropy of ∇I_{ref} required by the data-loss formulation. The baseline parameters $N_{\text{dot}} = 6500$, $d_{\text{min}} = 19$ px yield a realised area of ~ 645 px^2 per dot across the 2048×2048 canvas, well above the close-packing area $(\sqrt{3}/2)\, d_{\text{min}}^2 \approx 313$ px^2 at separation d_{min}; the realised dot density is thus roughly half the close-packing limit, so that the canvas is intentionally unsaturated (Fig. 3.3(a)).

Stage 2 — Dot rendering. Each accepted centre is rendered as a filled disc of radius $r_{\text{dot}} = 2$ px in black on a white background. The finite support $\pi r_{\text{dot}}^2 \approx 13$ px^2 is sufficient to resolve a meaningful finite-difference gradient under the five-point stencil of Eq. (3.7), whereas a single-pixel mark would produce impulsive gradients incompatible with the linearised optical-flow model.

Stage 3 — Gaussian blur. A Gaussian blur with $\sigma_{\text{blur}} = 2.1$ px is convolved with the binary dot field, serving a dual purpose: algorithmically, it softens the step-like disc edges and suppresses the sub-pixel aliasing otherwise inconsistent with the small-displacement linearisation; physically, it approximates the finite point-spread function of a real imaging lens. This value keeps the effective dot radius $r_{\text{dot}} + 2\sigma_{\text{blur}} \approx 6.2$ px well below the half-spacing $d_{\text{min}}/2 = 9.5$ px, preserving inter-dot contrast.

Stage 4 — Low-frequency intensity modulation. A smooth additive modulation, generated by low-pass-filtering a white-noise array with $\sigma_{\text{noise}} = 50$ px and rescaling to $[-\eta_{\text{LF}}/2, +\eta_{\text{LF}}/2]$ with $\eta_{\text{LF}} = 5$ grey levels, is superimposed on the blurred field to mimic the illumination and reflectivity drifts of experimental BOS backgrounds. Because the modulation is inherited by I_{def} through the backward warp of Eq. (3.10), it introduces no spurious contribution to $\mathbf{b}$ beyond the displacement-induced one.

The four generation stages and the resulting spatial gradient structure are summarised in Fig. 3.3.

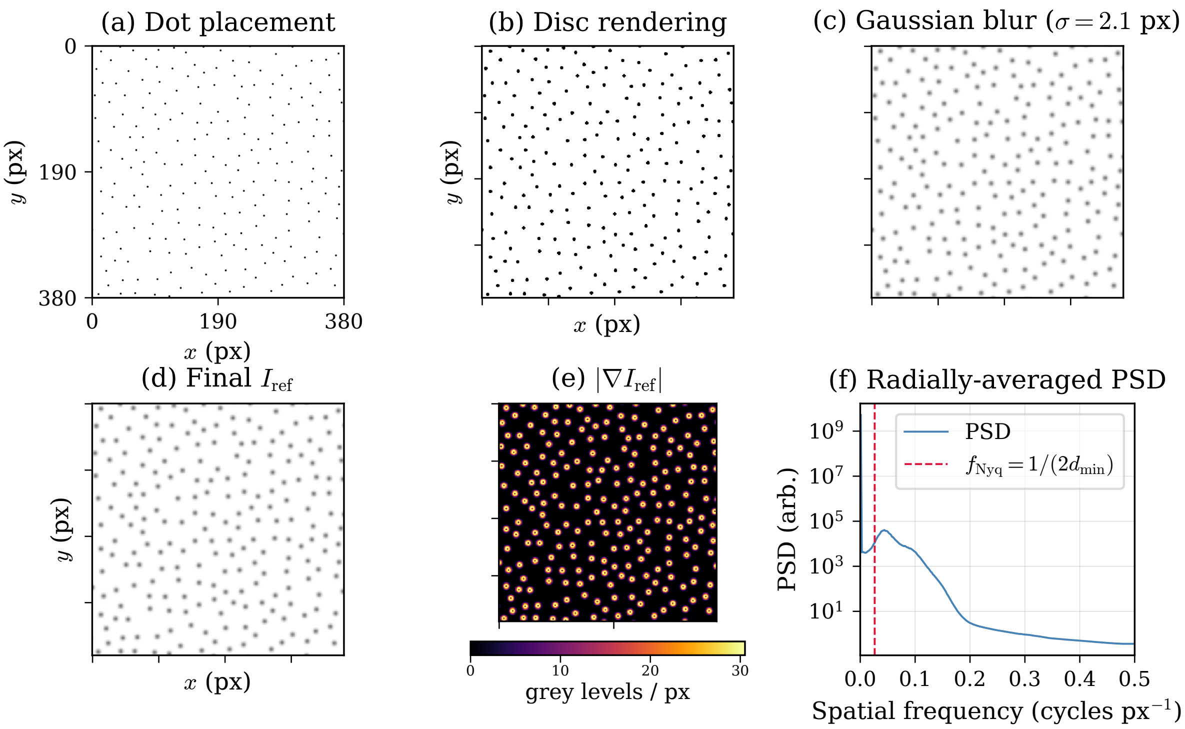


Figure 3.3: Reference background image generation pipeline (380×380 crop on 2048×2048 canvas). (a) Dot centres via blue-noise rejection-sampling ($N_{\mathrm{dot}} = 6500$, $d_{\mathrm{min}} = 19$ px). (b) Binary image: filled discs of radius $r_{\mathrm{dot}} = 2$ px. (c) Gaussian blur ($\sigma_{\mathrm{blur}} = 2.1$ px) removes sharp edges and aliasing. (d) Final I_{ref} with low-frequency modulation ($\eta_{\mathrm{LF}} = 5.0$, $\sigma_{\mathrm{noise}} = 50$ px). (e) Gradient magnitude $|\nabla I_{\mathrm{ref}}|$ (five-point stencil, Eq. (3.7)); confirms dense and isotropic gradient field. (f) Radially-averaged power spectral density, with dashed line at the exclusion frequency (d_{min}), indicating blue-noise suppression at low frequencies.

The complete set of generation parameters adopted for the baseline validation is collected in Table 3.1. A fixed random seed (42) is used throughout to ensure exact reproducibility of the image and of all downstream quantities derived from it. How each of these parameters influences the conditioning of the UBOS operator and the quality of the $\mathbf{b}$–s relation is the subject of the sensitivity study introduced in Section 3.3.

Table 3.1: Parameters of the reference background image generator used in the baseline synthetic validation.

Stage	Parameter	Symbol	Value
Dot placement	Canvas size (px)	$W \times H$	2048×2048
	Number of dots	N_{dot}	6500
	Minimum inter-dot separation	d_{min}	19 px
	Random seed	–	42
Dot rendering	Dot radius	r_{dot}	2 px
Gaussian blur	Blur standard deviation	σ_{blur}	2.1 px
LF modulation	Noise strength (grey levels)	η_{LF}	5.0
	Noise spatial scale	σ_{noise}	50 px

Optical System Constant and Pixel-Space Deflections

In the planar, paraxial setting adopted here, the line-of-sight integral of Eq. (2.6) reduces to a product with the effective probing length ΔZ_{grad}, which represents the axial extent of the region over which the density gradient is assumed to be coherent along the optical path. The collapse of the integral to a single product presupposes out-of-plane coherence, that is, a notionally two-dimensional flow uniform along the optical axis; ΔZ_{grad} is therefore a modelling choice tied to the 2D approximation, set to a value representative of the VKI cascade span. In the paraxial regime the ray coincides with its reference straight line, so that the curvilinear integration variable s may be identified with the optical-axis coordinate z ($\mathrm{d}s \approx \mathrm{d}z$). As a consequence, the system constant absorbs an extra length ΔZ_{grad} with respect to its general form in Eq. (2.6): dimensionally, C_{sys} goes from m^3/kg in the integrated formulation to m^4/kg in the planar one. The continuous BOS deflection model of Eq. (2.6) then collapses to

$$\delta_\alpha(x, y) = C_{\text{sys}} \frac{\partial \rho}{\partial \alpha}(x, y), \qquad \alpha \in \{x, y\}, \tag{3.8}$$

with

$$C_{\text{sys}} = \frac{G_{\text{GD}}\, Z_D\, M\, \Delta Z_{\text{grad}}}{n_0\, \ell_{\text{px}}}, \tag{3.9}$$

where the symbols G_{GD}, Z_D, M, n_0 and ℓ_{px} are those of Eq. (2.6) (Chapter 2), and ΔZ_{grad} is the effective density-gradient probing length introduced by the planar reduction. The values adopted in the synthetic case are summarised in Table 3.2 and reproduce the order of magnitude of the VKI optical configuration discussed in Section 2.4. With these values $C_{\text{sys}} \approx 2.66 \times 10^{-2}\ \text{m}^4/\text{kg}$, which, when multiplied by the density gradient $\partial\rho/\partial\alpha$

(units kg/m^4), yields a dimensionless deflection which is interpreted directly in pixel units thanks to the inclusion of $1/\ell_{\text{px}}$ in Eq. (3.9). The corresponding displacement magnitude $\|\boldsymbol{\delta}(x, y)\|$ is shown in Fig. 3.4.

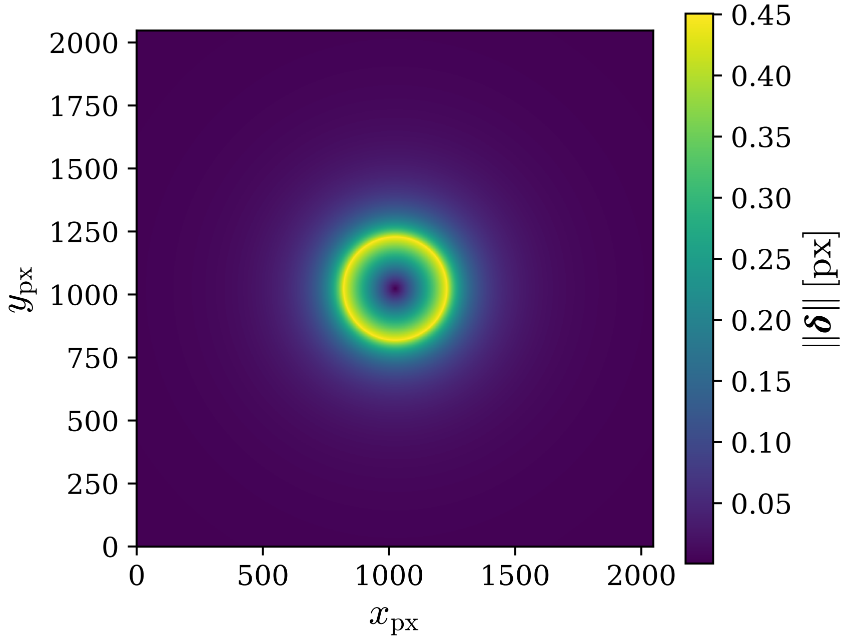


Figure 3.4: Magnitude of the synthetic BOS displacement field $\|\boldsymbol{\delta}(x, y)\|$ expressed in pixel units. The field is obtained from the pixel-space deflection components generated through Eq. (3.8).

Table 3.2: Optical parameters used to assemble the synthetic BOS system constant C_{sys} through Eq. (3.9).

Parameter	Symbol	Value
Gladstone–Dale constant	G_{GD}	2.26×10^{-4} m^3/kg
Background-to-schlieren distance	Z_D	0.116 m
Imaging magnification	M	0.066
Density-gradient probing length	ΔZ_{grad}	0.100 m
Physical pixel pitch	ℓ_{px}	6.5 μm

Backward Warping and Construction of I_{def}

The deflection field (δ_x, δ_y) obtained from Eq. (3.8) is used to synthesise the deformed image I_{def} from the reference image I_{ref}. Operationally, the construction uses a *backward warp*: each output pixel at location (X, Y) in I_{def} is assigned the value of I_{ref} at the source location $(X - \delta_x, Y - \delta_y)$, interpolated by a fifth-order spline. Formally,

$$I_{\text{def}}(X, Y) = I_{\text{ref}}\big(X - \delta_x(X, Y),\ Y - \delta_y(X, Y)\big). \tag{3.10}$$

The backward-warp convention is preferred to a forward scatter because every output pixel receives a value without gaps or overlaps, and I_{ref} can be evaluated at non-integer coordinates by a high-order interpolator. The fifth-order spline and the fourth-order stencil used in Section 3.2 for I_x, I_y keep interpolation and differentiation errors small relative to the leading optical-flow linearisation error: for the baseline sub-pixel displacements these errors are of high order in $|\boldsymbol{\delta}|$ and set the residual floor seen at small $|\boldsymbol{\delta}|$ in the sensitivity campaign of Section 3.3, well below the linearisation residual at the baseline amplitude.

Eq. (3.10) also fixes the sign convention of the pipeline. A positive deflection $\delta_\alpha > 0$ moves the apparent position of a background feature in the direction of increasing α, so that I_{def} appears shifted with respect to I_{ref} in the same direction. The sign is propagated consistently into the training signal defined in Section The Image Difference Signal $\mathbf{b} = I_{\text{ref}} - I_{\text{def}}$ and is verified a posteriori by the consistency check of Section 3.4.

Evaluation mask and noise model. An evaluation mask $E_{\text{val}} \in \{0, 1\}^{H \times W}$ identifies the region over which the UBOS residual and the physics loss are evaluated during training. For the present validation case this mask is set to the full domain: the Rankine vortex occupies the centre of the grid and smoothly decays toward the far field, so no spatial exclusion is needed. The pipeline nevertheless supports circular and rectangular submasks to enable localised sensitivity studies without altering the underlying synthetic data. An optional additive Gaussian noise model, parameterised as a fraction of the 16-bit dynamic range, can be superimposed on I_{ref} and I_{def} after warping; in the baseline validation it is disabled, so that the consistency check of Section 3.4 isolates discretisation and interpolation errors from measurement noise in a go/no-go test, making any failure attributable to the discrete operator alone.

BOS Training Signal and Training Package

With the image pair $(I_{\text{ref}}, I_{\text{def}})$, two further ingredients complete the synthetic dataset feeding the PINN models of the next chapters: the data vector $\mathbf{b} = I_{\text{ref}} - I_{\text{def}}$, which serves as the measurement input of the UBOS loss of Eq. (2.27), and the supplemental boundary information that lifts the null-space indeterminacy of Eq. (2.21).

The Image Difference Signal $\mathbf{b} = I_{\text{ref}} - I_{\text{def}}$

The data vector $\mathbf{b}$ is assembled directly from the image pair,

$$\mathbf{b}(x, y) = I_{\text{ref}}(x, y) - I_{\text{def}}(x, y), \tag{3.11}$$

so that $\mathbf{b} = -\mathbf{I}_t$ in the standard optical-flow notation $\mathbf{I}_t \equiv I_{\text{def}} - I_{\text{ref}}$, consistent with the UBOS operator of Eq. (2.20).

The resulting image pair and training signal are summarised in Fig. 3.5. The full-field views of I_{ref} and I_{def} (panels (a) and (b)) are visually indistinguishable, which is the expected signature of a sub-pixel displacement field; their central crops (panels (d) and (e)) confirm that the backward warp of Eq. (3.10) preserves the dot structure of the blue-noise background without visible artefacts. The image difference $\mathbf{b}$ (panels (c) and (f)) isolates the flow-induced component of the signal into an annular pattern around the vortex core, whose spatial support matches that of the displacement field in Fig. 3.4 and, through Eq. (3.8), that of the density-gradient field.

Supplemental Boundary Information

Beyond the image-difference signal $\mathbf{b}$, the training package must supply the supplemental boundary information intended to lift the null-space indeterminacy of Eq. (2.21). Motivated by the VKI cascade instrumentation, the synthetic case mimics two complementary sources of boundary data.

Static-pressure taps. A set of $N_{p_s} = 10$ synthetic pressure taps is placed on the computational grid, a fraction $\kappa = 0.60$ sampled uniformly within the rigid core $r \leq r_c$ and the remainder in the potential-vortex region. Two distinct, often conflated arguments favour the core. First, at its very centre $\partial\rho/\partial r \to 0$, so the optical signal is locally weak and a pressure measurement is maximally complementary; second, the core hosts the largest pressure variation in the Rankine geometry, carrying the most thermodynamic information per tap. This placement is, however, an idealisation: it uses knowledge of the ground-truth field to position the taps, whereas in a real campaign the tap locations are fixed by the hardware rather than chosen adaptively. Each tap stores the triplet $(i, j, p_{i,j})$, with (i, j) the pixel indices and $p_{i,j}$ the local ground-truth pressure, and contributes a pointwise soft Dirichlet term to the total loss in the models developed in the following chapters.

Total-temperature constraint. The field $T_0(x, y)$ defined in Eq. (3.6) is saved on the whole evaluation domain.

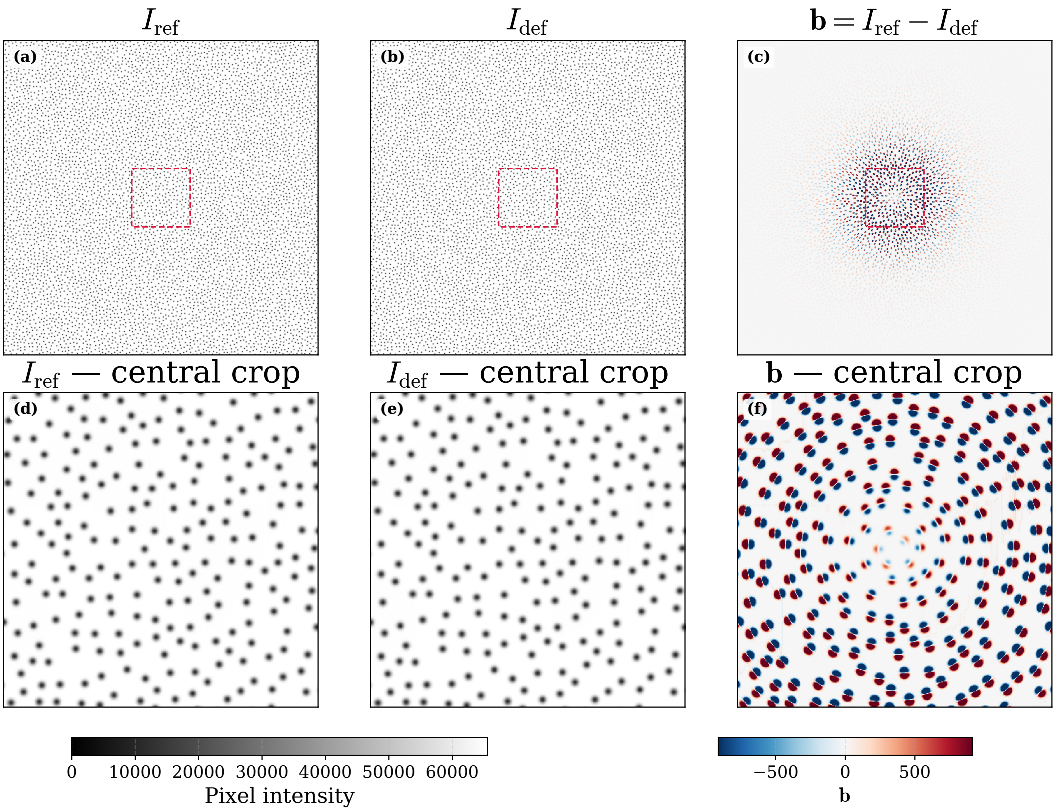


Figure 3.5: Synthetic BOS image pair and resulting training signal $\mathbf{b} = I_{\mathrm{ref}} - I_{\mathrm{def}}$. (a,b) Reference and deformed images (2048×2048 px), visually indistinguishable at sub-pixel maximum displacement (Fig. 3.4). (d,e) Central 380×380 px crops: backward warp (Eq. (3.10)) preserves blue-noise dot structure. (c,f) Training signal $\mathbf{b}$ on full domain and crop: annular flow-induced pattern around vortex; opposite-sign lobes reflect sub-pixel edge shifts (Eq. (3.11)). Intensities in native 16-bit range; $\mathbf{b}$ colour scale symmetric, clipped at 99^{th} percentile of $|\mathbf{b}|$.

Domain boundary conditions. A further mechanism allows pixelwise Dirichlet conditions on (ρ, u, v, p, T_0) on the sides of the frame, retained as a diagnostic tool: gradually reducing their spatial coverage maps the transition from a well-posed reconstruction (full inlet condition, as in Molnar et al. [26]) to the minimal VKI-inspired case.

The ground-truth fields are stored separately from the quantities legitimately available to the PINN models, so that post-training evaluation against the true state cannot contaminate the training pipeline.

It remains to verify that the two principal ingredients of the package, the signal $\mathbf{b}$ and the discrete operator meant to predict it, are mutually consistent: Section 3.2 formalises the linearised reference s and the agreement statistics, Section 3.3 maps the feasibility region and motivates the baseline, and Section 3.4 carries out the numerical validation.

3.2. Linearised Optical-Flow Model and Consistency Test Framework

The UBOS measurement loss of Eq. (2.27) treats the discrete operator A and the data vector $\mathbf{b}$ as independently assembled objects; verifying their numerical consistency in the sense of the linearised optical-flow identity $\mathbf{b} \approx s$ is a prerequisite to the PINN training of the next chapters. The present section assembles the framework for this verification: it formalises the analytical reference $s = I_x\delta_x + I_y\delta_y$, delimits the regime of background and flow parameters within which $\mathbf{b} \approx s$ is admissible, and introduces the three scalar statistics $(\alpha^\star, r_{\mathrm{bs}}(\mathbf{b}, s), \mathrm{RMSE}(\epsilon_{\mathrm{res}}))$ that quantify the agreement on the evaluation mask.

The Linear BOS Model: $s = I_x\,\delta_x + I_y\,\delta_y$

To verify that the signal $\mathbf{b}$ defined in Section 3.1 is consistent with the underlying BOS physics, an analytical prediction s that $\mathbf{b}$ should reproduce in the small-displacement limit must first be constructed. The row of the operator A corresponding to pixel (x, y) implements this prediction through the linearised optical-flow identity

$$s(x,y) = I_x(x,y)\,\delta_x(x,y) + I_y(x,y)\,\delta_y(x,y), \tag{3.12}$$

where s is the first-order prediction of the image change induced by the displacement (δ_x, δ_y). Equivalently, when the analytical deflection field is reconstructed from $\boldsymbol{\delta}_\alpha = C_{\mathrm{sys}} D_\alpha \boldsymbol{\rho}$ (Eq. (2.19)), the pointwise identity $s = I_x\delta_x + I_y\delta_y$ coincides with the corresponding row of $A\boldsymbol{\rho}$ in Eq. (2.20): the consistency test $\mathbf{b} \approx s$ therefore probes the assembled discrete form of the unified BOS operator without requiring the explicit construction of the matrix A.

The image gradients are computed from the reference image alone,

$$I_x = \mathcal{D}_x\, I_{\mathrm{ref}}, \qquad I_y = \mathcal{D}_y\, I_{\mathrm{ref}}, \tag{3.13}$$

where $\mathcal{D}_\alpha$ denotes the local one-dimensional convolution implementing the fourth-order, five-point, first-derivative stencil of Eq. (3.7), and is therefore distinct from the line-of-sight discrete-gradient operator D_α of Eq. (2.19), which acts as a sparse matrix on the discretised density coefficients. Using I_{ref} rather than the frame average $(I_{\mathrm{ref}} + I_{\mathrm{def}})/2$ preserves the statistical decoupling between operator and data that the experimental pipeline will require, since I_{ref} is a single, low-noise acquisition while I_{def} is contaminated by camera noise and optical drifts.

Validity Domain of the Linearised Approximation

The fundamental relation $\mathbf{b} \approx s$, that is, $I_{\text{ref}} - I_{\text{def}} \approx I_x\,\delta_x + I_y\,\delta_y$, is an approximation whose accuracy is conditioned by three independent requirements.

Small-displacement regime. The linearisation of $I_{\text{ref}}(X - \delta_x, Y - \delta_y)$ in Eq. (3.10) is a first-order Taylor expansion: it holds when the truncated quadratic term in I_{xx}, I_{xy}, I_{yy} is negligible relative to s, which requires $|\boldsymbol{\delta}| \ll \ell_{\text{grad}}$, with ℓ_{grad} the characteristic spatial scale of I_{ref}. With the parameters of Table 3.2, the peak displacement of the Rankine field remains below one pixel, placing the validation within the linear regime.

Consistent discretisation. As established at Eq. (3.7), the same five-point stencil differentiates image and density gradients, so that their truncation errors cancel insofar as the two fields share grid and sampling-to-feature-scale ratio; this is the requirement that fails when the background is left unblurred, as the sensitivity campaign shows below.

Background-to-displacement scale matching. Even at small displacement, $\mathbf{b} \approx s$ degrades if the dominant frequencies of I_{ref} approach $1/(2|\boldsymbol{\delta}|_{\max})$, since the gradient then varies appreciably over the displacement interval. This condition links the background design of Section The Reference Background Image to the flow parameters through the dimensionless ratio made explicit in Section 3.3.

These three requirements are not independent: together they define a feasibility region in the joint space of background parameters and flow parameters within which the UBOS measurement model is self-consistent. Within this region, the test statistic introduced in the next subsection quantifies numerically how closely $\mathbf{b}$ matches s; establishing the empirical boundaries of the region itself is the primary goal of the sensitivity analysis of Section 3.3.

Scalar Test Statistics on the Evaluation Mask

Within the validity domain identified in Section 3.2, the agreement between the training signal $\mathbf{b}$ and the linearised prediction s is quantified by three scalar statistics evaluated over the N_{mask} pixels of E_{val}. The first is the optimal scaling factor obtained as the least-squares slope of $\mathbf{b}$ on s,

$$\alpha^\star \;=\; \arg\min_{\alpha\in\mathbb{R}} \left\|\mathbf{b} - \alpha\, s\right\|^2_{2,E_{\text{val}}} \;=\; \frac{\langle \mathbf{b},\, s\rangle_{E_{\text{val}}}}{\langle s,\, s\rangle_{E_{\text{val}}}}, \tag{3.14}$$

where $\langle\cdot,\cdot\rangle_{E_{\text{val}}}$ denotes the inner product restricted to the evaluation mask. By construction, $\alpha^\star = 1$ is the value attained when the discrete implementation of A is in perfect numerical agreement with the data vector $\mathbf{b}$ assembled from the image pair; deviations from unity expose calibration mismatches in C_{sys}, sign-convention errors, or systematic stencil biases.

The second statistic is the Pearson correlation coefficient,

$$r_{\mathrm{bs}}(\mathbf{b}, s) \;=\; \frac{\langle \mathbf{b}-\bar{\mathbf{b}},\, s-\bar{s}\rangle_{E_{\mathrm{val}}}}{\left\|\mathbf{b}-\bar{\mathbf{b}}\right\|_{2,E_{\mathrm{val}}}\left\|s-\bar{s}\right\|_{2,E_{\mathrm{val}}}}, \tag{3.15}$$

which measures linear association independently of absolute scale and is therefore complementary to $\alpha^\star$: a calibration error shifts $\alpha^\star$ but leaves r_{bs} near unity, whereas a structural mismatch degrades r_{bs} regardless of scaling.

The third statistic is the root-mean-square of the residual $\epsilon_{\mathrm{res}} = \mathbf{b} - \alpha^\star s$,

$$\mathrm{RMSE}(\epsilon_{\mathrm{res}}) \;=\; \sqrt{\frac{1}{N_{\mathrm{mask}}}\sum_{(i,j)\in E_{\mathrm{val}}}\left[\mathbf{b}_{i,j}-\alpha^\star s_{i,j}\right]^2}, \tag{3.16}$$

which provides the absolute-magnitude counterpart to the dimensionless suppression ratio $\mathcal{S} \equiv \mathrm{RMS}(\mathbf{b})/\mathrm{RMSE}(\epsilon_{\mathrm{res}})$ reported in the figures of the following sections.

3.3. Sensitivity Analysis: Scope and Selection of the Baseline

The consistency triplet $(\alpha^\star, r_{\mathrm{bs}}(\mathbf{b}, s), \mathrm{RMSE}(\epsilon_{\mathrm{res}}))$ depends jointly on the spatial structure of I_{ref} (operator-side) and on the amplitude of $\boldsymbol{\delta}$ (data-side), so verifying consistency at a single ad-hoc point would expose only one slice of the response surface. The campaign of this section therefore maps the triplet across the synthesis parameter space, locates the feasibility-region boundaries empirically, and identifies a representative configuration safely inside it for the numerical validation of Section 3.4.

Experimental Design and Sensitivity Campaign

The campaign comprises 468 end-to-end runs. Two parameter families govern the consistency of the discrete UBOS operator through distinct mechanisms: the reference background I_{ref} sets the bandwidth of the gradients I_x, I_y populating A, hence the conditioning of the operator independently of the displacement; the Rankine parameters (v_c, r_c) set the amplitude and scale of $\boldsymbol{\delta}(x, y)$, hence the strength of the higher-order Taylor terms truncated by the linearisation of Section 3.2.

The campaign organises these dependencies along three axes, *background-driven*, *flow-driven* and a *cross-axis* probing their coupling through the dimensionless ratio below. The first two axes are explored in two complementary modes. Five one-dimensional sweeps perturb each governing parameter individually, keeping all the others at the baseline values of Section 3.1: the minimum dot separation $d_{\text{min}} \in [4, 45]$ px (15 samples), the Gaussian-blur standard deviation $\sigma_{\text{blur}} \in [0, 5]$ px (12 samples), the dot radius $r_{\text{dot}} \in [1, 5]$ px (5 samples), the peak core velocity $v_c \in [40, 240]$ m/s (15 samples), and the core radius $r_c \in [5, 20]$ mm (12 samples).

Two iso-grids then explore the joint variation of the dominant pair within each family: a $15 \times 12 = 180$-run grid on the $(d_{\text{min}}, \sigma_{\text{blur}})$ plane at fixed flow, and a $15 \times 12 = 180$-run grid on the (v_c, r_c) plane at fixed background. These iso-grids supply the two-dimensional iso-contour deliverables on which the feasibility regions of the next subsections are defined.

The third axis is realised by an explicit cross-axis sweep. Seven values of v_c uniformly spanning the flow range are paired with seven values of d_{min} uniformly spanning the background range, producing 49 runs along which both families vary simultaneously. This sweep is constructed so that the dominant background scale and the dominant flow scale change together, generating the broadest range of values of the dimensionless ratio defined below.

For every run the synthetic pipeline of Section 3.1 is executed in full, the consistency triplet $(\alpha^\star, r_{\text{bs}}(\mathbf{b}, s), \text{RMSE}(\epsilon_{\text{res}}))$ together with the suppression ratio $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}})$ and the displacement statistics $|\boldsymbol{\delta}|_{\text{max}}$, $P_{99}|\delta|$ are evaluated on the full mask, and the dimensionless ratio introduced below is computed for that configuration.

The dimensionless ratio is defined as

$$\xi = \frac{|\boldsymbol{\delta}|_{\text{max}}}{\ell_{\text{grad}}}, \qquad \ell_{\text{grad}} = \sqrt{d_{\text{min}}^2 + (2\,\sigma_{\text{blur}})^2}, \tag{3.17}$$

where the numerator is the peak pixel-space deflection produced by the Rankine field on the full mask, and the denominator is the characteristic spatial scale over which I_{ref} varies, constructed from the inter-dot spacing d_{min} combined in quadrature with the blur bandwidth $2\,\sigma_{\text{blur}}$.

The form of ℓ_{grad} embeds the two background mechanisms that contribute to the gradient scale: a sub-blur regime in which the spacing between dots dominates ($d_{\text{min}} \gg \sigma_{\text{blur}}$), and a blur-dominated regime in which the kernel bandwidth sets the scale ($\sigma_{\text{blur}} \gg d_{\text{min}}$). The quadrature combination is a heuristic estimate, valid in the band-limited regime $\sigma_{\text{blur}} \gtrsim 1$ px; in the limit $\sigma_{\text{blur}} \to 0$ it returns $\ell_{\text{grad}} \to d_{\text{min}}$, whereas the gradient scale that actually controls the linearisation there is set by the unresolved dot-edge width, not by $d_{\min}$. As anticipated here and confirmed in Section 3.3, this is the reason the $\sigma_{\text{blur}} = 0$ runs do not collapse onto the common ξ curve. The ratio ξ therefore quantifies, in a single dimensionless number, the relative weight of the displacement amplitude with respect to the spatial scale at which I_{ref} becomes Taylor-resolvable, and is the natural control parameter for the second-order Taylor argument of Section 3.2.

By design, the linearised optical-flow identity is asymptotically exact in the limit $\xi \to 0$, up to a numerical floor set by interpolation and discretisation that bounds the attainable suppression at small amplitude, and it degrades when $\xi \to \mathcal{O}(1)$. The numerical verification of this expectation is the object of the cross-axis discussion at the end of this section.

Background-Driven Sensitivity

The background-driven axis is examined while the Rankine parameters (v_c, r_c) are held at their baseline values. Under this constraint the displacement field, and therefore the numerator of ξ, is identical for every run reported in the present subsection: the only quantity that varies across runs is ℓ_{grad}, and any change in $(\alpha^\star, r_{\text{bs}}, \text{RMSE}(\epsilon_{\text{res}}))$ must be attributed to the structure of the operator A rather than to an amplitude effect. The one-dimensional behaviour along $d_{\min}$ and σ_{blur} is shown in Fig. 3.6. The two background axes display markedly different sensitivity profiles. Along $d_{\min}$, both $\alpha^\star$ and the suppression ratio are essentially flat over the explored range $[4, 45]$ px: the slope $\alpha^\star$ remains within a few $\times 10^{-3}$ of unity in every run and the suppression ratio remains in the band $11.8 \lesssim \text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}}) \lesssim 12.3$. The weakness of the $d_{\min}$ dependence is the operational signature of the requirement for a dense and isotropic gradient field of Section 3.1: as long as the blue-noise dot distribution remains non-degenerate, the spatial isotropy of ∇I_{ref} guarantees that no extended row of A is silenced, and the conditioning of the operator is preserved at the aggregate level. Along σ_{blur}, by contrast, the suppression profile is strongly non-monotonic. At $\sigma_{\text{blur}} = 0$ the discrete optical-flow identity collapses, with $\alpha^\star = 0.737$, $r_{\text{bs}}(\mathbf{b}, s) = 0.694$ and a suppression ratio of only 1.39; the suppression rises rapidly with σ_{blur}, crosses the value 10 at $\sigma_{\text{blur}} \simeq 1.4$ px, peaks at $\sigma_{\text{blur}} \simeq 2.7$ px with $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}}) = 12.66$, and decays back below 10 for $\sigma_{\text{blur}} \gtrsim 3.7$ px. A localised overshoot of $\alpha^\star$ to $\simeq 1.035$ is observed at $\sigma_{\text{blur}} \approx 0.45$ px, in the transition between the discretisation-failure regime at $\sigma_{\text{blur}} = 0$ and the bandlimited plateau; this transient excursion remains well within the $|\alpha^\star - 1| < 0.05$ linearity band and is consistent with the partial recovery of stencil resolvability as the dot edges begin to be smoothed.

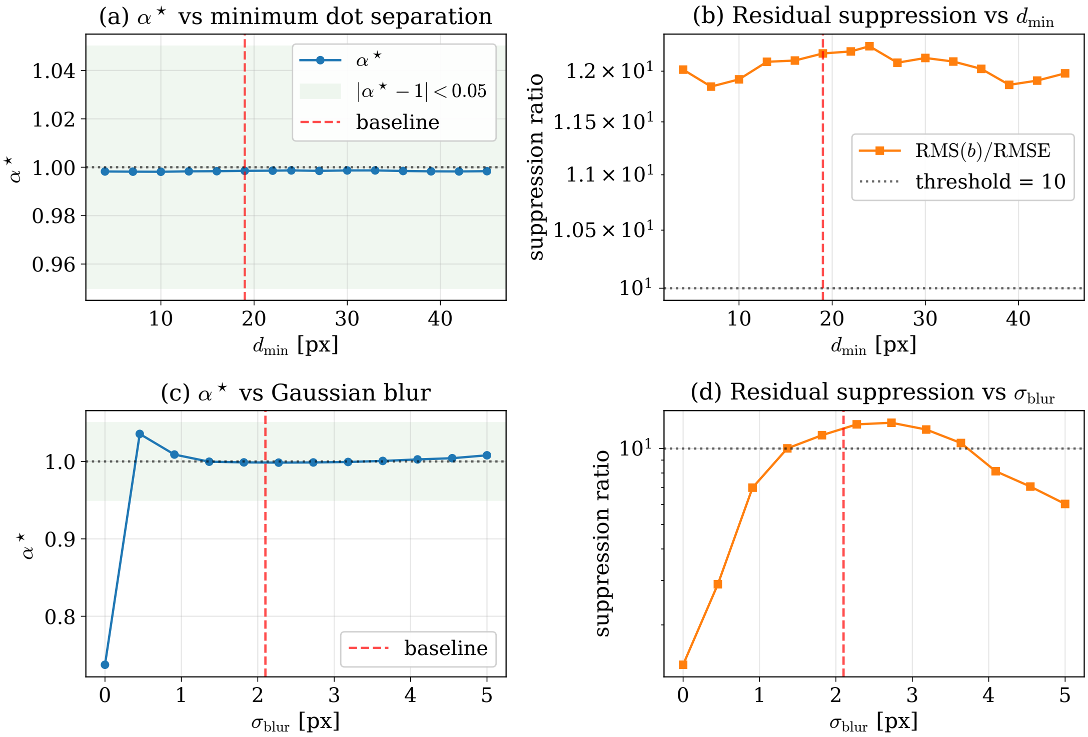


Figure 3.6: One-dimensional background sweeps: minimum dot separation $d_{\min} \in [4, 45]$ px (top), Gaussian-blur $\sigma_{\text{blur}} \in [0, 5]$ px (bottom) with baseline $v_c = 130$ m/s, $r_c = 10$ mm. (a,c): Scaling factor $\alpha^\star$ (Eq. (3.14), green: $|\alpha^\star - 1| < 0.05$). (b,d): Suppression ratio $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}})$ (log scale, dotted: threshold 10). $d_{\min}$ yields a flat plateau, while σ_{blur} shows a sharp drop at 0 and a peak near 2.7 px.

Mechanistically, the unblurred case violates the consistent-discretisation requirement of Section 3.2: the binary dot edges are not resolved by the five-point stencil of Eq. (3.7), so that the finite-difference I_x, I_y are dominated by aliasing rather than by the local image gradient that the linearisation of Eq. (3.12) assumes. The high-blur degradation, by contrast, is the asymptotic manifestation of the same scale-matching argument: when σ_{blur} approaches the inter-dot half-spacing $d_{\min}/2 = 9.5$ px, neighbouring dots merge and the gradient field loses contrast, concentrating a heavy-tailed residual in the narrow annulus where the local image gradient ∇I_{ref} is poorly resolved. Between these two failure modes lies the band $\sigma_{\text{blur}} \in [1.4, 3.7]$ px within which the suppression ratio remains above the $10 : 1$ threshold; the baseline value $\sigma_{\text{blur}} = 2.1$ px falls inside this band, in agreement with the bandwidth argument of Section 3.1 that fixes $r_{\text{dot}} + 2\sigma_{\text{blur}} \approx 6.2$ px below $d_{\min}/2$.

The joint behaviour of the two background parameters is reported in Fig. 3.7. The $(d_{\min}, \sigma_{\text{blur}})$ plane displays the iso-stratification anticipated by the one-dimensional sweeps: the $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}})$ iso-contours are nearly horizontal, depending almost exclusively

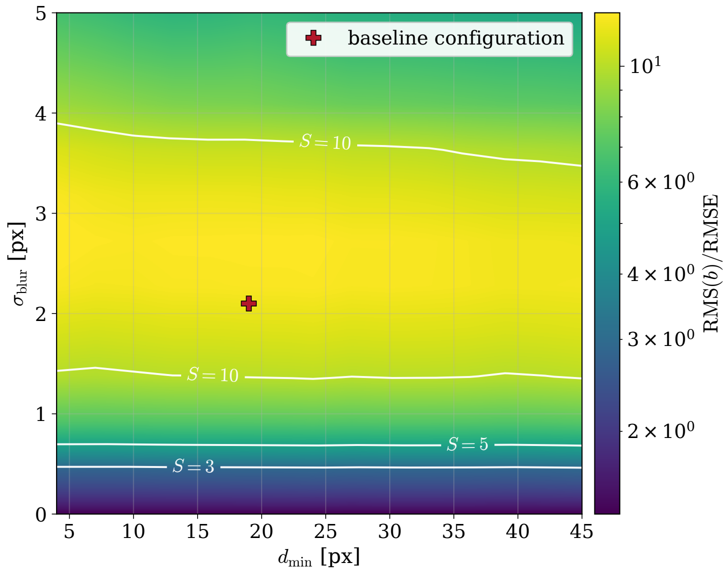


Figure 3.7: Joint background sensitivity on $(d_{\mathrm{min}}, \sigma_{\mathrm{blur}})$ iso-grid $(15 \times 12 = 180$ runs); flow fixed at $v_c = 130$ m/s, $r_c = 10$ mm. Colour: suppression ratio $\mathrm{RMS}(\mathbf{b})/\mathrm{RMSE}(\epsilon_{\mathrm{res}})$ (log scale). White iso-contours: levels $\{3, 5, 10, 20, 40\}$. Red cross: baseline $(19, 2.1)$ px. Nearly horizontal contours show suppression is mainly controlled by σ_{blur}; $\mathrm{RMS}(\mathbf{b})/\mathrm{RMSE}(\epsilon_{\mathrm{res}}) > 10$ for $\sigma_{\mathrm{blur}} \in [1.4, 3.7]$ px.

on σ_{blur} over the full range of d_{min}. The $\mathrm{RMS}(\mathbf{b})/\mathrm{RMSE}(\epsilon_{\mathrm{res}}) = 10$ iso-contour delimits a broad horizontal stripe extending across the whole sampled $d_{\mathrm{min}} \in [4, 45]$ px for σ_{blur} between $\simeq 1.4$ and $\simeq 3.7$ px, with the peak suppression of the grid attained at $(d_{\mathrm{min}}, \sigma_{\mathrm{blur}}) = (4, 2.73)$ px and the worst suppression ratio attained again at $\sigma_{\mathrm{blur}} = 0$, where the slope drops to $\alpha^\star = 0.738$. The baseline pair $(19, 2.1)$ px falls inside the favourable region, where the worst-case sensitivity to either coordinate is dominated by the σ_{blur} direction.

The combined evidence of Figures 3.6 and 3.7 settles two distinct points relevant for the next chapters. First, the failure modes of the background pattern are not symmetric: the discretisation breakdown at $\sigma_{\mathrm{blur}} \to 0$ is sharper than the bandwidth-loss degradation at $\sigma_{\mathrm{blur}} \gg d_{\mathrm{min}}/2$, so that the operational margin is determined by the lower bound on σ_{blur}. Second, the dot-spacing parameter d_{min} is a benign degree of freedom over the full sampled range; this is the formal counterpart of the qualitative argument for a dense gradient field anticipated in Section 3.1, and licenses the use of the baseline $d_{\mathrm{min}} = 19$ px as a representative value rather than a sharp tuning point.

Flow-Driven Sensitivity

The flow-driven axis is examined under the symmetric protocol: the background pattern is held at its baseline configuration, so that ℓ_{grad} is identical for every run, and only the displacement field varies across runs. Any change in $(\alpha^\star, r_{\text{bs}}, \text{RMSE}(\epsilon_{\text{res}}))$ must therefore be attributed to the amplitude of $\boldsymbol{\delta}$ and to the spatial scale on which it concentrates, both of which are controlled by the Rankine parameters (v_c, r_c) through Eqs. (3.1) and (3.5). The one-dimensional behaviour is shown in Fig. 3.8.

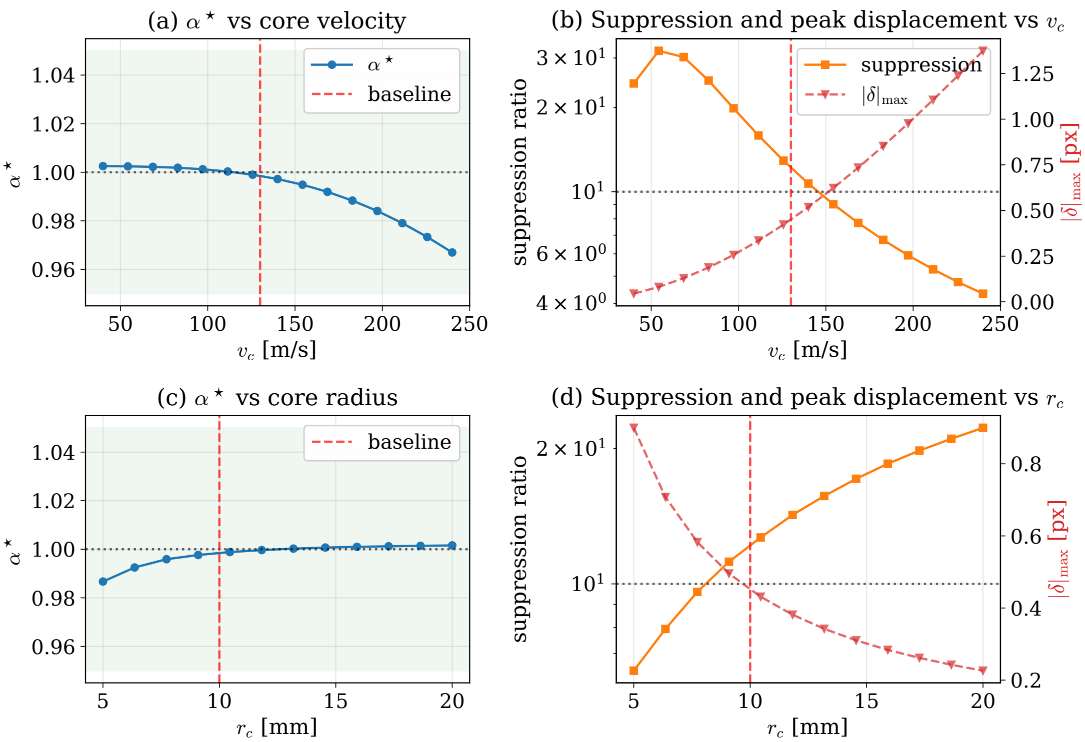


Figure 3.8: One-dimensional flow parameter sweeps: $v_c \in [40, 240]$ m/s (top), $r_c \in [5, 20]$ mm (bottom); background fixed at $(d_{\min}, \sigma_{\text{blur}}) = (19, 2.1)$ px. (a,c): Scaling factor $\alpha^\star$ (Eq. (3.14), green: $|\alpha^\star - 1| < 0.05$). (b,d): Suppression ratio $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}})$ (log scale), with peak displacement $|\boldsymbol{\delta}|_{\max}$ (red dashed, right axis). Baseline marked by dashed vertical. Both parameters control $|\boldsymbol{\delta}|_{\max}$, driving monotonic deterioration in suppression for larger v_c or smaller r_c.

The two flow axes drive the consistency triplet through $|\boldsymbol{\delta}|_{\max}$ in opposite directions but through the same mechanism. As v_c increases from 40 to 240 m/s the peak displacement grows from 0.044 to 1.37 px and the suppression ratio decreases monotonically over the bulk of the range, from its largest sampled value of 31.8 down to 4.33 at the unfavourable end. The maximum is attained at the interior sample $v_c = 54.3$ m/s rather than at $v_c = 40$ m/s, where the residual is already dominated by the numerical floor. Meanwhile

$\alpha^\star$ leaves the unit neighbourhood for $v_c \gtrsim 150$ m/s, settling at 0.967 at $v_c = 240$ m/s. Symmetrically, as r_c decreases from 20 to 5 mm the peak displacement grows from 0.23 to 0.90 px and the suppression ratio decreases from 22.2 to 6.41, with $\alpha^\star$ dropping to 0.987 at the unfavourable end. The two responses are the direct optical signature of the second-order term in the Taylor expansion of Section 3.2: because both an increase of v_c and a decrease of r_c raise the amplitude $|\boldsymbol{\delta}|$ relative to the fixed ℓ_{grad}, the truncated quadratic contribution $\frac{1}{2}[\delta_x^2 I_{xx} + 2\delta_x\delta_y I_{xy} + \delta_y^2 I_{yy}]$ becomes non-negligible relative to s, and the linearisation degrades.

The baseline configuration $v_c = 130$ m/s, $r_c = 10$ mm produces $|\boldsymbol{\delta}|_{\max} = 0.45$ px and a suppression ratio of 12.18: the margin in the suppression ratio itself is therefore modest on both flow axes, even though, as shown below, the margin in the control parameters ξ and σ_{blur} is comparatively wide.

The joint behaviour of the two flow parameters is reported in Fig. 3.9. In contrast with the iso-stratified structure of the background grid, the iso-contours of the (v_c, r_c) heat-map are clearly diagonal: they sweep from the upper-left corner (low v_c, large r_c) down to the lower-right corner (high v_c, small r_c) following a single one-parameter family. This geometry is the direct visual confirmation that the relevant control parameter is not v_c or r_c individually but their combination through $|\boldsymbol{\delta}|_{\max}$, which in turn depends on the radial gradient of ρ produced by Eq. (3.5) and is therefore driven by the combination of v_c and r_c that enters the radial momentum balance. The grid extremes confirm and extend the one-dimensional findings. The best joint configuration is $(v_c, r_c) = (68.6 \text{ m/s}, 20 \text{ mm})$ with $|\boldsymbol{\delta}|_{\max} = 0.065$ px, $\alpha^\star = 1.0025$ and a suppression ratio of 38.7, almost a factor four above the baseline; the worst joint configuration is $(v_c, r_c) = (240 \text{ m/s}, 5 \text{ mm})$ with $|\boldsymbol{\delta}|_{\max} = 2.74$ px, $\alpha^\star = 0.874$ and a suppression ratio of only 2.32, well below the feasibility threshold.

The amplitude of the displacement field varies by more than a factor of 40 across the grid, while the operator A remains fixed: this is the cleanest experimental separation of the two failure modes of $\mathbf{b} \approx s$ available to the present pipeline, and isolates the amplitude-driven mechanism that the next subsection analyses through the dimensionless ratio ξ.

Amplitude-Driven Collapse and the Linearity Regime

The two preceding subsections have isolated, axis by axis, the operator-side and the displacement-side contributions to the breakdown of the linearisation $\mathbf{b} \approx s$. The dimensionless ratio ξ defined in Eq. (3.17) was constructed precisely to merge these two contributions into a single control parameter: the question of whether ξ effectively organ-

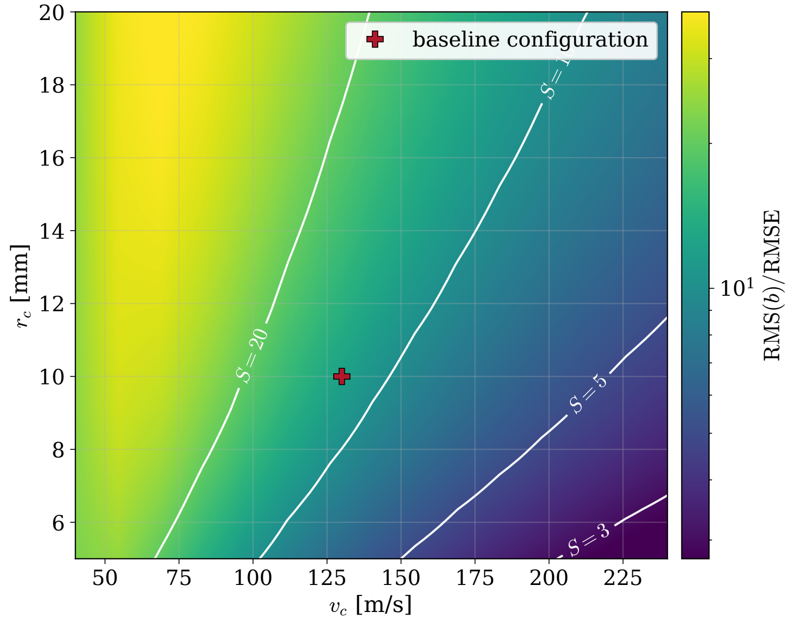


Figure 3.9: Joint flow sensitivity on (v_c, r_c) iso-grid $(15 \times 12 = 180)$, background fixed at $(19, 2.1)$ px. Colour: suppression ratio $\mathrm{RMS}(\mathbf{b})/\mathrm{RMSE}(\epsilon_{\mathrm{res}})$ (log scale); white contours: $\{3, 5, 10, 20, 40\}$. Red cross: baseline (130 m/s, 10 mm). Diagonal contours indicate suppression depends on combined v_c and r_c, not individually.

ises the cross-axis behaviour of the consistency triplet is addressed by replotting all 468 runs of the campaign on a common ξ axis, regardless of which family they belong to. The result is shown in Fig. 3.10.

The collapse onto the ξ axis is partial but informative. In panel (a), the runs of every sweep family align onto a common master curve in which $\alpha^\star$ remains in the unit neighbourhood for $\xi \lesssim 0.05$, departs smoothly from unity for $\xi \gtrsim 0.05$ and crosses the lower edge of the linearity band $|\alpha^\star - 1| = 0.05$ around $\xi \sim 0.07{-}0.1$. In panel (b), the same runs lie on a single descending curve in log–log coordinates, with the high-ξ tail consistent with a power-law slope close to -1, in line with the second-order Taylor prediction of Section 3.2. Explicitly, the first-order term scales as $\mathrm{RMS}(\mathbf{b}) \sim |\boldsymbol{\delta}|\, \mathrm{RMS}(\nabla I_{\mathrm{ref}})$, while the truncated second-order term contributes a residual of order $\mathrm{RMSE}(\epsilon_{\mathrm{res}}) \sim |\boldsymbol{\delta}|^2\, \mathrm{RMS}(\nabla^2 I_{\mathrm{ref}})$. Using the dimensional estimate $\mathrm{RMS}(\nabla^2 I_{\mathrm{ref}}) \sim \mathrm{RMS}(\nabla I_{\mathrm{ref}})/\ell_{\mathrm{grad}}$, the ratio of the two scales becomes

$$\frac{\mathrm{RMS}(\mathbf{b})}{\mathrm{RMSE}(\epsilon_{\mathrm{res}})} \sim \frac{|\boldsymbol{\delta}|\, \mathrm{RMS}(\nabla I_{\mathrm{ref}})}{|\boldsymbol{\delta}|^2\, \mathrm{RMS}(\nabla I_{\mathrm{ref}})/\ell_{\mathrm{grad}}} = \frac{\ell_{\mathrm{grad}}}{|\boldsymbol{\delta}|} = \xi^{-1}, \tag{3.18}$$

recovering the ξ^{-1} envelope traced by the high-ξ tail of panel (b): the dimensional argument fixes the slope at an exponent of -1, while the intercept is not, since it absorbs the ratio $\mathrm{RMS}|\boldsymbol{\delta}|/|\boldsymbol{\delta}|_{\max}$ and the curvature factors of $\nabla^2 I_{\mathrm{ref}}$. The agreement with this enve-

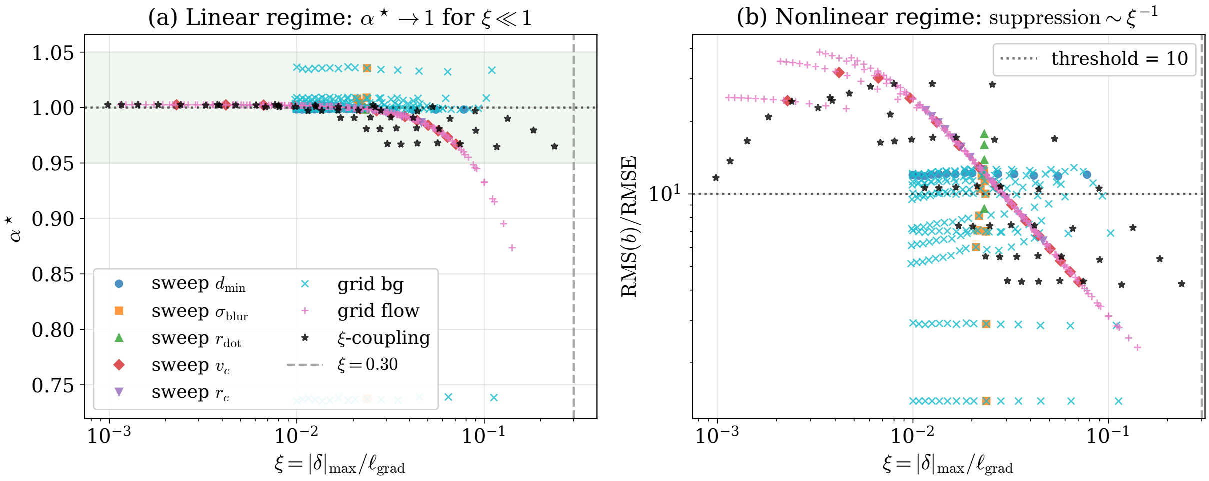


Figure 3.10: Consistency metrics collapse vs. $\xi = |\boldsymbol{\delta}|_{\max}/\ell_{\text{grad}}$ across all 468 parameter runs. (a) Scaling factor $\alpha^\star$; green: $|\alpha^\star - 1| < 0.05$. (b) Suppression ratio $\text{RMS}(\mathbf{b})/\text{RMSE}(\epsilon_{\text{res}})$ (log–log). Collapse occurs along amplitude axis; off-curve points at low ξ are $\sigma_{\text{blur}} = 0$ (operator-side breakdown).

lope, together with the alignment of the seven sweep families along it, is strong evidence that ξ is the correct amplitude-side control parameter for the linearisation regime. The ξ^{-1} relation describes this intermediate regime rather than the $\xi \to 0$ limit: as ξ decreases the suppression does not grow without bound but saturates against the interpolation and discretisation floor of Section Backward Warping and Construction of I_{def}, which is why the largest sampled suppression occurs at an interior point of the flow sweep.

The limit of the collapse is equally explicit. A cluster of points in the lower-left of panel (a) and at the bottom of panel (b) sits well off the master curve, with $\alpha^\star \simeq 0.74$ at $\xi \simeq 0.024$ and suppression ratios near 1.4: these are the $\sigma_{\text{blur}} = 0$ runs of the σ_{blur} sweep and of the background grid, in which the bandwidth requirement of Section 3.2 is violated. They also produce the linearity break at $\xi = 0.010$, $\alpha^\star = 0.735$: the heuristic criterion of Section 3.3 thus first activates on a *small*-ξ point, but the underlying mechanism is not an amplitude-driven Taylor breakdown, it is an operator-side discretisation failure in A, in which ∇I_{ref} itself is not properly resolved by the five-point stencil. The collapse is therefore universal along the amplitude axis but not along the bandwidth axis.

The aggregate picture is two-dimensional. Within the bandlimited regime $\sigma_{\text{blur}} \gtrsim 1$ px, ξ orders the linearisation quality and identifies the linear regime as $\xi \lesssim 0.05$; outside it the operator fails irrespective of ξ, as in the $\sigma_{\text{blur}} = 0$ runs. The two regimes correspond, respectively, to the small-displacement and consistent-discretisation conditions of Section 3.2.

The campaign thus delimits the joint feasibility region of the discrete UBOS pipeline and fixes a baseline well inside it, carried over to the numerical validation of Section 3.4: $d_{\min} = 19$ px, $\sigma_{\text{blur}} = 2.1$ px (background), $v_c = 130$ m/s, $r_c = 10$ mm (flow). This sits at $\xi \simeq 2.3 \times 10^{-2}$ and inside the bandlimited σ_{blur}-band $[1.4, 3.7]$ px over which the suppression ratio stays above the $10 : 1$ threshold. The wide margin (a factor of two to four) thus lies in the control parameters ξ and σ_{blur}, whereas the suppression-ratio margin is the more modest $\sim 1.2\times$ noted in Section 3.3. The linear regime $\xi \lesssim 0.05$ is intrinsically sub-pixel: the baseline $|\boldsymbol{\delta}|_{\max} = 0.45$ px and the whole feasibility band correspond to displacements well below one pixel. The O(pixel) displacements and the sharp gradients produced by shocks in transonic and supersonic cascade flows fall outside this band and are not represented by the present benchmark; the consequences of this gap are stated in the limitations below. The margins in ξ and σ_{blur} ensure that the numerical consistency assessed in the next section is representative of the bulk of the feasibility region rather than of a marginal point, and provide the methodological basis for the PINN benchmarking of the following chapters.

3.4. Numerical Validation of the Discrete UBOS Operator

The framework of Section 3.2 and the sensitivity campaign of Section 3.3 together fix the conditions for consistency between the discrete UBOS operator and the data vector $\mathbf{b}$ and the baseline selected inside the feasibility region ($\xi \simeq 2.3 \times 10^{-2}$). On this baseline the present section evaluates the consistency triplet $(\alpha^\star, r_{\text{bs}}(\mathbf{b}, s), \text{RMSE}(\epsilon_{\text{res}}))$ defined in Section 3.2, reports its observed values, and discusses the diagnostic use of the test statistics in detecting implementation-level inconsistencies that the sensitivity sweeps are, by construction, blind to (sign conventions, C_{sys} miscalibration, stencil mismatches between ρ and I_{ref}).

Observed Behaviour at the Selected Baseline

When the synthetic pipeline is executed at the baseline configuration selected in Section 3.3, i.e. with the parameters of Section 3.1 and the nominal optical constants of Table 3.2, the fields $\mathbf{b}$ and s exhibit strong pointwise agreement over the evaluation mask ($N_{\text{mask}} = 2048 \times 2048 = 4{,}194{,}304$ pixels). Fig. 3.11 compares the two maps with the residual $\epsilon_{\text{res}} = \mathbf{b} - \alpha^\star s$: $\mathbf{b}$ and s share the same annular support around the vortex core and exhibit nearly identical magnitudes, with $\text{RMS}(s) = 1.93 \times 10^2$ grey levels.

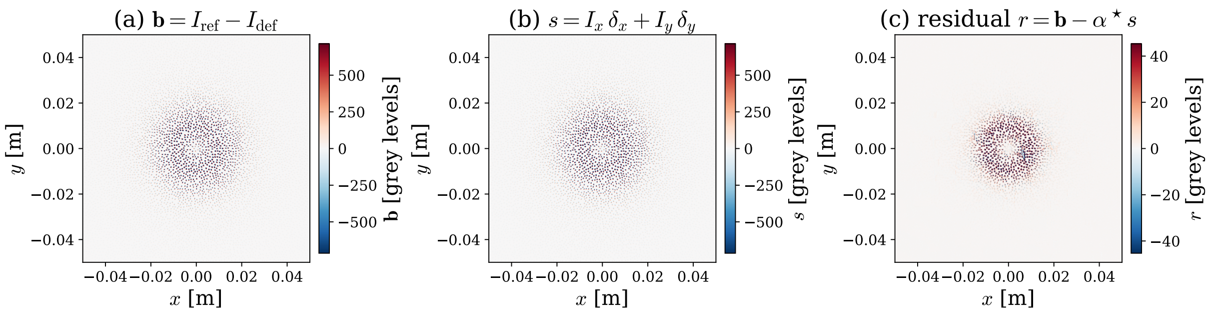


Figure 3.11: Discrete UBOS operator consistency on Rankine synthetic data. (a) Training signal $\mathbf{b} = I_{\text{ref}} - I_{\text{def}}$. (b) Optical-flow prediction $s = I_x\,\delta_x + I_y\,\delta_y$ (using I_{ref} gradients and analytical deflection). (c) Residual $\epsilon_{\text{res}} = \mathbf{b} - \alpha^\star s$ (independent colour scale; reported: $\text{RMS}(\mathbf{b}) = 1.93{\times}10^2$, $\text{RMSE}(\epsilon_{\text{res}}) = 1.585{\times}10^1$, suppression 12.18.) (a,b): same diverging colour scale, clipped at 99^{th} percentile of $|\mathbf{b}|$; values in 16-bit grey levels.

The pointwise behaviour of Fig. 3.11 is quantified by the joint distribution of $\mathbf{b}$ and s on the evaluation mask, shown in Fig. 3.12. The cloud collapses onto the bisector with negligible scatter: the least-squares slope of Eq. (3.14) evaluates to $\alpha^\star = 0.9985$, within 1.5×10^{-3} of unity, the Pearson correlation of Eq. (3.15) attains $r_{\text{bs}}(\mathbf{b}, s) = 0.9966$ (so that $R^2 \simeq r_{\text{bs}}^2 \simeq 0.993$, the two coinciding here because the fit has no intercept), and the residual RMSE of Eq. (3.16) is $\text{RMSE}(\epsilon_{\text{res}}) = 1.585 \times 10^1$ grey levels, matching the residual map of Fig. 3.11(c) and lying approximately one order of magnitude below $\text{RMS}(\mathbf{b})$. These three indicators jointly satisfy the consistency criterion of Section 3.2.

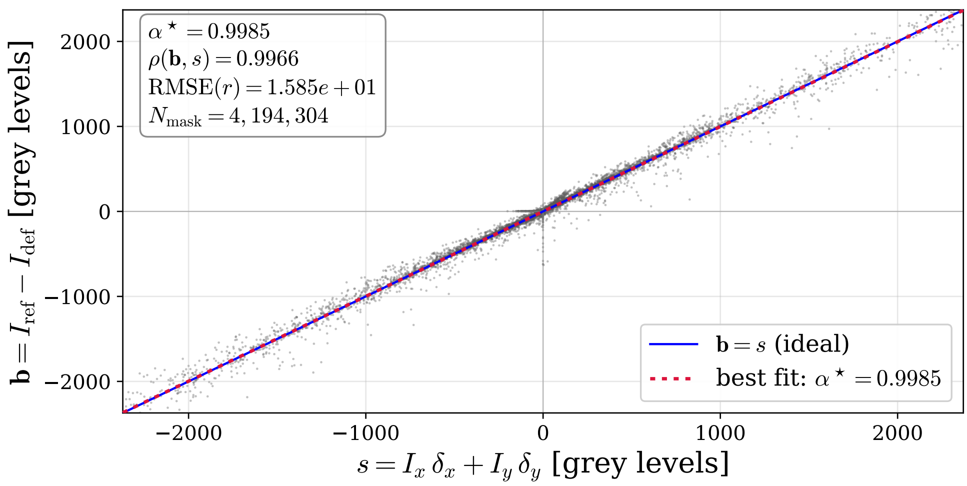


Figure 3.12: Pixelwise scatter of $\mathbf{b}$ vs. $s = I_x\,\delta_x + I_y\,\delta_y$ on the evaluation mask. Legend: best-fit slope $\alpha^\star = 0.9985$, Pearson correlation $r_{\text{bs}} = 0.9966$, residual RMSE $= 1.585 \times 10^1$ grey levels (full mask). Plotted: 2×10^5 points for readability. Solid: ideal $\mathbf{b} = s$; dashed: best linear fit ($\alpha^\star$).

The residual distribution and the corresponding L^2 suppression are reported in Fig. 3.13. The histogram of ϵ_{res} is sharply peaked around zero, with $\overline{\epsilon_{\text{res}}} \ll \sigma_{\epsilon_{\text{res}}}$ and $\sigma_{\epsilon_{\text{res}}} = 1.58 \times 10^1$ grey levels (so $\text{RMSE}(\epsilon_{\text{res}}) \simeq \sigma_{\epsilon_{\text{res}}}$); the bulk is approximately Gaussian, but the upper tail is heavier than Gaussian and asymmetric, with the 99.9^{th} percentile of $|\epsilon_{\text{res}}|$ reaching $\simeq 4.6\times$ the Gaussian envelope and $1.24\,\text{RMS}(\mathbf{b})$. This excess identifies a small population ($\lesssim 0.1\%$ of pixels) confined to a thin annulus around the rigid-core boundary, where the linearisation $\mathbf{b} \approx s$ is most strained.

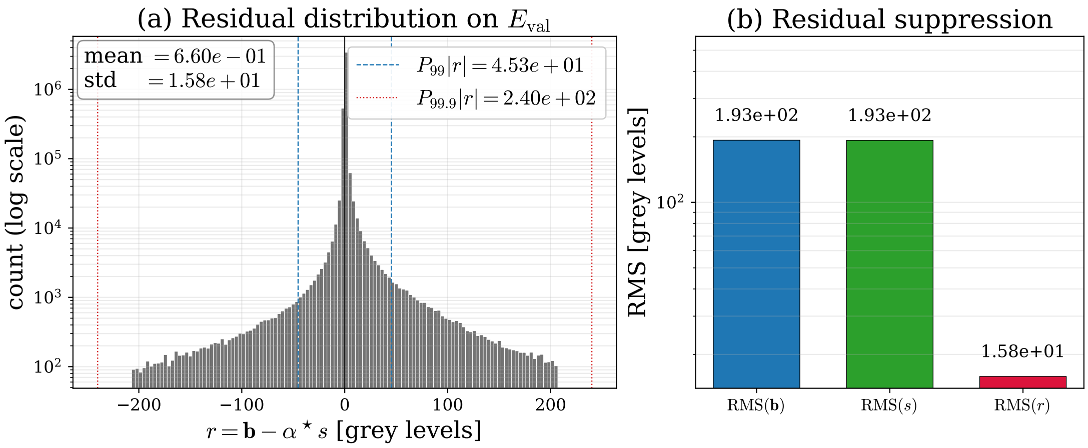


Figure 3.13: Residual statistics on the evaluation mask. (a) Histogram of $\epsilon_{\text{res}} = \mathbf{b} - \alpha^\star s$ (log counts). Inset: mean 6.60×10^{-1}, st. dev. 1.58×10^1; markers: 99^{th} ($|\epsilon_{\text{res}}| = 4.53 \times 10^1$) and 99.9^{th} ($|\epsilon_{\text{res}}| = 2.40 \times 10^2$) percentiles (heavy tail at rigid-core). (b) Comparison: $\text{RMS}(\mathbf{b}) = 1.93 \times 10^2$, $\text{RMS}(s) = 1.93 \times 10^2$, $\text{RMSE}(\epsilon_{\text{res}}) = 1.58 \times 10^1$ (log scale, 16-bit grey-levels).

The azimuthally averaged profiles of $\mathbf{b}$, s and ϵ_{res} as a function of r/r_c, shown in Fig. 3.14, confirm the picture. $\text{RMS}(\mathbf{b})$ and $\text{RMS}(s)$ are visually indistinguishable across the full radial extent, and $\text{RMS}(\epsilon_{\text{res}})$ remains uniformly suppressed below them by approximately one order of magnitude, with no localised loss of suppression: the consistency check is therefore not only globally satisfied but spatially uniform in the relative sense across the support of the deflection field. In absolute terms, $\text{RMSE}(\epsilon_{\text{res}})$ peaks at $r = r_c$, where both the deflection field and the linearisation strain attain their maxima: this co-location of signal and residual peaks is an intrinsic feature of the optical-flow model, and motivates the placement of the supplemental pressure-tap constraints described in Section 3.1.

This successful check, with the feasibility map of Section 3.3, justifies using the synthetic dataset as input to the PINN training of the next chapters: the measurement side of the UBOS/PINN loss is internally consistent at a baseline safely inside the joint feasibility region.

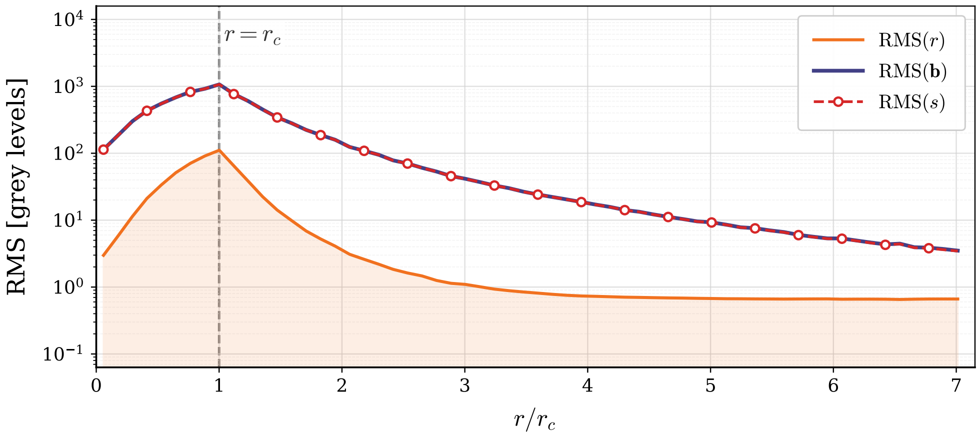


Figure 3.14: Azimuthally averaged RMS of $\mathbf{b}$, s, and residual $\epsilon_{\mathrm{res}} = \mathbf{b} - \alpha^\star s$ versus r/r_c. Dotted line: rigid-core boundary. RMS($\mathbf{b}$) and RMS(s) coincide at all radii; residual is uniformly suppressed, confirming no locally degraded consistency that could be missed by domain-averaged statistics.

3.5. Summary

This chapter has built and validated the synthetic benchmark grounding the next chapters: the compressible Rankine model supplies the full ground-truth state on a 2048×2048 grid, the image-synthesis pipeline produces $(I_{\mathrm{ref}}, I_{\mathrm{def}})$ and the signal $\mathbf{b}$, and the training package bundles these with (I_x, I_y), the T_0 map, the N_{p_s} pressure constraints and the ground-truth fields (stored separately for post-training evaluation only). The sensitivity campaign maps the feasibility region and motivates the baseline at $\xi \simeq 2.3 \times 10^{-2}$, at which operator and data vector are consistent at the 10^{-3} level on $|1-\alpha^\star|$ and $1-r_{\mathrm{bs}}(\mathbf{b}, s)$, with a suppression ratio of 12.18.

Several limitations bound this scope. The validation is purely synthetic (closed-form ground truth isolates pipeline and operator errors from reconstruction errors, but experimental VKI imagery is deferred and Table 3.2 only reproduces the order of magnitude of the configuration of Section 2.4); the measurement model is two-dimensional (collapsing the line-of-sight integral to a product with ΔZ_{grad} assumes out-of-plane coherence, consistent with the $\eta = 0$ Euler system, and neglects ray curvature and three-dimensional integration); the baseline is noise-free (the additive model of Section Backward Warping and Construction of I_{def} is disabled); the equation-of-state closure and K are shared between ground truth and physics loss by design, so the gauge-breaking is imposed rather than discovered, its persistence under measurement-based K and shock-bearing flows being undemonstrated here; and the dataset is a single realisation.

4 | Shared PINN Framework

The unified BOS forward model of Chapter 2 and the synthetic dataset of Chapter 3 delimit the inverse problem addressed here: the recovery of a two-dimensional steady compressible density field from a single BOS image pair. The three strategies of Chapters 5–7 instantiate one UBOS/PINN framework, sharing the flow-state representation, discrete unified BOS operator, data-fidelity loss, patch-based optimisation and evaluation protocol, and differing only in the supplementary information enlisted against the absolute-level indeterminacy of Eq. (2.21), collected in this chapter.

4.1. Shared Neural Field Representation

In all three models the reconstruction problem is the search for a parametric field

$$q_\theta : \mathbb{R}^2 \to \mathbb{R}^{n_q}, \qquad q_\theta(x, y) = \mathcal{N}_\theta(\tilde{x}, \tilde{y}), \tag{4.1}$$

minimising a composite objective that combines a unified BOS data loss with a model-specific selection of physical and supplementary constraints. Here θ collects the trainable parameters of a fully connected multilayer perceptron (MLP) $\mathcal{N}_\theta$, and $(\tilde{x}, \tilde{y})$ denote the normalised spatial coordinates introduced below. The approximator q_θ specialises the generic PINN field $\hat{u}_\theta$ to the present reconstruction problem.

Density-Centred Network Architecture

The density field ρ is the common reconstructed variable of the three models: the only quantity to which the BOS measurement is directly sensitive (Eq. (2.20)), whose absolute level the null-space property affects, and on which each model's supplementary information is tested. The additional variables predicted alongside ρ in some models, and the positivity parameterisations, are model-specific and introduced in the corresponding sections.

The shared architecture is summarised as follows: a fully connected MLP with $L = 8$ hidden layers of $W = 64$ units and the smooth swish activation $\mathrm{swish}(z) = z\,\sigma(z)$. A smooth activation is required because the loss differentiates residuals built on q_θ: it gives the continuous, non-trivial first derivatives that ReLU's piecewise-constant ones lack, and keeps well defined any higher-order derivative entering a model-specific residual, such as a Laplacian term. The output layer is linear; only its width n_q and the transformation of its components are model-specific, and are reported, with bias initialisations and loss-weight schedules, in the corresponding sections.

Table 4.1: Architecture of the shared MLP $\mathcal{N}_\theta$; the output dimension n_q is model-dependent and specified in the respective sections.

Hyperparameter	Value
Input dimension	2
Hidden layers L	8
Hidden width W	64
Activation	$\mathrm{swish}(z) = z\,\sigma(z)$
Output layer	linear

Spatial Input Normalisation

PINN optimisation is sensitive to the scale of its inputs and loss residuals [44]. The network input is therefore normalised uniformly across the three models: the physical coordinates (x, y), referred to an origin at the geometric centre of the computational domain of Chapter 3, are mapped to the dimensionless coordinates

$$\tilde{x} = \frac{x}{L_{\mathrm{ref}}}, \qquad \tilde{y} = \frac{y}{L_{\mathrm{ref}}}, \tag{4.2}$$

where L_{ref} is a reference length defined as half the largest extent of the domain. With this centred origin, Eq. (4.2) bounds $(\tilde{x}, \tilde{y})$ by unity over the domain (the bound attained along the axis of largest extent), keeping the swish-MLP well conditioned at initialisation and during training. Derivatives in physical units are obtained through the chain rule with the inverse scale $1/L_{\mathrm{ref}}$, so the discrete UBOS operator below and any model-specific residual remain dimensionally consistent. Output-side scaling, by contrast, is model-specific: which variables are predicted in dimensionless or physical form, and their reference scales, are specified in the corresponding section.

4.2. Unified BOS Data Loss

Discrete UBOS Operator and BOS Data Loss

The data-fidelity term tying the predicted density to the optical measurement, shared across the three models, is the trainable counterpart of the unified BOS consistency relation $A\boldsymbol{\rho} = \mathbf{b}$ of Eq. (2.20). On the structured grid assembled in Chapter 3, with grid spacings Δx and Δy, the discrete UBOS operator acts on the predicted density field as

$$\mathcal{A}[\rho_\theta] = I_x\, C_{\text{sys}}\, D_x\, \rho_\theta + I_y\, C_{\text{sys}}\, D_y\, \rho_\theta, \tag{4.3}$$

where ρ_θ is the density predicted by q_θ on the relevant grid (or on a patch thereof), I_x and I_y are the horizontal and vertical gradients of the reference background image already validated in Chapter 3, C_{sys} is the optical sensitivity coefficient introduced in Chapter 2, and D_x, D_y are discrete partial-derivative operators applied to ρ_θ. The latter are implemented as fourth-order, five-point centred first-derivative stencils,

$$(D_\alpha \rho_\theta)_i = \frac{-\rho_{\theta,i+2} + 8\rho_{\theta,i+1} - 8\rho_{\theta,i-1} + \rho_{\theta,i-2}}{12\,\Delta\alpha}, \qquad \alpha \in \{x, y\}, \tag{4.4}$$

where i runs along the α-direction and the other index is held fixed. The stencils are realised as one-dimensional convolutions with reflective padding, identical to those used in Chapter 3 for the analytical density gradients and for the discrete-operator/optical-flow consistency check. Sharing them across data generation, validation and training makes the discretisation error of $\mathcal{A}$ at training time coincide with the one already characterised on the synthetic baseline.

Written compactly, the operator returns the directional derivative $\mathcal{A}[\rho_\theta] = C_{\text{sys}}\, \nabla I \cdot \nabla \rho_\theta$, so that at each pixel only the component of $\nabla \rho_\theta$ aligned with the background-image gradient ∇I is observed, while the orthogonal component is not constrained pointwise. A spatially rich, near-isotropic background gradient (such as the speckle pattern of Chapter 3) varies the local direction of ∇I across the field, so that both components of $\nabla \rho_\theta$ become recoverable at field scale.

The common BOS data loss is the mask-normalised mean square mismatch between the discrete operator applied to the predicted density and the training signal $\mathbf{b}$ delivered with the synthetic data package,

$$\mathcal{L}_{\text{meas}}(\theta) = \frac{\sum_\Omega m\, \big(\mathcal{A}[\rho_\theta] - \mathbf{b}\big)^2}{\sum_\Omega m \,+\, \varepsilon}, \tag{4.5}$$

where Ω is the spatial evaluation domain, $m \in \{0,1\}^{H\times W}$ the validity mask provided by the training package and ε a small constant preventing division by zero where the mask weight vanishes. Throughout, $b(x,y)$ denotes the scalar unified-BOS signal field and $\mathbf{b}$ its vectorisation over the grid points of Ω. The mask normalisation makes $\mathcal{L}_{\text{meas}}$ comparable across patches of unequal mask coverage, essential in the patch-based training below.

Density Null-Space and Role of Supplemental Information

The trainable counterpart of the unified BOS operator inherits, by the same constant-annihilation and linearity argument already given for Eq. (2.21), the identical gauge property at the loss level:

$$\mathcal{L}_{\text{meas}}(\theta; \rho_\theta + c) \;=\; \mathcal{L}_{\text{meas}}(\theta; \rho_\theta) \quad \text{for any constant } c. \tag{4.6}$$

The data loss alone therefore determines the gradient structure of ρ but not its absolute level, and any reconstruction strategy must enlist supplementary information to break this gauge. The specific channels adopted by the three models, and the corresponding loss components, are introduced in the dedicated chapters.

4.3. Patch-Based Optimisation Strategy

The composite objective of each model aggregates terms that act on different mathematical objects: pointwise residuals evaluated through automatic differentiation of q_θ, grid-aware terms requiring the discrete UBOS operator of Eq. (4.3), and sparse or domain-restricted constraints that are model-specific. Reconciling these heterogeneous evaluation modes within a single training step motivates the spatially structured mini-batch strategy adopted uniformly by the three models.

A square spatial patch is the natural mini-batch unit for two reasons: the convolutional discrete derivatives of Eq. (4.4) require a contiguous two-dimensional region with halo cells, which random collocation scatters cannot provide; and the automatic-differentiation graph for higher-order derivatives of q_θ scales unfavourably with the number of points, making full-grid evaluation prohibitive at the resolutions of Chapter 3. A square patch of side P thus satisfies both requirements while supporting a vectorised $\mathcal{A}$. All three models use $P = 128$, $B = 4$ patches and $S = 256$ steps per epoch.

The three models also share the patch-selection strategy. Since the BOS signal is gradient-based, the energy of $\mathbf{b}$ concentrates where ρ varies non-trivially, while large portions of the domain carry $|b| \approx 0$; a uniform sampler would dilute the optimisation budget on uninformative patches. To avoid this, all three models adopt a *signal-weighted importance sampler*, with a pixel-level score defined from the training signal,

$$\text{score}(x, y) = |b(x, y)|\, m(x, y), \tag{4.7}$$

non-negative and vanishing on invalid pixels. To stabilise the dynamic range of the subsequent operations, the score is rescaled by its mean over valid pixels,

$$\widetilde{\text{score}}(x, y) = \frac{\text{score}(x, y)}{\bar{s} + \varepsilon}, \qquad \bar{s} = \frac{\sum_\Omega \text{score}}{\sum_\Omega m + \varepsilon}, \tag{4.8}$$

a normalisation that is multiplicative and therefore does not affect the resulting categorical distribution but prevents numerical degeneracies in the sampling step. The score is aggregated over candidate patches by a sliding-window summation,

$$s_{ij} = \sum_{(x,y) \in \mathcal{P}_{ij}} \widetilde{\text{score}}(x, y), \tag{4.9}$$

where $\mathcal{P}_{ij}$ is the patch of side P anchored at the upper-left pixel position (i, j), with $i \in \{0, \dots, H - P\}$ and $j \in \{0, \dots, W - P\}$. The sliding window is implemented as a two-dimensional cross-correlation with a unit-valued $P \times P$ kernel and valid padding, and produces a discrete map $\{s_{ij}\}$ of size $(H{-}P{+}1) \times (W{-}P{+}1)$. From the patch scores, a categorical sampling distribution is constructed by taking their logarithm and using it as the logit vector of a categorical sampler,

$$\ell_{ij} = \log\bigl(s_{ij} + \varepsilon\bigr), \qquad \Pr[(i, j)] = \frac{e^{\ell_{ij}}}{\sum_{k\ell} e^{\ell_{k\ell}}} \;\propto\; s_{ij}, \tag{4.10}$$

which is numerically stable and exactly equivalent to drawing patches with probabilities proportional to s_{ij}. At each training step, B patches are drawn *independently and with replacement* from this categorical distribution and used as the spatial mini-batch for the step; overlap of the B patches is therefore admissible and indeed expected when the bulk of the BOS energy is concentrated in a small region of the domain. The implementation is fully tensorised and the cost of importance sampling is dominated by the precomputation of $\{s_{ij}\}$, which is performed only once at the beginning of the training run; the per-step cost reduces to that of a single draw from a fixed categorical distribution. A side-by-side comparison of the resulting spatial coverage against uniform sampling is reported in Fig. 4.1.

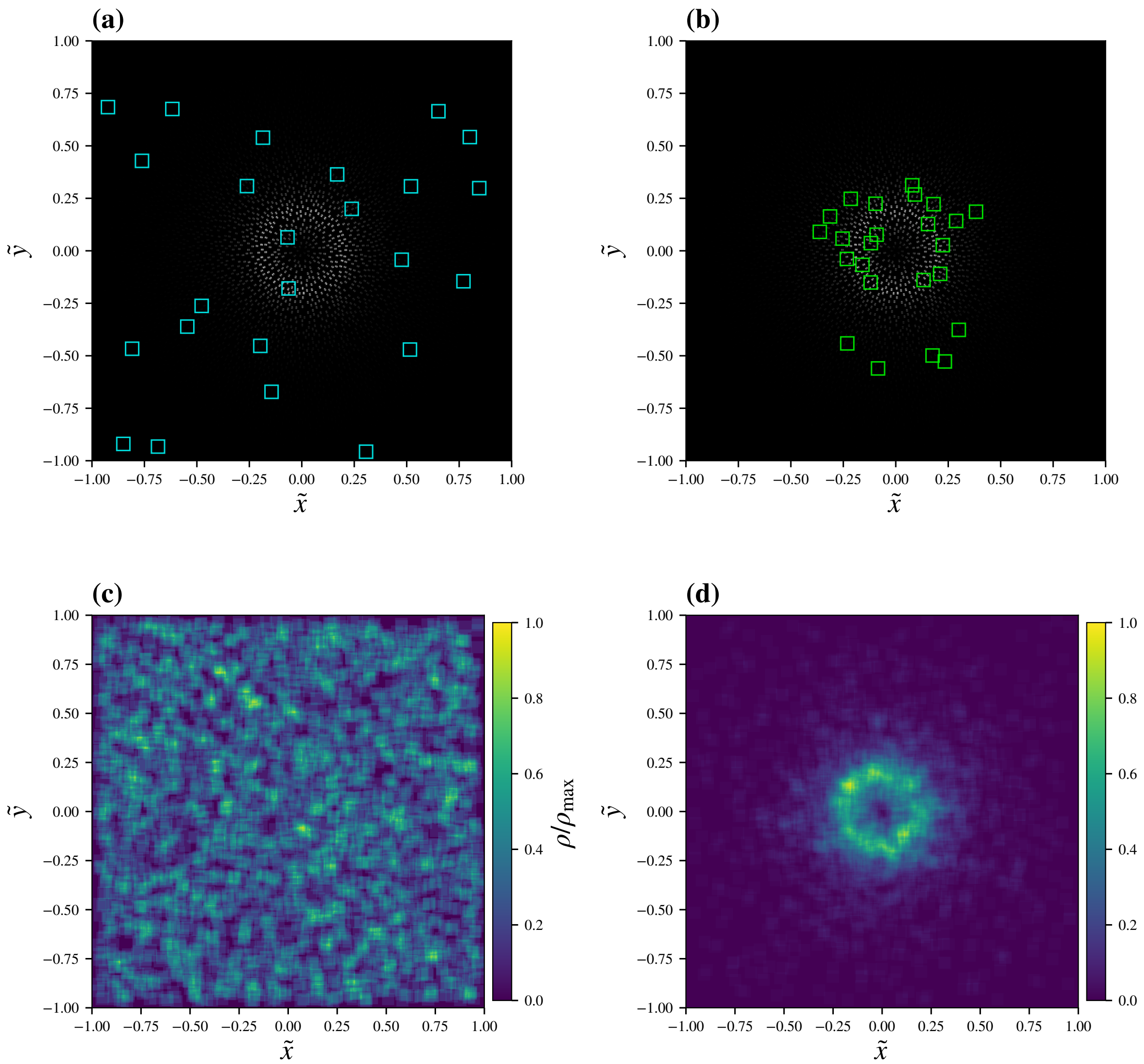


Figure 4.1: Comparison between uniform random patch sampling (**left column**) and the BOS-weighted importance sampling strategy adopted in the three reconstruction models (**right column**). **(a, b)** Single-realisation overlays of 24 patches drawn under the two strategies on top of the masked BOS signal $|b| \cdot m$. Uniform sampling produces patches scattered across the entire domain irrespective of the local signal energy, whereas the BOS-weighted sampler clusters its draws on the regions where $|b|$ is non-negligible. **(c, d)** Aggregate patch coverage density obtained by accumulating the indicator function of $N = 4096$ independent patch draws and normalising by the maximum count. Under uniform sampling the coverage is essentially flat away from the boundary, while under BOS-weighted sampling the coverage is concentrated on the high-$|b|$ structures, exposing the optimiser preferentially to the regions in which the gradient of ρ is constrained by the data.

Empirical Comparison: Uniform vs. BOS-Weighted Sampling

To quantify the contribution of the importance-sampling strategy, a dedicated cross-run campaign was conducted on Model 1 (Chapter 5), training the shared framework under two configurations differing only in the sampler: a uniform random patch baseline and the BOS-weighted sampler of Eqs. (4.7)–(4.10). With all other hyperparameters, optimisation settings and loss components held fixed, any difference in reconstruction quality is attributable to the sampling strategy alone. The figures reported below correspond to a single training realisation per configuration and to the best epoch selected by the validation score of Section 4.4; the percentages should therefore be read as indicative of the effect size rather than as statistically resolved estimates.

Fig. 4.2 compares the relative L_2 error of the two configurations across the optimisation history, for the density and for the three gradient channels. At the best epoch the BOS-weighted sampler reduces the error on the gradient channels by more than an order of magnitude with respect to the uniform baseline[1]: the improvement is about 95% on $\partial_x\rho$ (from 2.3×10^{-1} to 1.2×10^{-2}), about 94% on $\partial_y\rho$ (from 2.3×10^{-1} to 1.5×10^{-2}) and about 95% on $|\nabla\rho|$ (from 2.3×10^{-1} to 1.3×10^{-2}), and it also lowers the relative L_2 error on the absolute density level by about 76% (from 7.3×10^{-4} to 1.7×10^{-4}).

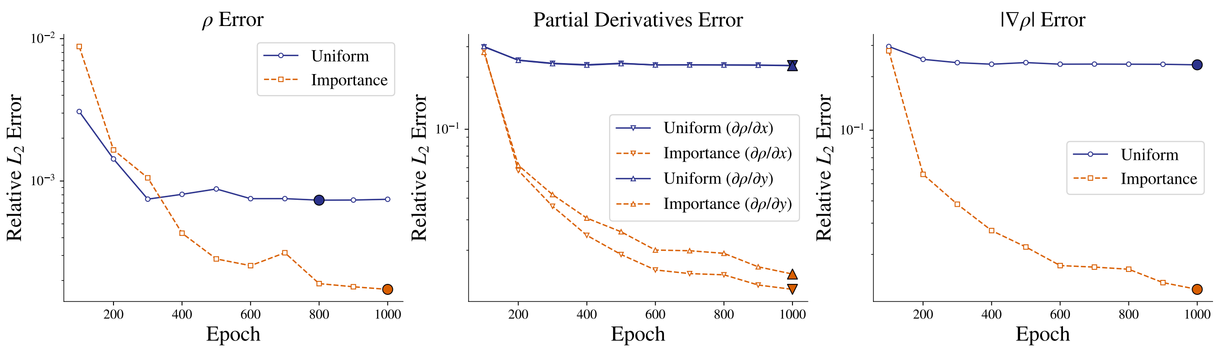


Figure 4.2: Relative L_2 error of Model 1 across the optimisation history for the density ρ and the three gradient channels $\partial_x\rho$, $\partial_y\rho$ and $|\nabla\rho|$, under uniform random patch sampling and under the BOS-weighted importance-sampling strategy of Eqs. (4.7)–(4.10).

The most pronounced effect of the sampler is spatial, exposed by the radial diagnostic of Fig. 4.3. In the Rankine core, where the BOS signal is strongest, the BOS-weighted sampler reduces the mean absolute error on $\partial_r\rho$ and $|\nabla\rho|$ by about 98% and on ρ by about 86%. This near-collapse is the geometric counterpart of the coverage maps of Fig. 4.1: concentrating the patch draws on the high-$|b|$ structures exposes the optimiser to the region where the gradient of ρ is most strongly constrained.

[1]The BOS-weighted figures of this benchmark differ from the headline Model 1 run of Chapter 5.

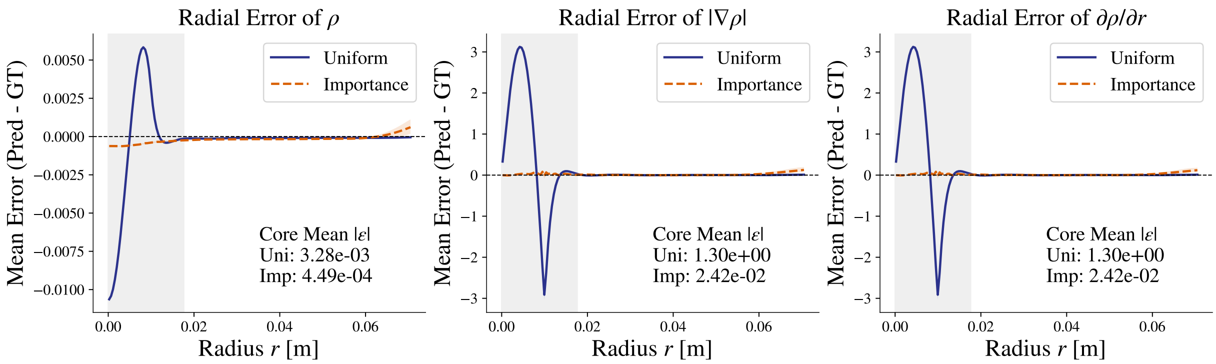


Figure 4.3: Mean absolute error in the central region of the domain ($r \leq r_{\text{core}}$) for the radial density gradient $\partial_r \rho$, the gradient magnitude $|\nabla \rho|$ and the density ρ, under uniform and BOS-weighted sampling.

Optimisation Workflow

The training workflow is shared across the three models. Each step consists of (i) the assembly of a mini-batch of B patches drawn from the categorical distribution of Eq. (4.10), (ii) the forward evaluation of q_θ on the patch coordinates, (iii) the application of the discrete UBOS operator of Eq. (4.3) and the evaluation of $\mathcal{L}_{\text{meas}}$ from Eq. (4.5), (iv) the evaluation of the model-specific loss components, (v) their combination through the model-specific weights, and (vi) back-propagation. All three models are trained with the Adam optimiser [21] with default momentum coefficients, initial learning rate $\alpha_{\text{lr},0}$, exponential decay factor r_{lr} every k_{decay} steps, and global gradient-norm clipping at g_{clip}. Each run comprises N_{ep} epochs of $S = 256$ steps.

Two methodological remarks on the BOS-weighted sampler are in order. First, the scheme of Eqs. (4.7)–(4.10) is informational rather than physical: since $\mathbf{b}$ depends on ρ only through its gradient, preferential sampling on high-$|b|$ regions does not lift the null-space indeterminacy of Eq. (4.6), but improves the conditioning of the gradient-domain component of the inverse problem. Second, no $1/p_{ij}$ inverse-probability correction is applied: the stochastic gradient is therefore an unbiased estimator of an *effective* loss in which each patch contributes with weight proportional to s_{ij}, rather than of the uniformly averaged $\mathcal{L}_{\text{meas}}$ of Eq. (4.5). The two sampling configurations compared in Section 4.3 thus minimise different effective objectives, though both are evaluated against the same uniform-mask metric; whether the deliberate bias improves the reconstruction even under that metric is precisely what the comparison quantifies.

4.4. Evaluation Protocol

The synthetic ground truth of Chapter 3 is the common reference for all three models, and the post-training protocol below is applied uniformly, so that comparisons rest on shared rather than model-specific definitions.

The primary evaluation field is the density ρ: the variable the null-space property of Eq. (2.21) affects and the only one predicted by all three models. Because the BOS measurement is intrinsically gradient-based, however, the evaluation also targets the components of the density gradient, all of which are computed from both the predicted and the ground-truth density field through the same fourth-order stencils of Eq. (4.4).

A common metric set is computed for each evaluated field on the valid mask. Denoting by $\hat{f}$ the prediction, by f the ground truth and by N the number of valid points, the metrics adopted are the relative L_2 error $\|\hat{f} - f\|_2/(\|f\|_2 + \varepsilon)$, the root mean square error, the mean absolute error, the mean absolute percentage error, the linear correlation coefficient $\mathrm{corr}(\hat{f}, f)$, and a battery of percentiles of the absolute error.

A scalar composite score, a weighted combination of the relative L_2 errors on ρ, on its gradient channels and on the BOS forward field $b_{\mathrm{pred}} = \mathcal{A}[\rho_\theta]$, is computed at every checkpoint and provides a uniform criterion for selecting a representative "best" epoch. Since it is built from relative L_2 errors against the synthetic ground truth, the resulting checkpoint selection is a *validation* procedure, not a field-deployable criterion: with no ground truth available, an experimental campaign would instead rely on measurable-only quantities, such as a plateau of the trainable losses ($\mathcal{L}_{\mathrm{meas}}$ and the model-specific supplementary and physical-residual terms). This caveat applies uniformly to the three models, to the uniform baseline and to the sampler comparison of Section 4.3.

One model-specific exception is admitted: for Model 3, whose objective is gauge-degenerate by construction (Chapter 7), the absolute-level component of the score is uninformative, and checkpoint selection is restricted to the gradient-channel terms; the gradient-domain figures of the three models remain selected under identical criteria and are therefore directly comparable.

5 | Model 1: A Strongly Anchored Proof of Concept

5.1. Purpose and Scope

In the controlled erosion of supplemental information that orders the three models of Chapter 4, Model 1 occupies the strongly informative end: the gauge of Eq. (4.6) is suppressed as far as the synthetic dataset permits, by exposing the network to ground-truth ρ and in-plane velocity on the four sides of the frame, with no interior supervision. Model 1 is therefore preparatory: it verifies that the shared framework recovers a self-consistent compressible state from a single BOS image pair under maximal gauge suppression, and provides the controlled upper-bound baseline calibrating Models 2 and 3.

5.2. Model Formulation

Model 1 instantiates the shared UBOS/PINN framework of Chapter 4 with the largest output set among the three models. The network is the shared swish-MLP of Table 4.1 ($L = 8$, $W = 64$), parameterised as a map

$$\mathcal{N}_\theta : (\tilde{x}, \tilde{y}) \longmapsto \big(\hat{\rho}^*,\, \hat{u}^*,\, \hat{v}^*,\, \hat{p}^*\big). \tag{5.1}$$

The output components of $\mathcal{N}_\theta$ are normalised primitive variables, recovered in physical units through the reference scales ρ_{ref}, u_{ref} and p_{ref}, associated respectively with the freestream density, a characteristic in-plane velocity and the ambient pressure of the synthetic Rankine vortex of Chapter 3:

$$\rho = \rho_{\text{ref}}\,\hat{\rho}^*, \qquad u = u_{\text{ref}}\,\hat{u}^*, \qquad v = u_{\text{ref}}\,\hat{v}^*, \qquad p = p_{\text{ref}}\,\hat{p}^*. \tag{5.2}$$

The reference scales adopted in Model 1 are $\rho_{\mathrm{ref}} = 1.225\ \mathrm{kg\,m^{-3}}$, $u_{\mathrm{ref}} = 110\ \mathrm{m\,s^{-1}}$, $p_{\mathrm{ref}} = 101325$ Pa and a heat capacity ratio $\gamma = 1.4$. With these scales the reference speed of sound is $c_{\mathrm{ref}} = \sqrt{\gamma p_{\mathrm{ref}}/\rho_{\mathrm{ref}}} \approx 340\ \mathrm{m\,s^{-1}}$, so that the characteristic Mach number is $M_{\mathrm{ref}} = u_{\mathrm{ref}}/c_{\mathrm{ref}} \approx 0.32$: the synthetic vortex is weakly compressible, the relative density variation scaling as $\mathcal{O}(M_{\mathrm{ref}}^2)$, i.e. a few percent: small but spatially structured, hence detectable because BOS probes the density *gradient*, not its level. The compressible Euler closure is thus retained as the consistent governing model despite the absence of shocks at this Mach number. The scale u_{ref} is used purely for non-dimensionalisation: it fixes the order of magnitude of the in-plane velocity and does not impose $|\hat{u}^*| \leq 1$, so local swirl velocities slightly above u_{ref} are admissible. To bias the optimiser toward physically admissible states from the first iteration, the bias of the output linear layer is initialised to the constant vector $[1, 0, 0, 1]^\top$ in normalised space, so that an untrained $\mathcal{N}_\theta$ predicts $\rho = \rho_{\mathrm{ref}}$, $u = v = 0$ and $p = p_{\mathrm{ref}}$, a uniform freestream-like state. A default zero-output initialisation would instead drive $\rho \to 0$, where the equation of state (EOS) $p = K\rho^\gamma$ and the Euler balances become singular through division by density; the freestream initialisation places the optimiser in a benign neighbourhood in which all residuals are smooth and well scaled.

The BOS data-fidelity term is built from the predicted density alone: at each step $\rho_\theta = \rho_{\mathrm{ref}}\,\hat{\rho}^*$ is passed to the discrete UBOS operator of Eq. (4.3) and $\mathcal{A}[\rho_\theta]$ is compared with the signal $\mathbf{b}$ on the same patch. The components $\hat{u}^*, \hat{v}^*, \hat{p}^*$ do not enter $\mathcal{A}[\rho_\theta]$ but are coupled to ρ through the Euler residuals below. The architecture is sketched in Fig. 5.1.

5.3. Active Loss Terms and Boundary Anchoring

The composite objective of Model 1 collects four contributions: the BOS data loss of Chapter 4, a compressible Euler residual loss, an isentropic equation-of-state loss and a soft Dirichlet boundary loss on density and in-plane velocity:

$$\mathcal{L}^{(1)}(\theta) = w_{\mathrm{meas}}\,\mathcal{L}_{\mathrm{meas}}(\theta) + w_{\mathrm{phys}}\,\mathcal{L}_{\mathrm{phys}}(\theta) + w_{\mathrm{EOS}}\,\mathcal{L}_{\mathrm{EOS}}(\theta) + w_{\mathrm{BC}}\,\mathcal{L}_{\mathrm{BC}}(\theta), \tag{5.3}$$

with the model-specific weights $w_{\mathrm{meas}} = 10^{-3}$, $w_{\mathrm{phys}} = 1$, $w_{\mathrm{EOS}} = 1$ and $w_{\mathrm{BC}} = 10$.

The BOS data loss $\mathcal{L}_{\mathrm{meas}}$ is the mask-normalised term of Eq. (4.5), evaluated on patches drawn from the BOS-weighted categorical distribution of Eq. (4.10), with $\mathcal{A}$ applied to ρ_θ as in Eq. (5.2).

The physics-informed loss $\mathcal{L}_{\mathrm{phys}}$ enforces the two-dimensional planar specialisation ($\eta = 0$)

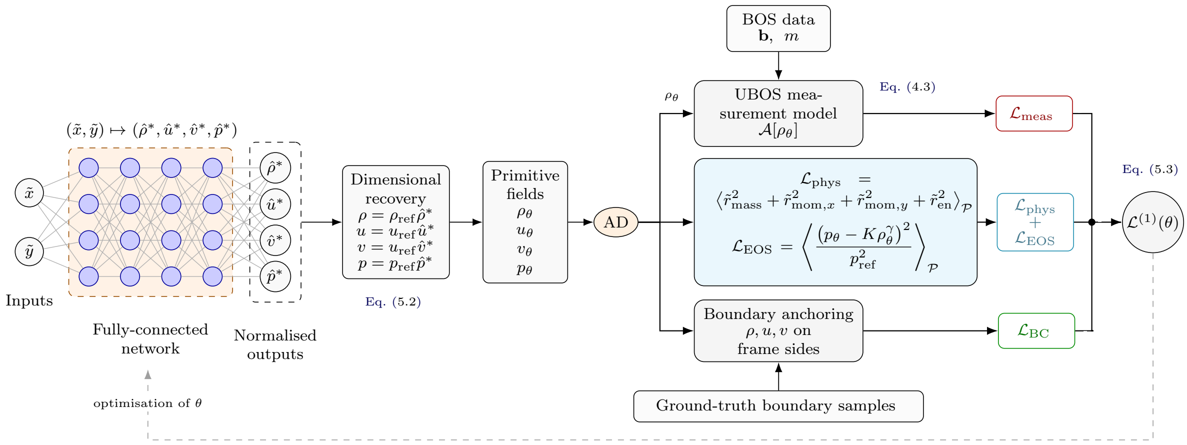


Figure 5.1: Architecture of the Model 1 UBOS/PINN reconstruction. Normalised coordinates $(\tilde{x}, \tilde{y})$ (Eq. (4.2)) feed a swish-MLP ($L = 8$, $W = 64$) whose output $(\hat{\rho}^*, \hat{u}^*, \hat{v}^*, \hat{p}^*)$ is rescaled via Eq. (5.2). The predicted density ρ_θ is passed through the discrete UBOS operator $\mathcal{A}$ (Eq. (4.3)) and compared patch-wise with $\mathbf{b}$ to define $\mathcal{L}_{\mathrm{meas}}$ (Eq. (4.5)). The primitive variables drive the Euler residual $\mathcal{L}_{\mathrm{phys}}$ and isentropic EOS residual $\mathcal{L}_{\mathrm{EOS}}$ via automatic differentiation. Strong anchoring is enforced by the soft Dirichlet loss $\mathcal{L}_{\mathrm{BC}}$ on ρ, u, v at the four boundaries.

of the compressible Euler system of Eqs. (2.22)–(2.25), evaluated on the patch collocation points through automatic differentiation of $\mathcal{N}_\theta$. Denoting by r_{mass}, $r_{\mathrm{mom},x}$, $r_{\mathrm{mom},y}$ and r_{en} the pointwise residuals of the steady mass, x- and y-momentum and energy balances built from the rescaled primitive variables of Eq. (5.2), the physics loss is

$$\mathcal{L}_{\mathrm{phys}}(\theta) = \langle \tilde{r}^2_{\mathrm{mass}} + \tilde{r}^2_{\mathrm{mom},x} + \tilde{r}^2_{\mathrm{mom},y} + \tilde{r}^2_{\mathrm{en}} \rangle_{\mathcal{P}}, \tag{5.4}$$

where $\langle\cdot\rangle_{\mathcal{P}}$ denotes the empirical mean over the patch collocation points and the tildes signal that each residual is divided by its natural dimensional scale, $(\rho_{\mathrm{ref}} u_{\mathrm{ref}})/L_{\mathrm{ref}}$ for the mass balance, $(\rho_{\mathrm{ref}} u^2_{\mathrm{ref}})/L_{\mathrm{ref}}$ for the momentum balances, and the sum $(\rho_{\mathrm{ref}} u^3_{\mathrm{ref}} + p_{\mathrm{ref}} u_{\mathrm{ref}})/L_{\mathrm{ref}}$ for the energy balance. As established in Chapter 2, this non-dimensionalisation brings the four residual contributions, which span several orders of magnitude in physical units, to a comparable scale. The gradients in Eq. (5.4) are obtained from the network outputs through a persistent automatic-differentiation tape, with the chain-rule factor $1/L_{\mathrm{ref}}$ applied explicitly. The irrotationality residual introduced with the UBOS/PINN framework in Chapter 2 is deliberately excluded from $\mathcal{L}_{\mathrm{phys}}$. The solid-body core of the Rankine vortex is rotational (vorticity $2\omega_0 \neq 0$, Chapter 3), so enforcing $\omega = \partial v/\partial x - \partial u/\partial y = 0$ would be inconsistent with the synthetic ground truth; the same exclusion is inherited by Models 2 and 3.

The equation-of-state loss $\mathcal{L}_{\mathrm{EOS}}$ enforces the isentropic closure already adopted to construct the synthetic ground truth in Eq. (3.3). Since the Rankine vortex of Chapter 3 is built under the assumption of an isentropic state, with isentropic constant $K = p_{\mathrm{ref}}/\rho_{\mathrm{ref}}^{\gamma}$, the same closure is imposed pointwise on the network output as

$$\mathcal{L}_{\mathrm{EOS}}(\theta) = \left\langle \left(p_\theta - K\,\rho_\theta^{\gamma}\right)^2 / p_{\mathrm{ref}}^2 \right\rangle_{\mathcal{P}}, \tag{5.5}$$

where the division by p_{ref}^2 keeps the residual on the same dimensionless footing as $\mathcal{L}_{\mathrm{phys}}$. This term complements the EOS of Eq. (2.26): collapsing pressure onto a single-parameter function of density removes one degree of freedom from the four-component output and ties $\hat{p}^*$ to $\hat{\rho}^*$ consistently with the synthetic dataset. Since the closure reuses the ground-truth K, a small residual is expected by design, and only its systematic non-zero mean is diagnostic, as taken up in Section 5.5.

The boundary anchoring is the model-specific component that breaks the absolute-level indeterminacy of Eq. (4.6). At each optimisation step a batch of $N_{\mathrm{BC}} = 128$ collocation points is drawn uniformly along the four sides of the rectangular frame, with one quarter on each side, and the network is evaluated on the corresponding normalised coordinates. The boundary loss is the soft Dirichlet term

$$\mathcal{L}_{\mathrm{BC}}(\theta) = \frac{1}{N_{\mathrm{BC}}} \sum_{k=1}^{N_{\mathrm{BC}}} \left[\left(\hat{\rho}_k^* - \rho_{k,\mathrm{GT}}^*\right)^2 + \left(\hat{u}_k^* - u_{k,\mathrm{GT}}^*\right)^2 + \left(\hat{v}_k^* - v_{k,\mathrm{GT}}^*\right)^2 \right], \tag{5.6}$$

where $(\rho_{k,\mathrm{GT}}^*, u_{k,\mathrm{GT}}^*, v_{k,\mathrm{GT}}^*)$ are the ground-truth values of the normalised density and of the in-plane velocity components, sampled from the optional pixelwise Dirichlet field of Section 3.1 on the four sides of the frame. Two methodological remarks are warranted regarding Eq. (5.6). First, the pressure is deliberately excluded from the boundary anchoring: the EOS loss $\mathcal{L}_{\mathrm{EOS}}$ collapses p_θ onto a known function of ρ_θ, so that anchoring the density automatically anchors the pressure and an explicit Dirichlet term on $\hat{p}^*$ would constitute redundant information. Second, ground-truth values are imposed only on the frame boundary: no ground-truth quantity ever reaches interior collocation points, so the bulk reconstruction is determined entirely by the BOS data, the Euler residuals and the EOS residual, the boundary anchoring acting solely as a pinning condition for the density gauge.

5.4. Results

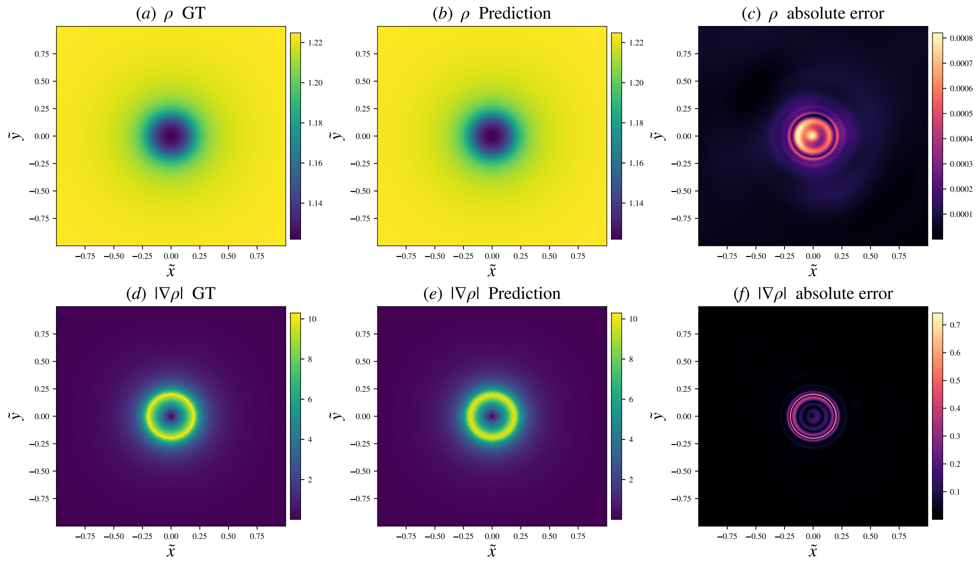
A representative checkpoint was selected through the uniform protocol of Chapter 4. The composite score combines, with comparable weights, the relative L_2 errors on ρ, on the two components of $\nabla\rho$ and its magnitude, and on $b_{\text{pred}} = \mathcal{A}[\rho_\theta]$, so that selection cannot favour the absolute level at the expense of gradient structure.

At the selected checkpoint, Table 5.1 reports a relative L_2 error on ρ of order 9.90×10^{-5}, an RMSE of $1.21\times10^{-4}\,\mathrm{kg\,m^{-3}}$ and a MAE of $7.86\times10^{-5}\,\mathrm{kg\,m^{-3}}$, with correlation essentially unity (1.00) and a global bias $-7.86\times10^{-5}\,\mathrm{kg\,m^{-3}}$. The residual MAPE of 0.0065% confirms that the network reproduces the absolute density level within a few thousandths of a percent. The corresponding spatial maps of ρ_{GT}, ρ_θ and of the absolute and relative errors are collected in Fig. 5.2a, and the density scatter plot of Fig. 5.2b shows that prediction and ground truth align along the identity line without visible saturation.

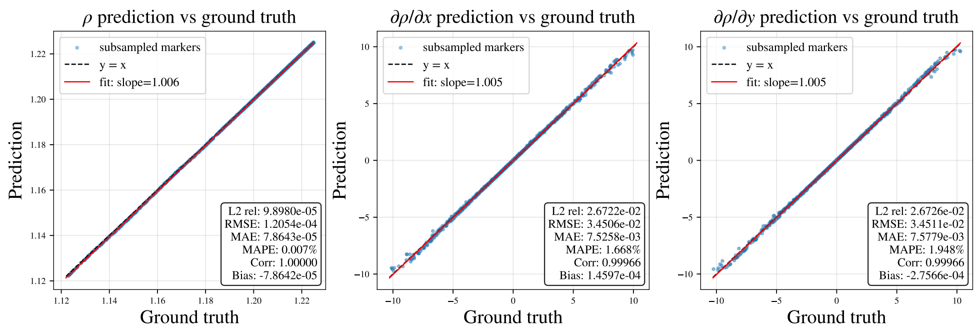
Table 5.1: Reconstruction metrics for Model 1 at the composite-score-selected checkpoint: errors on density, density gradients and BOS forward field against the synthetic reference. The MAPE is reported only for sign-definite fields.

Field	L_2 rel	RMSE	MAE	MAPE [%]	corr	bias
ρ	9.90×10^{-5}	1.21×10^{-4}	7.86×10^{-5}	0.0065	1.00	-7.86×10^{-5}
$\partial_x\rho$	0.0267	0.0345	0.0075	—	1.00	1.46×10^{-4}
$\partial_y\rho$	0.0267	0.0345	0.0076	—	1.00	-2.76×10^{-4}
$\lvert\nabla\rho\rvert$	0.0266	0.0486	0.0114	—	1.00	0.0044
b_{pred}	0.0527	7.07	1.41	—	0.999	-0.581

The accuracy on the components of $\nabla\rho$ is lower than on ρ but remains high. Both $\partial_x\rho$ and $\partial_y\rho$ show a relative L_2 error of about 0.0267, an MAE near $0.0075\,\mathrm{kg\,m^{-4}}$ and a correlation 1.00; the magnitude $|\nabla\rho|$ inherits comparable figures (relative L_2 error 0.0266, RMSE 0.0486, correlation 1.00; Table 5.1). The largest gradient errors, visible in Fig. 5.2a, are concentrated in narrow sub-pixel bands close to the discontinuity in $\partial_r\rho$ at the Rankine core radius introduced in Chapter 3; the methodological significance of this gradient-versus-level differential is discussed in Section 5.5. Re-applying the discrete UBOS operator of Eq. (4.3) to ρ_θ yields a predicted BOS field b_{pred} with relative L_2 error 0.0527 against the training signal $\mathbf{b}$ of Section 3.1, correlation 0.999 and a small negative bias -0.581 in the units of $\mathbf{b}$. Since $\mathcal{L}_{\text{meas}}$ already minimises $\|\mathcal{A}[\rho_\theta]-\mathbf{b}\|$, this low error largely restates that term's convergence rather than providing independent validation; its informative content is the residual *structure*. In Fig. 5.3a the observed $\mathbf{b}$ and the predicted $b_{\text{pred}} = \mathcal{A}[\rho_\theta]$ are visually indistinguishable on the common range $\sim\pm3\times10^3$, while the



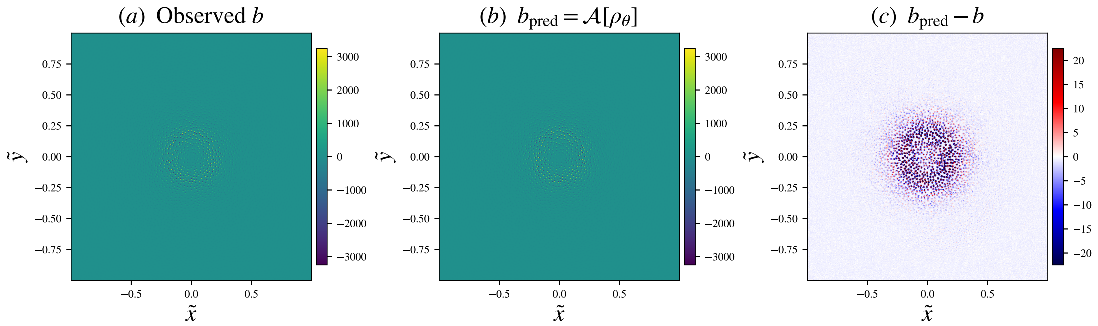
(a) Ground truth, prediction and absolute error for ρ (top) and $|\nabla\rho|$ (bottom), computed via the fourth-order stencils of Eq. (4.4).



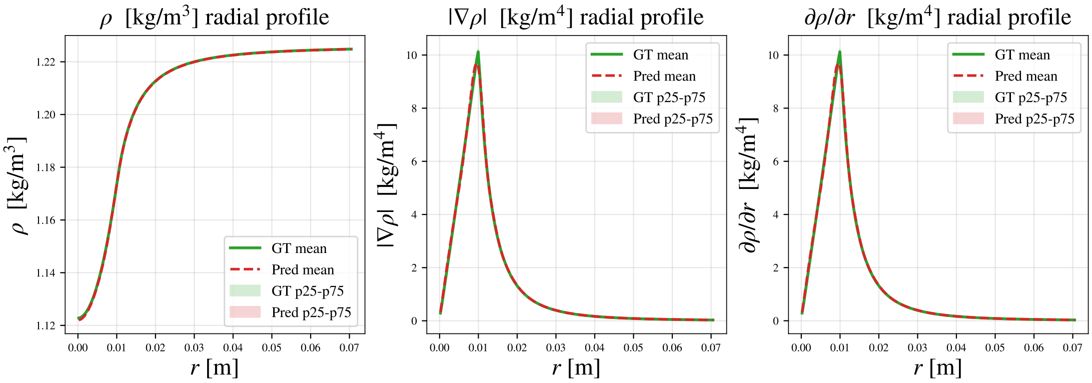
(b) Pixelwise scatter of predicted versus ground-truth values for ρ, $\partial_x\rho$ and $\partial_y\rho$.

Figure 5.2: Complete evaluation for Model 1, showing basic prediction maps (a) and scatter plots (b).

residual map $b_{\text{pred}} - \mathbf{b}$, resolved on a colour scale two orders of magnitude smaller, exposes a structured pattern confined to the gradient-sensitive annulus around the Rankine core with no large-scale offset: the reconstruction thus stays within the consistency region of the unified BOS model. The agreement between prediction and ground truth is further documented along interpretable spatial sections. Fig. 5.3b compares the azimuthally averaged profiles of ρ, $|\nabla\rho|$ and $\partial_r\rho$, against the analytical Rankine reference of Eqs. (3.1)–(3.5). The diagnostics on the soft Dirichlet anchoring of Eq. (5.6) are reported in Fig. 5.4. The MAE on ρ at the four sides of the frame is 2.89×10^{-5} at the top, 5.65×10^{-5} at the bottom, 4.94×10^{-5} at the left and 3.48×10^{-5} at the right, all smaller than the bulk MAE 7.86×10^{-5}, ranging from roughly one-third to three-quarters of the bulk value. This is the quantitative signature expected from the soft Dirichlet term in Eq. (5.6).



(a) BOS forward-consistency check. (a) Observed field $\mathbf{b}$ (Section 3.1). (b) Predicted field $b_{\text{pred}} = \mathcal{A}[\rho_\theta]$ (Eq. (4.3)). (c) Residual $b_{\text{pred}} - \mathbf{b}$, concentrated in the gradient-sensitive annulus around the Rankine core with no large-scale background bias.



(b) Azimuthally averaged profiles of ρ, $|\nabla\rho|$ and $\partial_r\rho$ for prediction and ground truth, with p_{25}–p_{75} bands.

Figure 5.3: Evaluation of Model 1 showing the BOS forward-consistency check (a) and azimuthally averaged radial profiles with error bands (b).

5.5. Interpretation

The metrics of Table 5.1 and the spatial diagnostics of Figures 5.2a–5.4 admit a coherent methodological reading. The discussion is organised around two interpretive points, from the gauge property motivating the model to the limitations motivating the two subsequent chapters.

Suppression of the gauge null-mode and division of labour among loss terms. The defining quantitative outcome of Model 1 is the recovery of the absolute level of ρ within a few thousandths of a percent, with a global bias of $-7.86 \times 10^{-5}\,\mathrm{kg\,m^{-3}}$. By the gauge property of Eq. (4.6), $\mathcal{L}_{\text{meas}}$ does not constrain the constant component of ρ_θ. Among the four terms, the boundary loss $\mathcal{L}_{\text{BC}}$ is the *only* one referencing an absolute ground-truth level: with the largest weight $w_{\text{BC}} = 10$ it exerts a quadratic restoring force

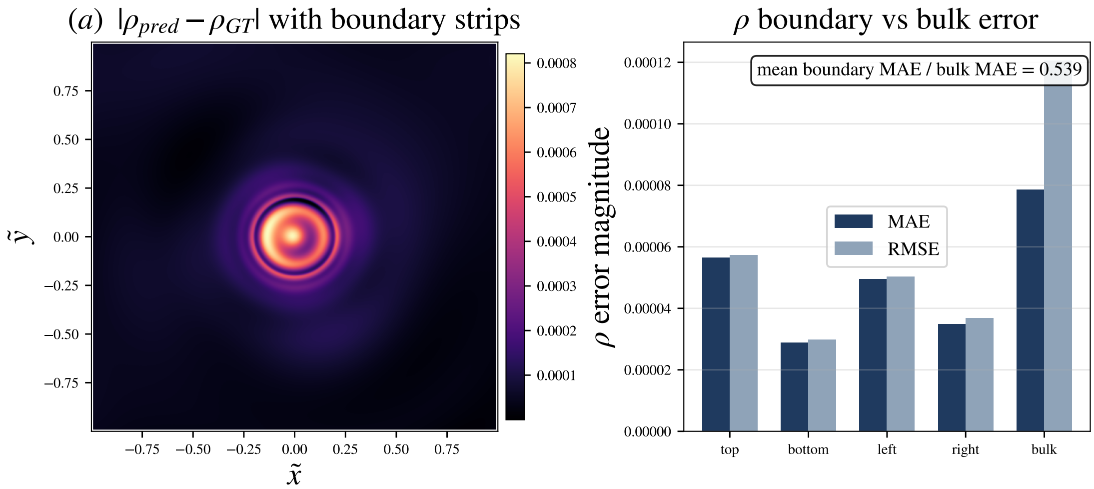


Figure 5.4: Left: absolute density error map with boundary strips highlighted. Right: boundary versus bulk MAE and RMSE for ρ. Boundary errors are consistently below bulk values, in agreement with the soft Dirichlet term of Eq. (5.6).

on that constant along the four sides and pins the gauge. Without boundary anchoring the constant would drift within the gauge-equivalent family; Model 1 is therefore a *strongly anchored proof of concept*, not evidence that BOS alone recovers the absolute density.

Methodological scope and limitations. The favourable picture must be qualified by what Model 1 does *not* prove. The EOS of Eq. (5.5) coincides structurally with the ground-truth closure, making the small EOS residual partly a self-consistency property; and the anchoring of Eq. (5.6) uses ground-truth ρ, u, v on all four sides, stronger side information than any realistic instrumentation provides. Three further qualifications concern the experimental setting and are common to all three models, treated in full in Chapter 8: the synthetic BOS signal $\mathbf{b}$ carries no measurement noise; all results come from a single geometry and a single fixed-seed run, without ensemble statistics; and the reconstructed state is two-dimensional and steady, exact only for a flow invariant along the line of sight, whereas a real BOS measurement is intrinsically three-dimensional.

6 | Model 2: Measurable Thermodynamic Constraints

6.1. Purpose and Scope

Model 2 occupies the intermediate, reduced-supervision position: the ground-truth boundary anchoring of Model 1, which no actual BOS deployment could provide, is replaced by a minimal set of *thermodynamic observables* drawn from the training package of Section 3.1. Motivated by the VKI-inspired instrumentation of Chapter 3, two observables are retained: the dense total-temperature map $T_0(x, y)$ of Eq. (3.6) and the sparse static-pressure taps $\{p_{s,k}\}_{k=1}^{N_{p_s}}$, which anchor the absolute level of p_θ at a few locations. No ground-truth value of ρ, u or v is supplied at any point of the domain. Model 2 asks whether $\mathcal{L}_{\text{meas}}$, $\mathcal{L}_{\text{phys}}$, $\mathcal{L}_{\text{EOS}}$ and the two measurement constraints together lift the indeterminacy of Eq. (4.6), i.e. whether a minimal measurable set built from T_0 and a few p_s can substitute for the dense inlet supervision of Molnar et al. [26]. It also operates without ambient reference scales, predicting ρ, u, v, p directly in physical units; the resulting heterogeneous loss scaling makes weight selection more delicate and places Model 2 between Models 1 and 3.

6.2. Model Formulation

Model 2 instantiates the shared UBOS/PINN framework with the Model 1 architecture, input normalisation (Eq. (4.2)) and patch-based optimisation (Eqs. (4.8)–(4.10)), but with a different output parameterisation, initialisation policy and composite objective. The swish-MLP of Table 4.1 is here parameterised as

$$q_\theta : (\tilde{x}, \tilde{y}) \longmapsto \big(r_\rho,\ r_u,\ r_v,\ r_p\big), \tag{6.1}$$

from dimensionless coordinates to a four-component vector of *raw* pre-activations. The output dimension $n_q = 4$ is preserved because the four primitive variables are jointly required to evaluate the physics residuals of Eqs. (2.22)–(2.25) and the isentropic closure of Eq. (2.26). The physical primitive variables are recovered from the raw output through the component-wise transformations

$$\rho_\theta = \mathrm{softplus}(r_\rho) + \varepsilon, \qquad u_\theta = r_u, \qquad v_\theta = r_v, \qquad p_\theta = \mathrm{softplus}(r_p) + \varepsilon, \tag{6.2}$$

where $\mathrm{softplus}(z) = \log(1 + e^z)$ and ε is a small positive constant. The softplus parameterisation enforces positivity of ρ_θ and p_θ by construction while preserving smooth differentiability under automatic differentiation, preventing the optimiser from drifting toward unphysical states in which the internal energy or the static temperature would lose meaning. The in-plane velocity components are signed kinematic quantities and are left as unconstrained linear outputs. Unlike Model 1, Model 2 applies no global output-side rescaling: the four primitives are predicted directly in physical units. The freestream-like initialisation is reformulated from the measurement channels alone; denoting by $\overline{T_0}$ and $\overline{p_s}$ the means of the total-temperature map and of the pressure taps, the output-layer biases are set so that, at $\theta = \theta_0$,

$$\rho_\theta\big|_{\theta_0} \approx \rho_0 = \frac{\overline{p_s}}{R_{\mathrm{gas}}\,\overline{T_0}}, \qquad u_\theta\big|_{\theta_0} = v_\theta\big|_{\theta_0} = 0, \qquad p_\theta\big|_{\theta_0} \approx \overline{p_s}, \tag{6.3}$$

through $\mathrm{softplus}^{-1}$ of Eq. (6.2). The reference values are drawn from the measurement channels themselves, at the cost of a markedly more heterogeneous optimisation landscape. The discrete UBOS operator is applied to ρ_θ exactly as in the shared framework (Eqs. (4.3)–(4.5)), yielding the mask-normalised $\mathcal{L}_{\mathrm{meas}}$. The components u_θ, v_θ, p_θ enter only through the physical and supplemental constraints, which couple the gauge-invariant data term to the absolute-level information carried by the measurement channels. The physics-informed loss $\mathcal{L}_{\mathrm{phys}}$ is built exactly as in Chapter 5 from the four primitive variables and enforces the steady compressible Euler system of Eqs. (2.22)–(2.25), with the total energy per unit mass defined as

$$E = \frac{p_\theta}{(\gamma - 1)\,\rho_\theta} + \tfrac{1}{2}\big(u_\theta^2 + v_\theta^2\big). \tag{6.4}$$

Spatial derivatives are obtained by automatic differentiation of q_θ, with the chain-rule factor $1/L_{\mathrm{ref}}$ applied explicitly. The aggregated $\mathcal{L}_{\mathrm{phys}}$ is the sum of the four squared-residual contributions, each independently mask-normalised on the patch with the validity mask m of Eq. (4.5).

Total-temperature constraint. The first supplemental constraint ties the predicted thermodynamic state and in-plane kinetic energy to the dense total-temperature map. Since ρ_θ and p_θ are predicted in physical units, the static temperature is recovered pointwise as

$$T_\theta = \frac{p_\theta}{\rho_\theta \, R_{\text{gas}}}, \tag{6.5}$$

and the predicted total-temperature field follows the definition of total (stagnation) temperature for a calorically perfect gas

$$T_{0,\theta} \; = \; T_\theta \; + \; \frac{u_\theta^2 + v_\theta^2}{2\, c_p}, \qquad c_p \; = \; \frac{\gamma \, R_{\text{gas}}}{\gamma - 1}. \tag{6.6}$$

Eq. (6.6) presupposes a calorically perfect gas (constant c_p) and the adiabatic conservation of total enthalpy implied by the steady, inviscid energy balance. The corresponding loss is mask-normalised:

$$\mathcal{L}_{T_0}(\theta) \; = \; \frac{\sum M_{T_0} \left(T_{0,\theta} - T_0\right)^2}{\sum M_{T_0} \; + \; \varepsilon}, \tag{6.7}$$

with $M_{T_0} \in \{0, 1\}$ the availability mask of T_0. Unlike $\mathcal{L}_{p_s}$ and $\mathcal{L}_{\text{EOS}}$, which are additionally gated by the patch validity mask m of Eq. (4.5), $\mathcal{L}_{T_0}$ is gated by M_{T_0} alone. Methodologically, Eq. (6.7) couples three predicted variables to a single measurable thermodynamic invariant through a relation depending on the ratio p_θ/ρ_θ: a uniform shift of ρ_θ admitted by Eq. (4.6) drives T_θ away from the ideal-gas-consistent value and incurs a quadratic penalty, contributing to gauge breaking in a thermodynamically motivated, measurement-realistic way.

Static-pressure tap constraint. The second supplemental constraint exploits the sparse pressure measurements supplied at N_{p_s} locations $\{(x_k, y_k)\}$. The information is encoded in a sparse map $p_s(x, y)$ and a binary mask M_{p_s} flagging the same locations. The tap loss is

$$\mathcal{L}_{p_s}(\theta) \; = \; \frac{\sum M_{p_s} \, m \left(\dfrac{p_\theta - p_s}{p_{\text{ref}}}\right)^2}{\sum M_{p_s} \, m \; + \; \varepsilon}, \tag{6.8}$$

with p_{ref} set to the mean of the available taps. Unlike $\mathcal{L}_{T_0}$, which constrains a thermodynamic ratio, $\mathcal{L}_{p_s}$ is the only loss component that pins, pointwise and unambiguously, the absolute scale of a single predicted variable (p_θ) to a directly measured value. Its action is, however, intrinsically sparse, and the propagation of the absolute level to the rest of the domain relies on the EOS closure introduced next.

Isentropic EOS closure. The synthetic ground truth of Chapter 3 is built with a single isentropic constant $K = p/\rho^\gamma$, and the same closure is imposed pointwise. The constant is in principle recoverable from the measurement channels alone: combining the mean tap pressure $\overline{p_s}$ with the mean total temperature $\overline{T_0}$ gives the reference density $\bar{\rho} = \overline{p_s}/(R_{\text{gas}}\,\overline{T_0})$ already used for the bias initialisation (Eq. (6.3)), and hence the estimate $K = \overline{p_s}/\bar{\rho}^\gamma$. The closure is enforced through

$$\mathcal{L}_{\text{EOS}}(\theta) \;=\; \frac{\sum m \left(\dfrac{p_\theta - K\,\rho_\theta^\gamma}{p_{\text{ref}}}\right)^2}{\sum m \;+\; \varepsilon}, \tag{6.9}$$

with the same p_{ref} as in Eq. (6.8). As in Model 1 (Eq. (5.5)), this closure removes one degree of freedom from the four-component output; its role specific to Model 2 is to propagate the absolute-pressure level fixed by $\mathcal{L}_{p_s}$ at the tap locations to ρ_θ across the entire valid domain: a pointwise constraint on p_θ, combined with a global functional relation between p_θ and ρ_θ, becomes a global constraint on ρ_θ. Thus $\mathcal{L}_{p_s}$ and $\mathcal{L}_{\text{EOS}}$ are not redundant but cooperative. Under the exact closure, the static temperature reduces to $T_\theta = K\,\rho_\theta^{\gamma-1}/R_{\text{gas}}$, a monotone function of ρ_θ; both $\mathcal{L}_{p_s}$ (through the EOS coupling) and $\mathcal{L}_{T_0}$ (through T_θ) therefore provide estimates of the absolute level of ρ_θ that are partially redundant but carry independent measurement noise.

Composite objective. The five contributions assemble into

$$\mathcal{L}^{(2)}(\theta) \;=\; w_{\text{meas}}\,\mathcal{L}_{\text{meas}} \;+\; w_{\text{phys}}\,\mathcal{L}_{\text{phys}} \;+\; w_{T_0}\,\mathcal{L}_{T_0} \;+\; w_{p_s}\,\mathcal{L}_{p_s} \;+\; w_{\text{EOS}}\,\mathcal{L}_{\text{EOS}}, \tag{6.10}$$

with five non-negative weights controlling the relative contribution to the optimiser gradient; their choice is non-trivial because of the heterogeneous scaling of the loss components. A fixed-weight policy analogous to Eq. (5.3) defines the baseline, while the adaptive strategy of Section 6.3 addresses the imbalance that varies during training. The static-weight baseline uses the fixed assignment $(w_{\text{meas}}, w_{\text{phys}}, w_{T_0}, w_{p_s}, w_{\text{EOS}}) = (100,\, 10,\, 3,\, 300,\, 1)$. The weights are not directly comparable to those of Model 1: since Model 2 predicts primitive variables in physical units rather than dimensionless ones, the magnitude of $\mathcal{L}_{\text{meas}}$ changes by several orders, and w_{meas} must be rescaled accordingly. The five terms play complementary roles: $\mathcal{L}_{\text{meas}}$ controls the gauge-invariant gradient content of ρ_θ; $\mathcal{L}_{\text{phys}}$ couples the four primitives through the Euler balances; $\mathcal{L}_{T_0}$ ties the thermodynamic ratio and kinetic energy to a measurable invariant; $\mathcal{L}_{p_s}$ pins p_θ at the taps; and $\mathcal{L}_{\text{EOS}}$ propagates that anchoring to ρ_θ globally. Jointly, $\mathcal{L}_{T_0}$, $\mathcal{L}_{p_s}$ and $\mathcal{L}_{\text{EOS}}$ substitute for the dense boundary anchoring of Eq. (5.6), as shown in Fig. 6.1.

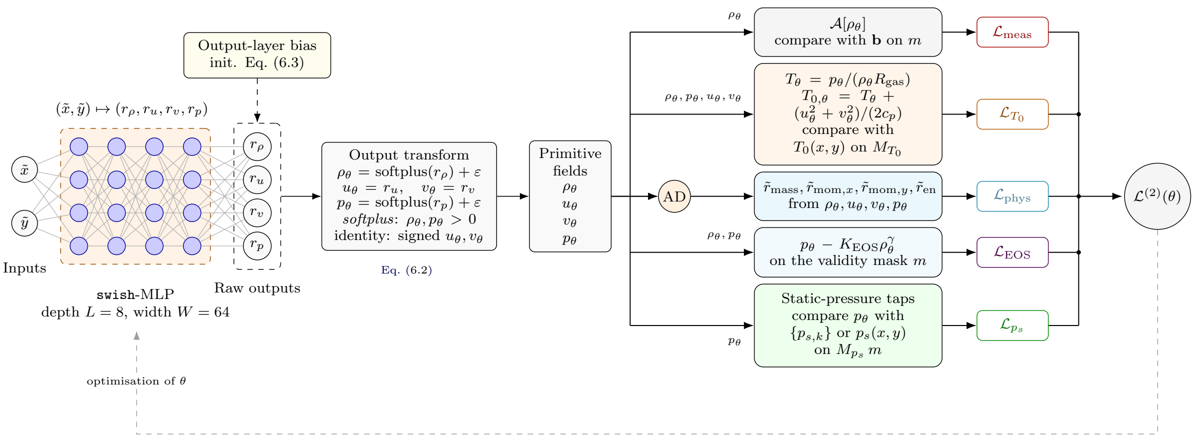


Figure 6.1: UBOS/PINN architecture for Model 2. Normalised coordinates $(\tilde{x}, \tilde{y})$ (Eq. (4.2)) feed a swish-MLP ($L = 8$, $W = 64$); the raw output (r_ρ, r_u, r_v, r_p) is mapped to primitive variables via Eq. (6.2), with softplus enforcing positivity of ρ_θ and p_θ and biases initialised from $\overline{T_0}$, $\overline{p_s}$ (Eq. (6.3)). The density feeds the UBOS operator $\mathcal{A}$ to define $\mathcal{L}_{\text{meas}}$; all four primitives enter $\mathcal{L}_{\text{phys}}$ and $\mathcal{L}_{\text{EOS}}$. The dense $T_0(x, y)$ and sparse $\{p_{s,k}\}$ define $\mathcal{L}_{T_0}$ and $\mathcal{L}_{p_s}$, replacing the boundary anchoring of Model 1.

6.3. Adaptive Weighting Heuristic Extension

The fixed-weight baseline mirrors Model 1, but its rigidity is the main weakness of the present configuration. The five terms of Eq. (6.10) act on heterogeneous scales, and a static weight vector cannot track the regime change between initialisation, when the gradient information is not yet assimilated, and convergence, when the residual imbalance concentrates on the absolute-level channels.

PINN optimisation landscapes are notoriously stiff due to the competing gradients of heterogeneously scaled loss components. Continuous self-adaptive loss weighting strategies based on Neural Tangent Kernel (NTK) analysis [45], gradient statistics [43], or soft-attention mechanisms [24] provide rigorous dynamic balancing and are, importantly, label-free, rebalancing the loss terms from quantities available during unsupervised training. The AWH adopted here is of a different kind: deliberately supervised, using a ground-truth-aware feedback score (Eq. (6.11)), and therefore not a cheaper substitute for those methods but a controlled proxy addressing a different, supervised problem. Its claimed advantages are implementation simplicity and a diagnostic role in exposing the gradient/level trade-off, not any superiority in computational cost. It introduces a staged, run-to-run update rule that lifts the rigidity of the static vector while leaving the loss components, architecture and constraint set of Eq. (6.10) unchanged.

The AWH replaces the static assignment by a staged sequence of $N_{\text{runs}} = 5$ consecutive runs of 1000 epochs each, with θ and Adam optimiser state retained across runs so that the trajectory of θ remains a single continuous descent. The first run uses a fixed weight vector $\boldsymbol{w}^{(0)}$; at the end of each subsequent run, the reconstruction is evaluated on the full domain, the metrics are condensed into an aggregated score, and the weights are updated through a feedback rule before the next run.

The aggregated score is built from the relative L_2 errors e_ρ, e_x, e_y, e_g of ρ_θ, $\partial_x\rho_\theta$, $\partial_y\rho_\theta$, $|\nabla\rho_\theta|$ against the synthetic ground truth, normalised by the respective targets T_ρ for the absolute density level and T_g for the gradient channels:

$$S \;=\; \alpha_\rho \frac{e_\rho}{T_\rho} \;+\; \alpha_x \frac{e_x}{T_g} \;+\; \alpha_y \frac{e_y}{T_g} \;+\; \alpha_g \frac{e_g}{T_g}, \tag{6.11}$$

The signed gap $\Delta S = \operatorname{clip}(S-1, -0.5, 2.0)$ acts as the primary feedback variable. At the end of each run the current weights $w_k^{[0]} = w_k$ are mapped to raw updated weights $\tilde{w}_k$ through a sequence of four multiplicative corrections: (i) a *global correction*; (ii) a *gradient-specific correction*; (iii) an *absolute-density correction*; and (iv) a *stagnation safeguard*. These corrections introduce a sizeable set of hyperparameters beyond the score coefficients $\alpha_\rho, \alpha_x, \alpha_y, \alpha_g$ and the targets T_ρ, T_g; the exact update equations are given in Appendix A.

Finally, to prevent the loss landscape from drifting arbitrarily across runs and to avoid degenerate scale collapses, the raw updated weights $\tilde{w}_k$ are renormalised to the initial weight budget $B = \sum_j w_j^{(0)}$ and then constrained within predefined bounds $[w_k^{\min}, w_k^{\max}]$. Budget conservation therefore holds exactly only while no bound is active; whenever the clipping saturates one or more weights, the post-clip sum departs from B by the clipped excess:

$$w_k^{(n+1)} \;=\; \operatorname{clip}\left(\frac{\tilde{w}_k}{\sum_j \tilde{w}_j}\, B, \quad w_k^{\min},\; w_k^{\max}\right). \tag{6.12}$$

Methodologically, the AWH score (Eq. (6.11)) uses relative L_2 errors against the synthetic ground truth and is therefore *not* computable from the BOS observation and the measurement channels alone: it assimilates a supervisory signal outside the minimal measurable set on which Model 2 is otherwise predicated, complementing rather than supplanting the baseline (Section 6.5).

6.4. Comparative Reconstruction Results

Both training configurations were run end-to-end on the same synthetic baseline. The static-weight configuration uses the fixed assignment of Section 6.2, while the AWH configuration uses the same $\boldsymbol{w}^{(0)}$ with the staged update of Eqs. (6.11)–(6.12). The two configurations, reported side by side in Table 6.1, differ in status: the static-weight baseline respects the measurable-only premise of Model 2, whereas the AWH column is the controlled, supervised proxy of Section 6.3, bounding what a ground-truth-aware reweighting can recover rather than documenting a measurable-only outcome. For both runs, checkpoint selection used the uniform composite validation score $\mathcal{V}$ of Chapter 4 (Section 4.4). Both configurations therefore rely on the synthetic ground truth for this single, downstream model-selection step; this differs in kind from the AWH feedback score S of Eq. (6.11), which assimilates the ground truth *during* training.

Density, gradient and BOS forward field

The reconstruction metrics at the selected checkpoints are reported side by side in Table 6.1. The pixel-wise error structure of the two configurations differs only on the absolute-level component of ρ. Under the static-weight policy the error distribution on ρ is almost entirely supported by a single negative shift of the field with respect to the ground truth: the absolute value of the bias coincides with the MAE and the RMSE within displayed precision, and the p_{95} and maximum absolute errors of 0.0519 and 0.0521 $\mathrm{kg\,m^{-3}}$ differ from the MAE only at the fourth significant digit. Compared with the sub-percent bias of Model 1 ($-7.86 \times 10^{-5}\,\mathrm{kg\,m^{-3}}$), this offset is nearly three orders of magnitude larger, reflecting both the replacement of the dense Dirichlet anchoring of Eq. (5.6) by the sparse, thermodynamically motivated constraints of Eqs. (6.7)–(6.9), and the heterogeneous loss scaling induced by the absence of fixed ambient reference scales. Under the AWH schedule the coherent negative offset is suppressed: L_2 rel and MAPE drop by approximately one order of magnitude ($0.0425 \to 0.0041$ and $4.26\% \to 0.406\%$), and the bias contracts from -0.0518 to $0.0049\,\mathrm{kg\,m^{-3}}$, with sign reversal. The absolute-level component is thus brought to an accuracy class comparable to the gradient channel, closing the gap that Eq. (4.6) had opened in the baseline. On the gradient channels the ranking is reversed but the variation is much smaller. The baseline relative L_2 errors on $\partial_x\rho$, $\partial_y\rho$ and $|\nabla\rho|$ are about 40% lower than those of Model 1 (0.0156 vs. 0.0267 on $\partial_x\rho$); in the absence of run-to-run statistics this difference is not interpreted here as a structural improvement over Model 1. Under AWH they degrade slightly (by approximately one accuracy class), while linear correlations and bulk biases remain within the same range as in the baseline. The largest gradient errors, in both configurations, are confined to the narrow annulus surrounding the core radius. The BOS forward-consistency metrics are essentially unchanged across the two configurations: the relative L_2 error of b_{pred} against $\mathbf{b}$ varies by only a few percentage points, with correlation of order 0.999 in both cases.

Table 6.1: Reconstruction metrics at the selected checkpoints for the static-weight (Base) and AWH training schedules of Model 2, grouped by field against the synthetic reference.

Field	Config	L_2 rel	RMSE	MAE	MAPE [%]	corr	bias
ρ	Base	0.0425	0.0518	0.0518	4.26	1.00	-0.0518
	AWH	0.0041	0.005	0.0049	0.406	1.00	0.0049
$\partial_x \rho$	Base	0.0156	0.0201	0.0098	—	1.00	3.11×10^{-4}
	AWH	0.0254	0.0328	0.0151	—	1.00	-0.0031
$\partial_y \rho$	Base	0.0153	0.0198	0.0095	—	1.00	-9.27×10^{-4}
	AWH	0.0262	0.0339	0.0154	—	1.00	-0.0041
$\|\nabla \rho\|$	Base	0.0151	0.0276	0.0143	—	1.00	4.09×10^{-4}
	AWH	0.0236	0.043	0.0186	—	1.00	-0.0016
b_{pred}	Base	0.0486	6.52	1.61	—	0.999	-0.582
	AWH	0.0527	7.07	1.80	—	0.999	-0.581

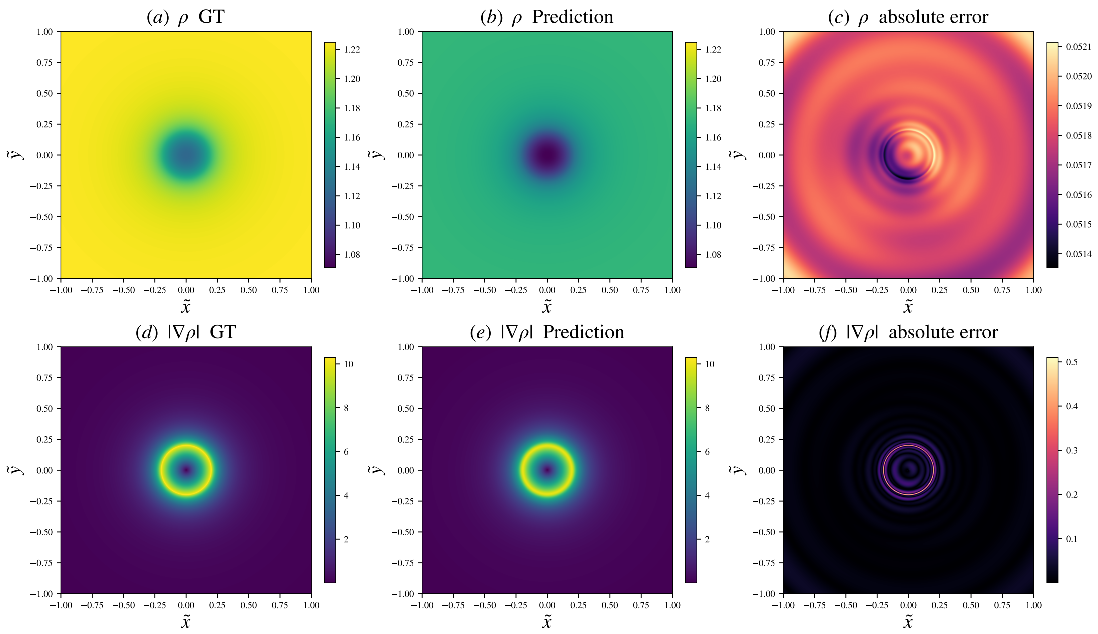
Spatial diagnostics

The spatial structure of these effects is documented in Figures 6.2–6.3, each of which juxtaposes the baseline and AWH outputs in two stacked panels (a) and (b) so that the gauge-aware signature is read directly by comparison.

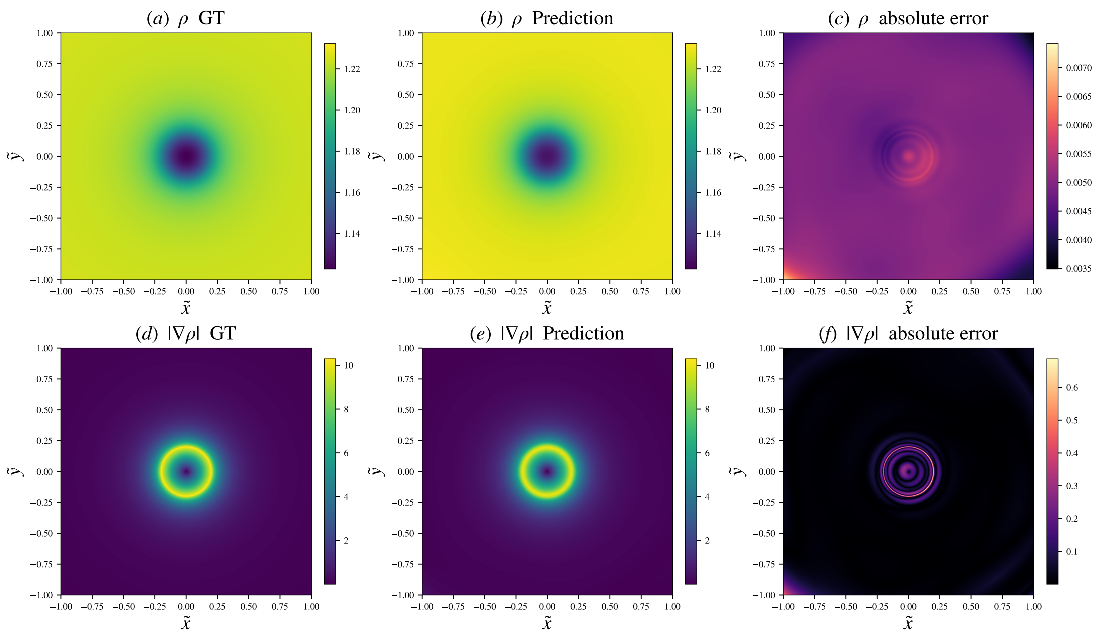
The maps of Fig. 6.2 and the scatter plots of Fig. 6.3 make the gauge-aware character of the trade-off explicit: under the baseline the ρ error map shows a near-uniform negative bulk offset and the ρ scatter is a cloud parallel to but rigidly translated below the identity, whereas under AWH the error map is essentially structureless on a colour scale one order of magnitude tighter and the ρ scatter lies on the identity. The component-gradient scatter clouds retain identity alignment in both configurations.

Physics residuals and measurement-channel constraints

The physical self-consistency of the reconstructed state is documented by the Euler residuals (Eqs. (2.22)–(2.25)) and the isentropic closure (Eq. (3.3)), evaluated by automatic differentiation on the full grid for each configuration. In both cases the four Euler residuals remain comparatively small in the bulk, with the largest excursions, and an analogous

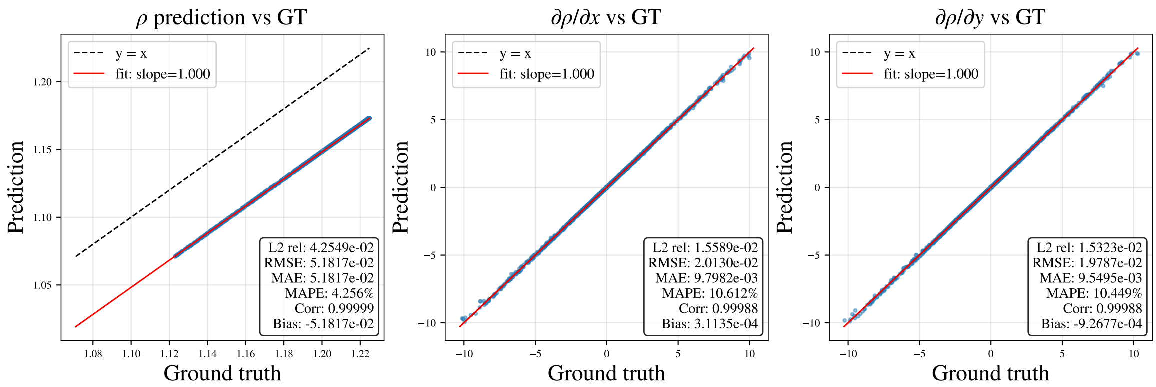


(a) Static-weight baseline.

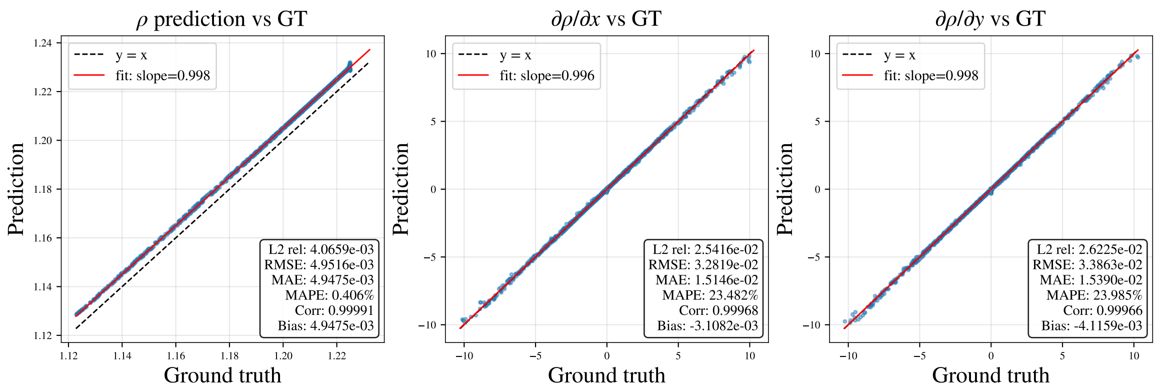


(b) AWH-configured training.

Figure 6.2: Spatial maps of ground truth, network prediction and absolute error for ρ (top row of each panel) and $|\nabla\rho|$ (bottom row of each panel), evaluated through the fourth-order stencils of Eq. (4.4) on prediction and ground truth. (a) Static-weight baseline: the absolute-error map on ρ is dominated by a near-uniform negative bulk offset. (b) AWH-configured training: the ρ error map is essentially structureless on a colour scale approximately one order of magnitude tighter, while the $|\nabla\rho|$ error retains its baseline localisation on the annulus surrounding the Rankine core radius.



(a) Static-weight baseline.



(b) AWH-configured training.

Figure 6.3: Pixelwise scatter of prediction vs. ground truth for ρ and the in-plane components of $\nabla\rho$. (a) Baseline: the ρ cloud is parallel to but rigidly translated below the identity; gradient clouds align tightly along it. (b) AWH: the ρ cloud lies on the identity; gradient clouds retain alignment with marginally broader dispersion (cf. Table 6.1).

EOS pattern, concentrated in the annulus around the Rankine core. Two asymmetric trends emerge. The pressure-tap constraint is satisfied *in the mean* to numerical precision (bias ≈ -1 Pa, five orders below the ambient level), while a pointwise scatter of order 3×10^3 Pa ($\sim 3\%$) persists at the tap locations.

The total-temperature residual, by contrast, behaves in the opposite direction. Already under the baseline the predicted $T_{0,\theta}$ of Eq. (6.6) retains a small but coherent negative bias, indicating that the network slightly underestimates T_0 there; under AWH this bias deepens, MAE and RMSE grow, and the relative L_2 error increases. Quantitatively, the total-temperature MAE rises from 6.12 K under the baseline to 18.3 K under AWH, and the bias deepens from -4.87 K to -18.0 K (L_2 rel $0.0233 \rightarrow 0.0643$), whereas the pressure-tap residual is essentially unchanged (bias $-1.16 \rightarrow -1.03$ Pa).

This deepening of the T_0 residual under AWH is reported openly and admits a double reading. On one side it shows that the schedule channels the absolute-level information primarily through the pressure-anchoring pathway, leaving the structurally weaker T_0 channel to relax. On the other it is a genuine warning: T_0 belongs to the measurable set on which Model 2 is predicated, and the AWH solution, while closest to the ground-truth density, is the *least* consistent with this measurement channel (18.3 K, i.e. $\sim$6% of the ambient level). In an experimental campaign, where the ground truth is unavailable and the measurement channels are the only arbiters, such a configuration could not be recognised as superior; a measurable-only weighting strategy would in fact penalise it. Part of the residual, finally, is structural rather than weighting-induced: with the velocity field unresolved, the dynamic-temperature contribution is missing from $T_{0,\theta}$ under both schedules.

6.5. Interpretation

The metrics of Table 6.1 and the spatial diagnostics of Figures 6.2–6.3 support a single coherent reading, developed below along four lines.

Sufficiency of the minimal measurable constraint set. The combined evidence of the baseline and AWH reconstructions shows that the information content of the set $\{T_0, \{p_{s,k}\}, \mathrm{EOS}\}$ is, on the present synthetic baseline, in principle sufficient to lift the indeterminacy of Eq. (4.6) to an accuracy class comparable to that achieved on the gradient channels by the data-fidelity loss alone. This sufficiency is conditional: it holds only when the constraint weights evolve through the ground-truth-aware AWH proxy. A fully unsupervised, measurable-only pipeline that breaks the gauge to the same accuracy therefore remains to be demonstrated.

Gauge-aware character of the gradient/level trade-off. Two distinct facts should be separated here. The gauge of Eq. (4.6) is *verified* by the near-invariance of the BOS forward residual under the change in ρ between the two schedules: a near-uniform shift of ρ_θ leaves b_{pred} essentially unchanged, which is the direct empirical signature of the null mode. Separately, the AWH *exploits* the channels orthogonal to this gauge to remove the offset: the order-of-magnitude reduction of the error on ρ, jointly with the bounded degradation on the gradient channels, is a clear quantitative illustration of how the trade-off of Eq. (4.6) can be broken without disturbing the forward-consistent content.

Anisotropy of the gauge-breaking pathway. The absolute level is carried mainly by the pressure-anchoring pathway, while the T_0 channel plays a structurally weaker role; the asymmetric response of the two measurement-channel residuals makes this anisotropy explicit. The AWH schedule reinforces the chain $\mathcal{L}_{p_s} \to \mathcal{L}_{\mathrm{EOS}} \to \rho_\theta$, whose combined effect of pointwise pressure anchoring and global functional coupling propagates the absolute pressure level to ρ_θ across the domain; and simultaneously relaxes the orthogonal $\mathcal{L}_{T_0}$ channel, whose residual grows. This reading is consistent with the weight dynamics of the staged update: the absolute-density correction of Section 6.3 places its strongest emphasis on w_{EOS}, while w_{T_0} is not inflated by the stagnation safeguard.

Methodological scope and limitations. Several limitations bound the scope of these results. (i) The AWH result is obtained through a ground-truth-aware score and is therefore not within the strict measurable-constraint scope of Model 2; it is best read as the response of the framework to a controlled proxy of a properly supervised training regime. (ii) The study uses purely synthetic data with no measurement noise on the BOS images, on T_0 or on the pressure taps; the robustness of the $p_s \to \mathrm{EOS} \to \rho$ gauge-breaking pathway to realistic noise is therefore untested. (iii) The reconstruction is two-dimensional and single-line-of-sight, and (iv) all results come from a single fixed-seed run; both limitations are common to the three models and are consolidated in Chapter 8. (v) The AWH introduces a sizeable set of hyperparameters that are not individually ablated. (vi) The *minimality* of the constraint set is not established: no single-channel ablation is reported, so the present results demonstrate sufficiency of the full set $\{T_0, \{p_{s,k}\}, \mathrm{EOS}\}$ under AWH.

The dependence on K and on global isentropy deserves particular emphasis, because the gauge-breaking mechanism rests on it. The chain $\mathcal{L}_{p_s} \to \mathcal{L}_{\mathrm{EOS}} \to \rho_\theta$ propagates the tap-anchored pressure to the absolute density only because K is known and the flow is globally isentropic, so that a single constant relates p and ρ throughout the domain. In a transonic VKI cascade with shocks this assumption fails across the shock surfaces, where entropy rises and a single global K no longer holds. The mechanism would then have to be reconsidered with a non-isentropic closure or with K treated as a spatially varying field or as an unknown, and its effectiveness in that regime cannot be inferred from the present synthetic, isentropic case.

The recurring fragile point across these limitations is the absolute level of ρ, which depends on the supplemental constraints and on their relative weighting, whereas the gradient content recovered by the BOS data alone is robust. This observation motivates the transition to Model 3, which isolates the gradient channels and treats the absolute level separately, following the controlled-erosion progression that runs through the three models.

7 | Model 3: A Reduced UBOS Formulation

7.1. Purpose and Scope

Model 3 occupies the weakly constrained end of the framework. The network is reduced to a scalar density approximator, and the only data-fidelity term retained is the BOS data loss of Chapter 4, complemented by a weak Laplacian regulariser. No component of the objective acts on the absolute level: the BOS loss is gauge-invariant by Eq. (4.6) and the Laplacian annihilates constant shifts, so Model 3 does *not* break the null-space gauge but isolates the gradient-domain information the BOS measurement alone delivers.

7.2. Model Formulation

Model 3 instantiates the shared UBOS/PINN framework with the architecture, input normalisation (Eq. (4.2)) and patch-based optimisation (Eqs. (4.8)–(4.10)) of Models 1 and 2, but with a density-only output, a single positivity-constrained parameterisation and a deliberately reduced objective. The swish-MLP of Table 4.1 is here parameterised as a scalar map

$$q_\theta : (\tilde{x}, \tilde{y}) \;\longmapsto\; r_\rho, \tag{7.1}$$

from the dimensionless coordinates of Eq. (4.2) to a single raw pre-activation. The output dimension is reduced to $n_q = 1$ because the only retained terms, the BOS data term and the smoothness regulariser, act on the density alone. The physical density fed to the BOS data loss is recovered through the transformation

$$\rho_\theta \;=\; \mathrm{softplus}(r_\rho) \;+\; \varepsilon, \tag{7.2}$$

with softplus$(z) = \log(1 + e^z)$ and ε a small positive constant. The motivation is the one given in Section 6.2. Beyond the positivity and smooth-differentiability motivation common to Model 2. The output-layer bias is initialised at $\text{softplus}^{-1}(1)$, so an untrained q_θ predicts a uniform $\rho_\theta \approx 1\,\mathrm{kg\,m^{-3}}$.

The discrete UBOS operator is applied to ρ_θ exactly as in Models 1 and 2: $\mathcal{A}$ (Eq. (4.3)), realised through the fourth-order stencils of Eq. (4.4), is evaluated on patches drawn from Eq. (4.10) and compared with the training signal $\mathbf{b}$.

A first methodological departure from Eq. (4.5) concerns how the patch residuals $\mathcal{A}[\rho_\theta] - \mathbf{b}$ are aggregated into a scalar data loss. Where the two preceding models adopt the mask-normalised mean square mismatch of Eq. (4.5) directly, weighting every valid patch pixel equally, Model 3 replaces this uniform average by a $|b|$-weighted one that emphasises the patch pixels carrying the largest absolute BOS signal. Specifically, on each sampled patch a weighting map is constructed as

$$w(\mathbf{x}) \;=\; m(\mathbf{x})\left(1 \;+\; \lambda_b \frac{|\mathbf{b}(\mathbf{x})|}{\overline{|\mathbf{b}|} \;+\; \varepsilon}\right), \qquad \overline{|\mathbf{b}|} \;=\; \frac{\sum_{\mathcal{P}} m\,|\mathbf{b}|}{\sum_{\mathcal{P}} m \;+\; \varepsilon}, \tag{7.3}$$

where $m \in \{0,1\}^{P\times P}$ is the validity mask of Eq. (4.5) restricted to the patch, $\overline{|\mathbf{b}|}$ denotes the mask-averaged absolute BOS signal on the same patch and $\lambda_b \geq 0$ is a non-negative scalar that modulates the intensity of the additional emphasis. The corresponding data loss takes the form of a weighted mask-normalised mean square mismatch,

$$\mathcal{L}_{\text{meas}}(\theta) \;=\; \frac{\sum_{\mathcal{P}} w(\mathbf{x})\,\big(\mathcal{A}[\rho_\theta] - \mathbf{b}\big)^2}{\sum_{\mathcal{P}} w(\mathbf{x}) \;+\; \varepsilon}, \tag{7.4}$$

which reduces to the shared $\mathcal{L}_{\text{meas}}$ of Eq. (4.5) in the limit $\lambda_b \to 0$ and progressively shifts the per-pixel emphasis toward high-$|b|$ regions as λ_b is increased. Eq. (7.4) acts in synergy with the BOS-weighted patch sampler of Eq. (4.10): the sampler determines *which* patches the optimiser sees most often, while the weighting of Eq. (7.3) determines which pixels within a sampled patch dominate the patch gradient.

The second component of the Model 3 objective is a weak smoothness regulariser that penalises the discrete Laplacian of the predicted density on the patch grid. The Laplacian is implemented through the five-point centred stencil

$$(\Delta\rho_\theta)_{ij} \;=\; \big(\rho_{\theta,i,j+1} \;-\; 2\rho_{\theta,i,j} \;+\; \rho_{\theta,i,j-1}\big) \;+\; \big(\rho_{\theta,i+1,j} \;-\; 2\rho_{\theta,i,j} \;+\; \rho_{\theta,i-1,j}\big), \tag{7.5}$$

realised as a two-dimensional convolution with reflective padding, and the associated smoothness loss is the mask-normalised mean square of the discrete Laplacian on the patch,

$$\mathcal{L}_{\text{smooth}}(\theta) \;=\; \frac{\sum_{\mathcal{P}} m \left(\Delta\rho_\theta\right)^2}{\sum_{\mathcal{P}} m \,+\, \varepsilon}, \tag{7.6}$$

where m is once again the validity mask of Eq. (4.5). The stencil of Eq. (7.5) is written in discrete form, the grid-spacing factors being absorbed in w_{smooth}, since $\mathcal{L}_{\text{smooth}}$ suppresses small-scale roughness rather than approximating a physical operator. Two properties matter: the discrete Laplacian annihilates affine functions, so the term shares the null-space of Eq. (4.6) and does not break the gauge; and the squared Laplacian is a convex isotropic curvature penalty that disfavours small-scale oscillations without biasing global structure. $\mathcal{L}_{\text{smooth}}$ is thus a conditioning aid, not a source of supplemental information.

The composite objective minimised by Model 3 collects these two contributions in the weighted form

$$\mathcal{L}^{(3)}(\theta) \;=\; w_{\text{meas}}\,\mathcal{L}_{\text{meas}}(\theta) \;+\; w_{\text{smooth}}\,\mathcal{L}_{\text{smooth}}(\theta), \tag{7.7}$$

with the model-specific weights $w_{\text{meas}} = 1$, $w_{\text{smooth}} = 10^{-4}$ and the data-reweighting parameter $\lambda_b = 5$, whose relative magnitudes reflect the methodological hierarchy intended in Model 3. The unit-order w_{meas} lets the $|b|$-reweighted BOS data loss of Eq. (7.4) drive the optimisation and provide the dominant gradient signal on ρ_θ. The four-orders-smaller w_{smooth} keeps $\mathcal{L}_{\text{smooth}}$ a weak conditioning aid that does not compete with the data term on its own gradient-domain ground: a unit-order value would systematically attenuate the steep gradient features of ρ_θ that the BOS observation is meant to recover, whereas a vanishing one would leave the optimiser exposed to small-scale roughness that the second-order differentiation of the discrete Laplacian would amplify. The value 10^{-4} rescales the smoothness contribution to a fraction of the data gradient in the network's typical operating regime, so the regulariser intervenes only where the BOS data are locally insufficient to determine ρ_θ unambiguously. The resulting architecture is sketched in Fig. 7.1.

7.3. Results

A representative checkpoint was selected through the uniform protocol of Chapter 4, with the score restricted to the relative L_2 errors on the components of $\nabla\rho_\theta$ and its magnitude. The reconstruction of ρ at the selected epoch exhibits the quantitative signature anticipated by the gauge analysis of Eq. (4.6) in its most pronounced form. Table 7.1 reports a relative L_2 error on ρ of 0.549, an RMSE of $0.668\,\mathrm{kg\,m^{-3}}$ and a MAE of $0.668\,\mathrm{kg\,m^{-3}}$, with a linear correlation coefficient with the ground truth that is essentially unity and a MAPE of 54.9%. The global bias of $-0.668\,\mathrm{kg\,m^{-3}}$ coincides with the MAE, the RMSE and the

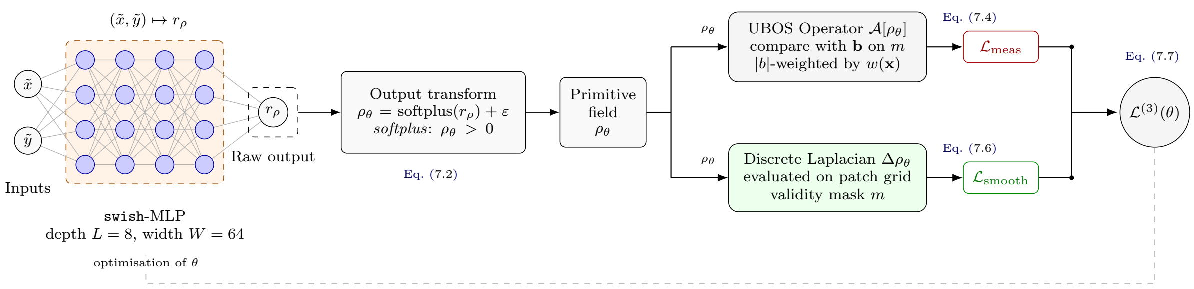


Figure 7.1: Reduced UBOS/PINN architecture of Model 3. Normalised coordinates $(\tilde{x},\tilde{y})$ (Eq. (4.2)) feed a swish-MLP ($L=8$, $W=64$). The predicted density is passed through the discrete UBOS operator $\mathcal{A}$ (Eq. (4.3)) and compared patch-wise with $\mathbf{b}$ via the $|b|$-weighted loss $\mathcal{L}_{\mathrm{meas}}$ (Eqs. (7.4)–(7.3)). A weak Laplacian regulariser $\mathcal{L}_{\mathrm{smooth}}$ (Eq. (7.6)) suppresses small-scale roughness. Composite objective (Eq. (7.7)): $w_{\mathrm{meas}}=1$, $w_{\mathrm{smooth}}=10^{-4}$, $\lambda_b=5$. No physics-informed, equation-of-state, boundary, total-temperature or static-pressure terms are included, in contrast with Models 1 and 2.

p_{95} absolute error of $0.668\,\mathrm{kg\,m^{-3}}$ to within the displayed precision, while the maximum absolute error of $0.669\,\mathrm{kg\,m^{-3}}$ differs from them only at the third significant digit. Because the pointwise discrepancy on ρ is almost exactly the uniform bias $-0.668\,\mathrm{kg\,m^{-3}}$, the relative L_2 error on ρ is set by the size of that offset relative to the mean ground-truth density, and it is this ratio that places the level error above fifty percent. The pointwise error distribution on ρ is therefore concentrated, almost entirely, on a single negative shift of the predicted field with respect to the ground truth. The density scatter plot of Fig. 7.2 confirms that the prediction-versus-ground-truth point cloud is aligned along a line parallel to the identity but rigidly translated below it by an essentially constant offset, with no slope distortion or visible saturation.

Table 7.1: Reconstruction metrics for Model 3 at epoch 800, selected by the gradient-domain composite score. Errors on density, density-gradient components and BOS forward field against synthetic references.

Field	L_2 rel	RMSE	MAE	MAPE [%]	corr	bias
ρ	0.549	0.668	0.668	54.9	1.00	-0.668
$\partial_x\rho$	0.0083	0.0107	0.0065	—	1.00	-0.0011
$\partial_y\rho$	0.0111	0.0143	0.0084	—	1.00	-7.73×10^{-4}
$\lvert\nabla\rho\rvert$	0.0095	0.0173	0.0108	—	1.00	0.0084
b_{pred}	0.0464	6.23	1.48	—	0.999	-0.577

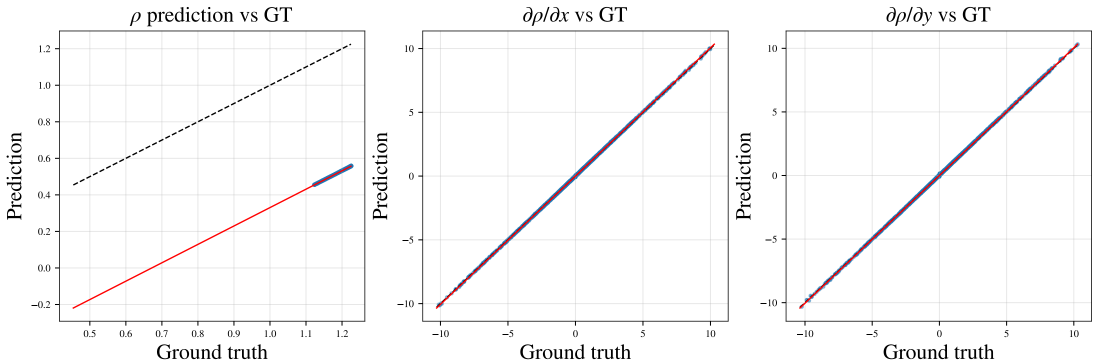


Figure 7.2: Pixelwise prediction vs. ground truth for ρ and the two in-plane density-gradient components at the best Model 3 checkpoint. The density panel shows a uniform shift below the identity, directly expressing the gauge null-mode of Eq. (4.6); both gradient panels collapse onto the identity across the full dynamic range.

The accuracy on the components of $\nabla\rho$ is, in sharp contrast with the level-channel behaviour, the highest of the three reconstructions analysed so far. The relative L_2 errors for $\partial_x\rho$, $\partial_y\rho$, and $|\nabla\rho|$ in Model 3 (0.0083, 0.0111, 0.0095) are uniformly lower than the corresponding values in Model 1 (0.0267, 0.0267, 0.0266) and Model 2 (0.0156, 0.0153, 0.0151). The juxtaposition of the density and gradient figures isolates, in its sharpest possible form, the gauge-aware signature of the reduced formulation. The spatial maps of Fig. 7.3 document the same picture in two-dimensional form. The gradient field is recovered to sub-percent relative error away from the core, while the residual concentrates on the narrow annulus where the gradient is steepest. This localised degradation at the sharpest feature is the single-discontinuity manifestation of the same spectral-bias limitation that the transonic shocks of a real campaign would accentuate, and it is taken up in that wider context in the concluding chapter. The corresponding p_{95} absolute errors of 0.0226 and 0.031 on the two gradient components confirm that the bulk of the gradient error remains far below the maximum value of 0.127, the MAPE is omitted consistently with the convention of Chapters 5–6: where computed, it is dominated by the small-magnitude tails of $\nabla\rho$ in the freestream region rather than by errors on the structured part of the field, and carries no additional information beyond the relative L_2 error.

The forward consistency with the BOS observation is essentially preserved with respect to the two preceding models. Re-applying the discrete UBOS operator of Eq. (4.3) to ρ_θ yields a predicted field b_{pred} with relative L_2 error 0.0464, correlation 0.999 and bias -0.577, within a few percentage points of the Model 1 (0.0527) and Model 2 baseline (0.0486) figures, despite the order-of-magnitude separations on the absolute-level channel in Table 7.1.

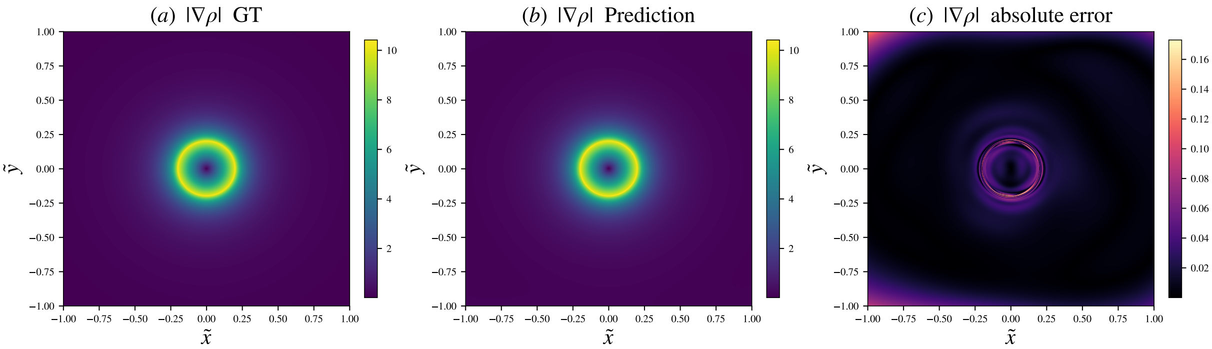


Figure 7.3: Ground truth, prediction and absolute error of $|\nabla\rho|$ at the best Model 3 checkpoint, evaluated via fourth-order stencils (Eq. (4.4)). The residual concentrates in the narrow annulus around the Rankine core radius; the bulk is recovered within sub-percent relative error. No companion density map is shown: the absolute error on ρ is essentially uniform by gauge construction (Eq. (4.6); Table 7.1) and carries no spatially resolved information.

7.4. Patch-Geometry Sweep and Operating Point

Because Model 3 isolates the gauge-invariant UBOS component without the confounding physics-informed and constitutive residuals, it provides the cleanest setting in which to probe the dependence of the reconstruction on the two governing hyperparameters of the patch-based strategy of Chapter 4: the batch size B and the patch side P. A parametric sweep was conducted on Model 3 over seven configurations: $B{=}4$ for $P \in \{32, 64, 128, 256\}$ and $B{=}8$ for $P \in \{32, 64, 128\}$. Each run is scored by a composite metric defined as the unweighted sum of the relative L_2 errors on $\partial_x\rho$, $\partial_y\rho$ and $|\nabla\rho|$. Fig. 7.4 reports the loss-convergence behaviour against the computational budget. The total loss at the best epoch decreases as the patch side grows. Larger patches expose the discrete UBOS operator to a larger contiguous region per optimisation step and therefore yield a lower-variance estimate of the data-loss gradient. The benefit is bought at a steeply rising cost: the number of pixels processed at the best epoch grows from 1.05×10^9 at $P = 32$ to 2.68×10^{10} at $P = 256$, and the wall time per epoch rises from approximately 4 s to 13 s at the largest patch. Increasing the batch size to $B = 8$ accelerates the per-epoch convergence but, at fixed P, yields no gain in final gradient accuracy and even slightly degrades it at $P = 128$, an indication that beyond a certain point the additional patches are drawn from an increasingly overlapping high-$|b|$ region and contribute redundant gradient information.

The resulting cost–quality trade-off, made explicit by the Pareto analysis of Fig. 7.5, is what turns this sweep into a methodological guideline. Under the per-epoch cost criterion (time per epoch versus score) the Pareto front is delineated by the configurations

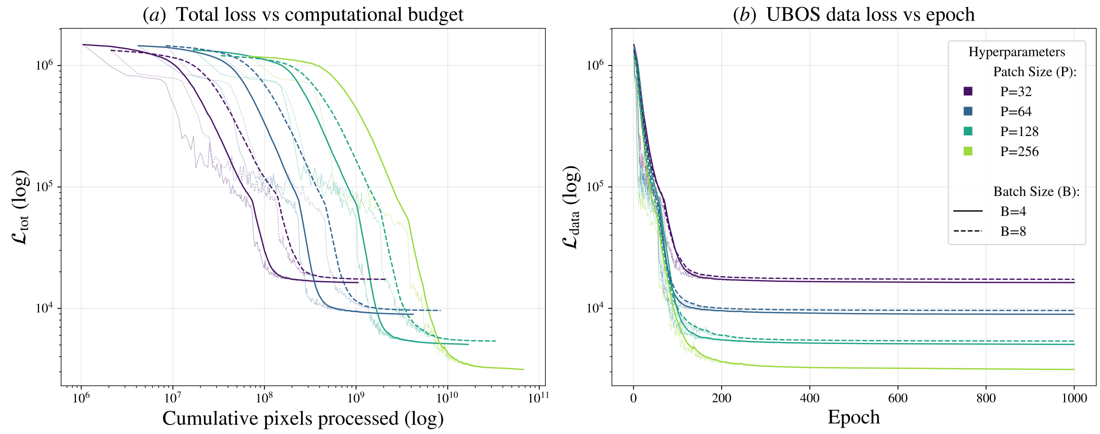


Figure 7.4: Convergence of the total and data loss of Model 3 for the seven (B, P) configurations of the parametric sweep, plotted against the computational budget expressed in processed pixels. The loss at the best epoch decreases with the patch side P, a qualitative and non-homogeneous comparison, while the pixel budget and the wall-clock cost grow super-linearly, framing the efficiency trade-off analysed in Fig. 7.5.

$B{=}4, P{=}64$, $B{=}4, P{=}128$ and $B{=}4, P{=}256$: no other configuration is simultaneously faster and more accurate than these three. Under the global time-to-convergence criterion (total time to the best epoch versus score) the front shifts to $B{=}8, P{=}32$, $B{=}4, P{=}128$ and $B{=}4, P{=}256$, since the larger batch reduces the number of epochs needed to reach the stationary regime. The configuration $B{=}4, P{=}128$ is the only one lying on both Pareto fronts, providing an objective basis for identifying it as the most robust operating point and recommending it for future campaigns within this framework.

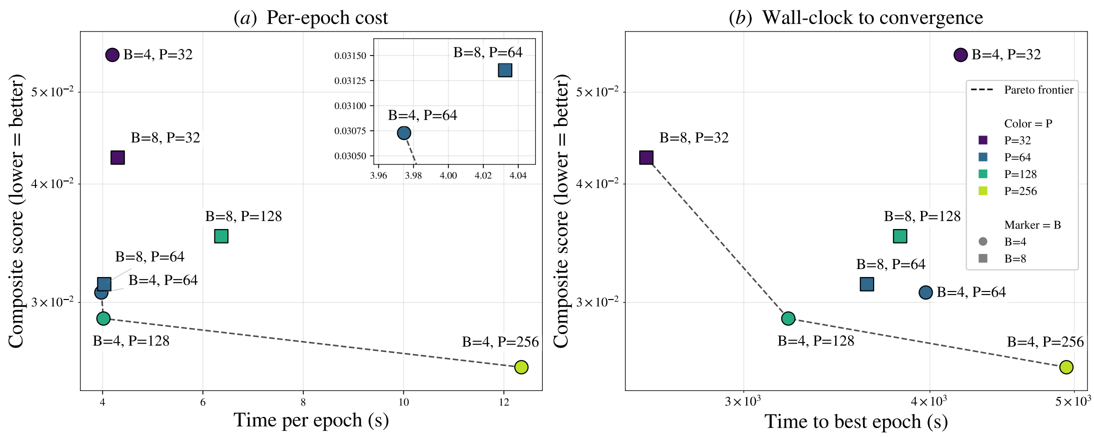


Figure 7.5: Composite gradient-domain score of the seven Model 3 configurations against the computational cost, for the per-epoch cost criterion and the global time-to-convergence criterion.

7.5. Interpretation

The metrics of Table 7.1, the spatial diagnostics of Figures 7.2–7.3 and the patch-geometry sweep of Section 7.4 admit a coherent methodological reading, organised below around three interpretive points, from the gauge property in its pure form to the channel-resolved bounding role assigned to Model 3 in the chapter's ordered progression.

Saturation of the gauge null-mode. The defining quantitative outcome of Model 3 is that the relative L_2 error on ρ exceeds fifty percent and is entirely accounted for by a single uniform shift: the bias coincides with the MAE, the RMSE and the p_{95} absolute error of Table 7.1 within the displayed precision. This is a literal numerical realisation of Eq. (4.6): neither $\mathcal{L}_{\text{meas}}$ nor $\mathcal{L}_{\text{smooth}}$ constrains the constant component of ρ_θ, so the optimisation lowers the data residual to its irreducible gradient-domain value while leaving the level free, its sign and magnitude set only by the bias initialisation and the null-mode dynamics. The near-unity correlation 1.00 makes the diagnostic doubly informative: the zero-mean spatial structure of ρ_θ is recovered with near-perfect fidelity, and the discrepancy is concentrated entirely on the single dimension that $\mathcal{L}_{\text{meas}}$ leaves invariant. Model 3 therefore does not measure a failure of the framework but the irreducible content of the BOS observation itself, isolated from any supplementary information.

Concentration of the optimisation budget on the gradient channel. Model 3 attains the highest gradient-domain accuracy of the three reconstructions, improving uniformly and channel by channel on Model 1 and the Model 2 baseline of Table 7.1, by roughly a factor of one-and-a-half to three depending on the channel. This ordering admits a natural, though not uniquely identified, explanation.

With respect to Models 1 and 2, Model 3 differs in three respects at once: the scalar output retires the coupling of ρ_θ to $u_\theta, v_\theta, p_\theta$ through $\mathcal{L}_{\text{phys}}$ and $\mathcal{L}_{\text{EOS}}$; the data loss is reweighted toward high-$|b|$ pixels through λ_b (Eq. (7.3)); and the loss-weight vector changes accordingly. The hypothesis favoured here is that the dominant factor is the first: the retired physics residuals act on the gradient channel as a smoothing constraint that attenuates the steepest density gradients, so that their removal frees every network-gradient component to act on the field the BOS observation directly probes. A controlled ablation separating this effect from the $|b|$-reweighting (e.g. Model 3 at $\lambda_b = 0$, and Models 1–2 with the reweighted data loss) has not been performed and is required before the attribution can be considered established. Under this (unconfirmed) attribution, the Model 3 gradient pathway is the most direct the framework allows, which is the empirical motivation for the hybrid pipeline of Chapter 8.

Methodological scope and channel-resolved bounding role. The gradient-domain picture must be qualified by what Model 3 does *not* prove. The reconstruction is gauge-degenerate by design and no claim is made about the absolute level of ρ: the offset in Table 7.1 is not a failure mode but the expected outcome of an optimisation problem whose objective admits a one-dimensional null-space.

Within this delimited scope, Model 3 and Model 1 bound the framework from opposite directions, but the bound is meaningful only channel by channel, not on a single global notion of reconstruction performance. On the absolute-level channel, Model 3 is the lower reference: with the gauge left open, its level error is the largest of the three models, whereas Model 1, anchored by dense Dirichlet supervision, is the upper reference. On the gradient channel the ordering is reversed: Model 3 attains the highest gradient-domain accuracy of the three and is therefore the upper reference there, while the physics-constrained models sit below it. The two roles are held on the very same network architecture, input normalisation, patch-based optimisation strategy and discrete UBOS operator, which is what makes the comparison controlled.

Because the ordering is channel-dependent, the benefit of a supplementary channel cannot be summarised as a single displacement of Model 3 toward Model 1: any such channel, the dense Dirichlet anchoring of Model 1, the minimal measurable set $\{T_0, \{p_{s,k}\}, \mathrm{EOS}\}$ of Model 2, or any future addition, improves the absolute level while, as the gradient figures show, its physics residuals degrade the gradient channel. The calibrating function exploited by the joint reading of the three models in the concluding chapter therefore applies specifically to the absolute-level channel, where Model 3 fixes the lower reference against which each supplementary channel is measured.

The data-reweighting intensity $\lambda_b = 5$ is, moreover, adopted without a sensitivity study: its interaction with the patch sampler of Eq. (4.10), which already biases the optimisation toward high-$|b|$ regions, is not individually ablated.

8 | Conclusions and Future Work

This chapter draws the three reconstruction experiments of Chapters 5–7 together as a single investigation of the density inverse problem: it synthesises the main findings, develops the gradient/level dichotomy into a hybrid reconstruction proposal, and outlines the transition toward the experimental BOS data of the VKI transonic cascade.

8.1. Summary of the Main Findings

The principal outcome is that the UBOS operator can be embedded directly into a physics-informed pipeline, with no intermediate deflection field. Coupling the discrete operator of Eq. (4.3) to a compressible flow representation, the framework of Chapter 4 ties the optical measurement to the density field through a single differentiable relation, a one-stage alternative to the conventional two-stage chain.

The second finding concerns the role of thermodynamic information. The intrinsic null-space property of the UBOS operator leaves the absolute density level undetermined: the data loss constrains the gradient structure of ρ but is invariant under uniform shifts. Model 2 suggests that the $\{T_0, \{p_{s,k}\}, \mathrm{EOS}\}$ constraint set is, in principle, sufficient to lift this indeterminacy. This evidence, however, rests on the AWH, a supervised proxy scored against the synthetic ground truth and the only offset-removing configuration. The thermodynamic constraints thus appear to carry the absolute-level information only conditional on an unsupervised weighting strategy that a fully measurable-only pipeline has yet to demonstrate; the algebraic level-calibration of the two-step pipeline proposed in Section 8.2 (Step 2) would constitute the decisive, supervision-free test of this sufficiency.

The third finding is the explicit characterisation of the trade-off between gradient fidelity and absolute-level accuracy, demonstrated by Model 3 (Chapter 7). With every absolute-level channel retired and the objective reduced to a gauge-invariant data loss plus a weak Laplacian regulariser, the reconstruction recovers the density gradient with the highest accuracy of the three models (relative L_2 errors of order 10^{-2} on $\partial_x\rho$, $\partial_y\rho$ and $|\nabla\rho|$), while the relative L_2 error on the level of ρ exceeds fifty percent and is entirely accounted for by a single uniform shift. Conversely, the supervision-rich Models 1 and 2 recover the absolute level at the cost of degraded steepest-gradient features (relative L_2 errors up to about a factor of three above Model 3). Read in sequence, the three models thus delimit operationally the information content of the BOS inverse problem: the optical measurement chain guarantees the *gradient structure* of ρ, whereas the *absolute level* requires the injection of thermodynamic information.

A fourth finding establishes that this intrinsic gradient channel is actively controllable: its recovery quality depends measurably on how the patch-based optimiser is fed. The BOS-weighted importance-sampling scheme of Eqs. (4.7)–(4.10), supported by the cross-run campaign of Section 4.3 (Figures 4.2–4.3), reduces the relative L_2 error on the gradient channels by more than an order of magnitude with respect to a uniform baseline, with the largest gains in the gradient-rich Rankine core.

8.2. Methodological Implications

The third finding carries the central methodological message of this thesis: the density inverse problem separates into an intrinsic gradient channel, recovered optically, and an extrinsic absolute-level channel requiring thermodynamic information.

The gradient channel itself was further characterised parametrically. The patch-geometry sweep of Section 7.4 varied the batch size B and the patch side P over seven configurations of Model 3, and identified $B{=}4, P{=}128$ as the configuration lying on both Pareto fronts.

The AWH of Section 6.3 provides a revealing perspective on the gradient/level dichotomy. As recalled in Section 8.1, its order-of-magnitude reduction of the error on ρ, obtained while the gradient channels degrade only marginally, cannot by itself separate a static-weighting limitation from a structural inadequacy of the $\{T_0, \{p_{s,k}\}, \mathrm{EOS}\}$ constraint set, because the AWH score is supervised; what can be stated unambiguously is that the channel on which AWH improves most is precisely the one the gauge leaves invariant. This motivates a working hypothesis: since the data loss alone already recovers the gradient structure to high accuracy, a procedure acting *exclusively* on the absolute-level channels should in principle cancel the density offset without disturbing the gradient-domain reconstruction.

A further implication concerns the spatial granularity of the pressure-anchoring channel. In Model 2 the static-pressure taps anchor the absolute level through a reference scale $\overline{p_s}$, defined as the mean of the available taps over the whole domain. To the extent that the absolute level reaches ρ_θ predominantly through the local pressure–EOS chain, it would be worth deriving the reference values *dynamically and regionally*, averaging the tap measurements over physically meaningful subdomains: a region-resolved $\overline{p_s}$ could adapt the anchoring to the local thermodynamic state and mitigate the bias a single global reference incurs whenever the pressure field departs appreciably from spatial uniformity.

A final implication closes the loop with the optical layer on which the framework ultimately rests. The two information channels are meaningful only insofar as the discrete UBOS operator faithfully encodes the optical-flow physics, and the 468-run sensitivity campaign of Chapter 3 showed this fidelity to be bounded by a markedly asymmetric hardware–software coupling: of the operator-side breakdown at $\sigma_{\text{blur}} \to 0$ (sharp) and the bandwidth-loss degradation at $\sigma_{\text{blur}} \gg d_{\min}/2$ (gradual) characterised there, the low-blur discretisation failure is the sharper and more dangerous. The methodological consequence is concrete: the gradient-fidelity guarantees of Model 3, and hence the gradient/level partition itself, are conditional on operating well inside the bandlimited regime, so the optical hardware of any experimental campaign must keep the realised background pattern and displacement statistics within the empirical feasibility band $\xi \lesssim 0.05$. The data-assimilation conclusions of this thesis cannot be divorced from the optical-design constraints that generate their input.

These considerations converge on a concrete methodological proposal: a *two-step hybrid reconstruction pipeline* that exploits the strengths of the reduced and constrained formulations while avoiding the gradient/level trade-off of Models 1 and 2. In the first step, Model 3 acts as a pure gradient extractor: its reduced UBOS formulation, run at the operating point $B{=}4, P{=}128$, recovers $\partial_x\rho$, $\partial_y\rho$ and $|\nabla\rho|$ at the highest accuracy the framework allows and returns a gauge-degenerate field ρ_1, the absolute level left explicitly free as the additive uniform shift undetermined by Eq. (4.6). In the second step, that shift is fixed not by re-optimising the network under competing residuals but by an algebraic post-processing of the measurable thermodynamic constraints. The pipeline thus proposes an operational synthesis of the three models, and the benchmarking campaigns support its individual stages rather than the integrated whole: the importance-sampling result certifies the Step 1 gradient extractor and the patch-size sweep fixes its operating point. Its end-to-end validation against Models 1 and 2 on the same synthetic data is left to future work.

8.3. Future Work

The framework has been validated entirely on the synthetic isentropic Rankine vortex of Chapter 3: a subsonic, globally isentropic, shock-free flow with sub-pixel displacements. The natural continuation is therefore the transfer to the experimental BOS data acquired in the CT2-SS transonic wind tunnel of the von Karman Institute, where exactly those phenomena are present. This transition raises four challenges, each defining a concrete direction for future research, to be read against the limitations set out below.

Limitations of the present study. Five limitations bound the scope of the results above and shape the priorities for future work. First, every reconstruction is a *single run*. Second, the entire validation is *noise-free*: no measurement noise is applied to the BOS images, to the T_0 field or to the pressure taps. Third, the reconstruction is two-dimensional and single-line-of-sight. Fourth, as analysed in Section 6.5, the gauge-breaking mechanism rests on a known K and on global isentropy, both of which fail across transonic shocks and would require a non-isentropic closure. Fifth, no head-to-head quantitative comparison against the conventional two-stage pipeline (optical-flow displacement extraction followed by Poisson integration, Chapter 2) is reported on the same synthetic data; since the one-stage formulation is the central methodological claim of this work, such a controlled comparison is a necessary validation left to future work.

Replacing the AWH with unsupervised weighting heuristics. The favourable absolute-level results of Model 2 rest on the AWH, whose feedback score (Eq. (6.11)) is built from relative L_2 errors against the synthetic ground truth and is therefore not computable in a real experimental campaign (Section 6.3). The first priority is thus the development of dynamic weighting algorithms operating exclusively on measurable quantities, such as the physical residuals and the measurement-channel mismatches. The label-free candidates already noted in Section 6.3 (NTK- and gradient-statistic-based rules [43, 45]) diagnose loss imbalance from the optimisation dynamics themselves and would render the training of Model 2 fully autonomous and experimentally deployable.

Architectural transition towards Neural Operators. The three models studied here are instance-based: each reconstruction requires a complete training run for a single test case, a *per-instance* cost acceptable for a methodological study but prohibitive for an experimental campaign comprising many operating points, all the more so given the steep super-linear scaling of pixel budget and wall-clock cost with patch size (Section 7.4). The decisive architectural step is the transition from PINNs to Neural Operators, such as the Deep Operator Network (DeepONet) [23] or the Fourier Neural Operator (FNO) [40], which learn the physical solution operator rather than a single field: a trained operator would map the BOS image and the pressure-tap data directly to the density field, generalising across flow configurations without per-case retraining.

Generation of a synthetic CFD meta-dataset. Training a Neural Operator requires a large and varied corpus of input–output pairs, which a single synthetic case cannot supply. A dedicated meta-dataset should be generated through high-fidelity CFD simulations spanning a broad range of inlet Mach numbers, back-pressure conditions and shock topologies representative of the VKI transonic cascade, augmented with optical-noise models. The target is not generic: the midspan reference of Chapter 2 exhibits the characteristic three-shock pattern at the design point [7], each shock imposing a nearly discontinuous transient on the density field, so the meta-dataset must densely sample the operating points at which these features appear, migrate and interact, testing the Neural Operator on their reproduction rather than merely on the smooth bulk flow.

Real-time zero-shot inference for the CT2-SS facility. The ultimate objective of this research line is the deployment of a trained Physics-Informed Neural Operator within the acquisition system of the CT2-SS facility, enabling *zero-shot* reconstruction of the density, pressure and velocity fields directly from the BOS imagery, without any test-specific optimisation. Real-time inference would let the experimentalist visualise quantitative flow data concurrently with the run, shortening the campaign turnaround and ultimately supporting active control of the test parameters during acquisition.

8.4. Closing Remarks

The demonstrated contribution is the quantitative partition of the density inverse problem holding within the ξ-defined feasibility envelope and supported by the importance-sampling and patch-size benchmarks. The principal limitations are equally explicit: the operator-side failure modes of the discrete UBOS forward map, the planar reduction of an intrinsically three-dimensional flow, and the distance between the subsonic shock-free validation case and the transonic, shock-dominated target. Against this backdrop the two-step hybrid pipeline is offered not as a demonstrated method but as a proposal: a route by which the gradient and absolute-level channels, structurally orthogonal, might be separated in practice. Its end-to-end validation, together with the move toward unsupervised weighting, Neural Operators and a synthetic CFD meta-dataset, marks the natural continuation of this work toward a deployable diagnostic instrument.

Bibliography

[1] S. Arridge, P. Maass, O. Oktem, and C.-B. Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.

[2] B. Atcheson, W. Heidrich, and I. Ihrke. An evaluation of optical flow algorithms for background oriented schlieren imaging. *Experiments in fluids*, 46(3):467–476, 2009.

[3] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18 (153):1–43, 2018.

[4] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. Compressed sensing using generative models. In *International Conference on Machine Learning (ICML)*, pages 537–546. PMLR, 2017.

[5] J. C. Butcher. *Numerical methods for ordinary differential equations.* John Wiley & Sons, 2016.

[6] S. Cai, Z. Mao, Z. Wang, M. Yin, and G. E. Karniadakis. Physics-informed neural networks (pinns) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12): 1727–1738, 2021.

[7] F. Cospito. Background-oriented schlieren (bos) on a supersonic compressor linear cascade. Research Master Project Report RM-Report 18, von Karman Institute for Fluid Dynamics, Turbomachinery and Propulsion Department, Sint-Genesius-Rode, Belgium, June 2024. Project part of the WINGS program (Walloon Innovations for Green Skies).

[8] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

[9] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.

[10] M. G. Dissanayake and N. Phan-Thien. Neural-network-based approximations for solving partial differential equations. *communications in Numerical Methods in Engineering*, 10(3):195–201, 1994.

[11] B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of computation*, 51(184):699–706, 1988.

[12] GE Reports. Gaining altitude: Airbus a380 to test cfm's open-fan architecture in flight, 7 2022.

[13] E. Goldhahn and J. Seume. The background oriented schlieren technique: sensitivity, accuracy, resolution and application to a three-dimensional density field. *Experiments in fluids*, 43(2):241–249, 2007.

[14] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6):3055–3071, 2018.

[15] M. J. Hargather and G. S. Settles. A comparison of three quantitative schlieren techniques. *Optics and Lasers in Engineering*, 50(1):8–17, 2012.

[16] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial intelligence*, 17 (1-3):185–203, 1981.

[17] A. D. Jagtap, E. Kharazmi, and G. E. Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.

[18] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017.

[19] X. Jin, S. Cai, H. Li, and G. E. Karniadakis. Nsfnets (navier–stokes flow nets): Physics-informed neural networks for the incompressible navier–stokes equations. *Journal of Computational Physics*, 426:109951, 2021.

[20] E. Kharazmi, Z. Zhang, and G. E. Karniadakis. hp-vpinns: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374:113547, 2021.

[21] B. J. Kingma, Diederik P. Adam: A method for stochastic optimization. 2015. *arXiv preprint arXiv:1412.6980*, 15:6–5, 2015.

[22] X. Li, M. Gao, J. Li, C. Pan, J. Wang, and Y. Xiong. Improving spatial resolution of background-oriented schlieren based on directional rays. *Experiments in Fluids*, 67(4):40, 2026.

[23] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.

[24] L. D. McClenny and U. M. Braga-Neto. Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474:111722, 2023.

[25] J. P. Molnar and S. J. Grauer. Flow field tomography with uncertainty quantification using a bayesian physics-informed neural network. *Measurement Science and Technology*, 33(6):065305, 2022.

[26] J. P. Molnar, L. Venkatakrishnan, B. E. Schmidt, T. A. Sipkens, and S. J. Grauer. Estimating density, velocity, and pressure fields in supersonic flows using physics-informed bos. *Experiments in Fluids*, 64(1):14, 2023.

[27] J. P. Molnar, E. J. LaLonde, C. S. Combs, O. Léon, D. Donjat, and S. J. Grauer. Forward and inverse modeling of depth-of-field effects in background-oriented schlieren. *AIAA journal*, 62(11):4316–4329, 2024.

[28] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett. Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory*, 1(1):39–56, 2020.

[29] P. Perdikaris, M. Raissi, A. Damianou, N. D. Lawrence, and G. E. Karniadakis. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2198), 2017.

[30] P. Pilar and N. Wahlström. Physics-informed neural networks with unknown measurement noise. In *6th Annual Learning for Dynamics & Control Conference*, pages 235–247. PMLR, 2024.

[31] M. Raffel. Background-oriented schlieren (bos) techniques. *Experiments in Fluids*, 56(3):60, 2015.

[32] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

[33] E.-Z. Rui, G.-Z. Zeng, Y.-Q. Ni, Z.-W. Chen, and S. Hao. Time-averaged flow field reconstruction based on a multifidelity model using physics-informed neural network (pinn) and nonlinear information fusion. *International Journal of Numerical Methods for Heat & Fluid Flow*, 34(1):131–149, 2024.

[34] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[35] P. G. Saffman. Vortex dynamics. In *Theoretical Approaches to Turbulence*, pages 263–277. Springer, 1992.

[36] B. E. Schmidt, B. F. Bathel, S. J. Grauer, M. J. Hargather, J. T. Heineck, and M. Raffel. Twenty-five years of background-oriented schlieren: advances and novel applications. *AIAA journal*, 63(12):5028–5058, 2025.

[37] G. S. Settles and M. J. Hargather. A review of recent developments in schlieren and shadowgraph techniques. *Measurement Science and Technology*, 28(4):042001, 2017.

[38] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[39] A. N. Tikhonov and V. Arsenin. Solutions of ill-posed problems, 1977.

[40] A. Tran, A. Mathews, L. Xie, and C. S. Ong. Factorized fourier neural operators. *arXiv preprint arXiv:2111.13802*, 2021.

[41] R. Ulichney. *Digital halftoning*. MIT press, 1987.

[42] L. Venkatakrishnan and G. Meier. Density measurements using the background oriented schlieren technique. *Experiments in Fluids*, 37(2):237–247, 2004.

[43] S. Wang, Y. Teng, and P. Perdikaris. Understanding and mitigating gradient pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43 (5):A3055–A3081, 2021.

[44] S. Wang, S. Sankaran, and P. Perdikaris. Respecting causality is all you need for training physics-informed neural networks. *arXiv preprint arXiv:2203.07404*, 2022.

[45] S. Wang, X. Yu, and P. Perdikaris. When and why pinns fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, 449:110768, 2022.

[46] L. Yang, X. Meng, and G. E. Karniadakis. B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *Journal of Computational Physics*, 425:109913, 2021.

A | Weight-Update Rules

This appendix details the staged weight-update rule of the adaptive weight heuristic (AWH) introduced in Section 6.3, whose aggregated score S is defined in Eq. (6.11). At the end of each run the signed gap $\Delta S = \mathrm{clip}(S-1, -0.5, 2.0)$ acts as the primary feedback variable. To compute the updated raw weights $\tilde{w}_k$, the AWH applies a sequence of four multiplicative corrections controlled by a set of positive hyperparameter coefficients. Let $w_k^{[0]} = w_k$ denote the current weights before the update, and let $w_k^{[j]}$ denote the intermediate weights after step $j \in \{1, 2, 3, 4\}$. The final raw weights are then $\tilde{w}_k = w_k^{[4]}$:

1. **Global correction.** Shifts the budget between the data-fidelity channel and the gauge-breaking channels depending on the sign of ΔS. If $\Delta S \geq 0.0$ (model underperforming global targets), the weights are expanded using the positive coefficients $c_k^+ > 0$:
$$\begin{aligned} w_{\mathrm{meas}}^{[1]} &= w_{\mathrm{meas}}^{[0]}(1 + c_{\mathrm{meas}}^+ \Delta S) \\ w_{\mathrm{phys}}^{[1]} &= w_{\mathrm{phys}}^{[0]}(1 + c_{\mathrm{phys}}^+ \Delta S) \\ w_{\mathrm{EOS}}^{[1]} &= w_{\mathrm{EOS}}^{[0]}(1 + c_{\mathrm{EOS}}^+ \Delta S) \\ w_{T_0}^{[1]} &= w_{T_0}^{[0]}(1 + c_{T_0}^+ \Delta S) \\ w_{p_s}^{[1]} &= w_{p_s}^{[0]}(1 + c_{p_s}^+ \Delta S) \end{aligned}$$
If $\Delta S < 0.0$ (model outperforming targets), the budget is redistributed away from the data-fidelity channel and towards the gauge-breaking channels: w_{meas} is reduced while $w_{\mathrm{phys}}, w_{\mathrm{EOS}}, w_{T_0}, w_{p_s}$ are increased, under the positive coefficients $c_k^- > 0$; letting $R = -\Delta S$:
$$\begin{aligned} w_{\mathrm{meas}}^{[1]} &= w_{\mathrm{meas}}^{[0]}(1 - c_{\mathrm{meas}}^- R) \\ w_{\mathrm{phys}}^{[1]} &= w_{\mathrm{phys}}^{[0]}(1 + c_{\mathrm{phys}}^- R) \\ w_{\mathrm{EOS}}^{[1]} &= w_{\mathrm{EOS}}^{[0]}(1 + c_{\mathrm{EOS}}^- R) \\ w_{T_0}^{[1]} &= w_{T_0}^{[0]}(1 + c_{T_0}^- R) \\ w_{p_s}^{[1]} &= w_{p_s}^{[0]}(1 + c_{p_s}^- R) \end{aligned}$$

2. **Gradient-specific correction.** Reinforces the measurement weight when the mean raw gradient error $\bar{e}_g = (e_x + e_y + e_g)/3$ exceeds T_g, or when gradient errors exhibit an anisotropy $A_g = \max(e_x, e_y)/\min(e_x, e_y)$ beyond a threshold T_A. The remaining weights are left unchanged ($w_k^{[2]} = w_k^{[1]}$ for $k \neq$ meas):

$$w_{\text{meas}}^{[2]} = w_{\text{meas}}^{[1]}\big[1 + \eta_g \cdot \text{clip}(\bar{e}_g/T_g - 1, 0, D_{\max})\big] \cdot \begin{cases} \min(M_A, A_g^{\kappa_g}) & \text{if } A_g > T_A \\ 1 & \text{otherwise} \end{cases}$$

3. **Absolute-density correction.** Reinforces the physics and measurement constraints when the absolute density error exceeds the target ($e_\rho > T_\rho$), with strongest emphasis on w_{EOS} because the isentropic closure propagates the tap-anchored pressure level to ρ_θ. Letting $D_\rho = \text{clip}(e_\rho/T_\rho - 1, 0, D_{\max})$, and introducing the scaling coefficients $\beta_k > 0$, the updated weights are:

$$\begin{aligned} w_{\text{phys}}^{[3]} &= w_{\text{phys}}^{[2]}(1 + \beta_{\text{phys}} D_\rho) \\ w_{\text{EOS}}^{[3]} &= w_{\text{EOS}}^{[2]}(1 + \beta_{\text{EOS}} D_\rho) \\ w_{T_0}^{[3]} &= w_{T_0}^{[2]}(1 + \beta_{T_0} D_\rho) \\ w_{p_s}^{[3]} &= w_{p_s}^{[2]}(1 + \beta_{p_s} D_\rho) \end{aligned}$$

with $w_{\text{meas}}^{[3]} = w_{\text{meas}}^{[2]}$.

4. **Stagnation safeguard.** Triggered when the relative improvement of the aggregated score S with respect to the previous run (S_{prev}) falls below a minimal threshold $I_{\min}$. If triggered, the affected weights are multiplied by the constant inflation factors $\sigma_k > 1$:

$$\begin{aligned} w_{\text{meas}}^{[4]} &= w_{\text{meas}}^{[3]} \cdot \sigma_{\text{meas}} \\ w_{\text{phys}}^{[4]} &= w_{\text{phys}}^{[3]} \cdot \sigma_{\text{phys}} \\ w_{\text{EOS}}^{[4]} &= w_{\text{EOS}}^{[3]} \cdot \sigma_{\text{EOS}} \end{aligned}$$

Otherwise, $w_k^{[4]} = w_k^{[3]}$ for all k. The final raw weights are then set as $\tilde{w}_k = w_k^{[4]}$. The pointwise anchoring channels w_{T_0} and w_{p_s} are excluded from this inflation: they are already saturated against their directly measured targets, so inflating them would not relieve a stagnation that originates in the gauge-breaking channels.

The raw updated weights $\tilde{w}_k = w_k^{[4]}$ are then renormalised to the initial weight budget and constrained within the predefined bounds $[w_k^{\min}, w_k^{\max}]$, as specified by Eq. (6.12) of Section 6.3.