

ORLICZ SPACE RELAXATION OF TOTAL VARIATION DENOISING

Christian Clason*

Tobias Unterberger†

Abstract We consider variational image denoising with spatially dependent Orlicz regularization and introduce an Orlicz–Sobolev approximation of the ROF model based on a scaled $L \log L$ -type density. For a general class of uniformly superlinear integrands satisfying a Δ_2 -condition, we establish existence and uniqueness of minimizers and derive a Fenchel dual problem with dual attainment and pointwise optimality conditions. We then specialize the framework to a logarithmic density whose correction is activated above a prescribed local gradient scale. For this model, we obtain explicit expressions of the Fenchel dual and optimality conditions as well as a pointwise radial formula for the dual proximal map involving the Lambert W -function. As the logarithmic parameter tends to zero, we prove equicoercivity and Γ -convergence to the ROF functional, together with convergence of the corresponding minimizers. Numerical illustrations indicate that the logarithmic correction can reduce staircasing for diffuse transitions, while retaining behavior comparable to ROF for sharp interfaces and a natural image.

2020 Mathematics Subject Classification: 49J45, 46E30, 49M29

Keywords and phrases: generalized Orlicz spaces; image denoising; total variation; ROF model; Γ -convergence; Fenchel duality; primal–dual splitting.

1 INTRODUCTION

Total variation regularization is a standard tool in variational image processing, since it suppresses noise while allowing for discontinuities representing edges. Given a noisy image $f \in L^2(\Omega)$ and a regularization parameter $\alpha > 0$, the Rudin–Osher–Fatemi (ROF) model [17] consists in minimizing

$$\mathcal{F}_0(u) = \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \alpha |Du|(\Omega)$$

over $BV(\Omega) \cap L^2(\Omega)$. The one-homogeneous growth of the regularization term permits singular derivatives and thus avoids smoothing across jump discontinuities as strongly as quadratic Sobolev regularization. At the same time, total variation regularization tends to favour piecewise constant reconstructions and may create artificial staircasing in regions containing smooth or diffuse transitions [3, 13]. This has motivated, among other approaches, higher-order regularizers such as total generalized variation [2].

A different possibility, aimed at reducing staircasing in smooth or diffuse regions while retaining a first-order model, is to replace the linear growth of total variation by a suitably chosen non-standard growth condition. This leads naturally to the question of whether one can introduce a superlinear

This research was funded in whole or in part by the Austrian Science Fund (FWF) 10.55776/F100800.

*Department of Mathematics and Scientific Computing, University of Graz c.clason@uni-graz.at

†Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria, tobias.unterberger@tuwien.ac.at

correction that discourages the concentration of diffuse transitions into narrow regions of large gradient, while remaining sufficiently close to linear growth to recover total variation in a suitable limiting regime. Orlicz spaces provide a natural framework for such models, since their defining modulars allow growth laws beyond fixed powers; see, for instance, [15, 11]. Their generalized, spatially dependent counterparts additionally permit the regularization behavior to vary across the image. Spatially dependent first-order energies have previously been considered through variable-exponent models [5] and double-phase regularization [12]. More recently, a general framework for bounded variation spaces with generalized Orlicz growth was developed in [9].

Motivated by this question, we introduce and analyze a spatially scaled $L \log L$ -type regularization that provides a superlinear approximation of total variation. Let

$$\kappa \in L^\infty(\Omega), \quad 0 < \kappa_- \leq \kappa(x) \leq \kappa_+ < \infty \quad \text{for a.e. } x \in \Omega,$$

and, for $\gamma > 0$, define

$$\varphi_\gamma(x, t) := \alpha t \left(1 + \gamma \log^+ \frac{t}{\kappa(x)} \right), \quad t \geq 0,$$

where $\log^+ r := \max\{0, \log r\}$ for $r > 0$, with $\log^+ 0 := 0$. For the sake of presentation, we omit here the dependence on κ as this is not the focus of this work. We consider the corresponding functional

$$(1.1) \quad \mathcal{F}_\gamma(u) = \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \int_\Omega \varphi_\gamma(x, |\nabla u(x)|) \, dx.$$

The scale function κ determines the gradient magnitude at which the logarithmic correction becomes active. More precisely, $\varphi_\gamma(x, t) = \alpha t$ for $0 \leq t \leq \kappa(x)$, whereas gradients above this local scale receive an additional logarithmic penalty. Thus the model retains the total-variation density for smaller gradients while discouraging the concentration of a diffuse transition into narrow regions of large gradient. The parameter γ controls the strength of this correction.

For every fixed $\gamma > 0$, the density φ_γ has superlinear growth of $L \log L$ -type, uniformly in the spatial variable. The corresponding minimization problem is naturally posed in an Orlicz–Sobolev space. Functions with finite energy belong to $W^{1,1}(\Omega)$, so their distributional derivatives have no singular part. In particular, jump discontinuities are excluded for fixed $\gamma > 0$, whereas they are admissible in the ROF model. On the other hand,

$$\varphi_\gamma(x, t) \longrightarrow \alpha t \quad \text{as } \gamma \downarrow 0$$

locally uniformly in t and uniformly with respect to x . This suggests viewing the family (1.1) as an Orlicz–Sobolev approximation of the ROF functional, a connection made rigorous in Section 4.4 through Γ -convergence.

The functional-analytic treatment of this model requires some care. Although the scaled logarithmic density satisfies a uniform Δ_2 -condition, its conjugate has exponential growth and does not satisfy a corresponding Δ_2 -condition. In particular, the natural Orlicz spaces need not be reflexive, and the full dual Orlicz space must be distinguished from the finite-modular domain of the conjugate functional. While the general theory of convex integral functionals on Orlicz spaces is classical, see, e.g., [16, 14], these distinctions are essential for obtaining the correct dual problem for the present model.

We therefore first establish a convex-analytic framework for functionals of the form

$$\mathcal{F}(u) = \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \int_\Omega \varphi(x, |\nabla u(x)|) \, dx,$$

where the integrand $\varphi(x, t)$ depends measurably on the spatial variable x , has superlinear growth in t uniformly in x , and satisfies a Δ_2 -condition in t with constants independent of x . Using uniform

integrability and weak compactness in L^1 , we prove compactness of energy-bounded sequences and existence of a minimizer without relying on reflexivity of the underlying Orlicz space. The strict convexity of the fidelity term yields uniqueness. We further identify the dual space and the Fenchel conjugate of the modular, including the case in which the conjugate integrand fails the Δ_2 -condition. This leads to dual attainment, an explicit Fenchel dual problem, and a pointwise optimality system in which the homogeneous normal boundary condition is encoded in a weak integration-by-parts identity.

We then specialize these results to the scaled logarithmic density. Its scalar conjugate and the relevant subdifferentials can be computed explicitly, which gives a concrete dual formulation and optimality system. We also identify the formal primal–dual splitting structure associated with the model. As in ROF denoising, the proximal map of the fidelity term is explicit and the dual proximal map acts pointwise and radially. In the scaled logarithmic case, its radial profile admits an explicit piecewise formula involving the principal branch of the Lambert W -function. The model therefore retains the basic computational structure of the primal–dual method of [4], despite the additional nonlinear growth.

The relation to total variation is made rigorous by proving that $\mathcal{F}_\gamma$ is equicoercive and Γ -converges to the ROF functional $\mathcal{F}_0$ with respect to strong convergence in $L^1(\Omega)$ as $\gamma \downarrow 0$. Since both the relaxed and limiting problems have unique minimizers, this also yields convergence of the minimizers to the ROF solution. Finally, numerical experiments explore the finite- γ behavior of the regularizer by comparing the scaled logarithmic model with ROF and quadratic H^1 regularization. A blurred-disk experiment indicates that the logarithmic correction can reduce staircasing for genuinely diffuse transitions, while the sharp-interface and natural-image examples yield results comparable to ROF.

Overview. Section 2 collects the required background on uniform integrability, Fenchel duality, Orlicz spaces, and the ROF functional. In Section 3, we develop the general convex-analytic framework for spatially dependent Orlicz regularization, including existence and uniqueness, the conjugate modular, Fenchel duality, and the associated optimality conditions. Section 4 introduces the scaled logarithmic model and derives its basic properties, explicit dual formulation, optimality system, and formal primal–dual splitting structure. The final part of that section proves the Γ -convergence to the ROF functional as $\gamma \downarrow 0$. The numerical realization and the blurred-disk and natural-image illustrations are presented in Section 5. The paper concludes with a summary of the analytical and numerical results.

2 PRELIMINARIES

2.1 NOTATION AND CONVENTIONS

We write $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$. Unless stated otherwise, $\Omega \subset \mathbb{R}^N$ is a bounded Lipschitz domain. In the imaging applications motivating the paper, $N = 2$. We denote by $\partial\Omega$ the boundary of Ω and by ν its outer unit normal, which is defined almost everywhere on $\partial\Omega$.

The N -dimensional Lebesgue measure is denoted by $\mathcal{L}^N$, and we abbreviate

$$|E| := \mathcal{L}^N(E)$$

for every Lebesgue-measurable set $E \subset \mathbb{R}^N$. Unless another measure is specified, all integrals are taken with respect to $\mathcal{L}^N$. Inequalities between measurable functions are understood almost everywhere unless explicitly stated otherwise. We write χ_E for the characteristic function of a measurable set E and $\mu \ll \nu$ for absolute continuity of the measure μ with respect to the measure ν .

For $m \in \mathbb{N}$, the space $L^0(\Omega; \mathbb{R}^m)$ consists of equivalence classes of Lebesgue-measurable functions from Ω to $\mathbb{R}^m$, where functions agreeing almost everywhere are identified. In the scalar-valued case, we omit the target space and write $L^0(\Omega)$. The same convention is used for the spaces $L^p(\Omega; \mathbb{R}^m)$,

$W^{1,p}(\Omega; \mathbb{R}^m)$, and their Orlicz counterparts. We denote by $C_c^\infty(\Omega)$ the space of smooth functions with compact support in Ω .

The symbol $|\cdot|$ denotes the absolute value on $\mathbb{R}$ and the Euclidean norm on $\mathbb{R}^m$, as determined by the context, while $z \cdot \xi$ denotes the Euclidean inner product. If X is a normed space, its norm is denoted by $\|\cdot\|_X$; the subscript is occasionally omitted when the underlying space is clear. We write

$$u_n \rightarrow u \quad \text{in } X$$

for strong convergence in X and

$$u_n \rightharpoonup u \quad \text{in } X$$

for weak convergence. The symbol $X \hookrightarrow Y$ denotes a continuous embedding. Whenever a subsequence is extracted, it is not relabelled unless this is necessary for clarity.

For $u \in W^{1,1}(\Omega)$, the symbol ∇u denotes its weak gradient. For $u \in BV(\Omega)$, we denote its distributional derivative by Du and its total variation measure by $|Du|$. In particular, $|Du|(\Omega)$ denotes the total variation of u on Ω . If $u \in W^{1,1}(\Omega)$, then

$$Du = \nabla u \mathcal{L}^N, \quad |Du|(\Omega) = \int_{\Omega} |\nabla u| dx.$$

The divergence of a vector field p is denoted by $\operatorname{div} p$ and is understood in the distributional sense unless additional regularity is specified.

For a Banach space X , we write X^* for its continuous dual and $\langle x^*, x \rangle_X$ for the duality pairing between $x^* \in X^*$ and $x \in X$. If $A : X \rightarrow Y$ is continuous and linear, then $A^* : Y^* \rightarrow X^*$ denotes its adjoint. For a set S in a topological space, S° denotes its interior. If $H : X \rightarrow \mathbb{R}_\infty$ is an extended-real-valued functional, then

$$\operatorname{dom} H := \{x \in X : H(x) < +\infty\}$$

is its effective domain. A superscript $*$ applied to a functional denotes its Fenchel conjugate, whereas a superscript $*$ applied to a Banach space denotes its continuous dual. The convex subdifferential of H at x is denoted by $\partial H(x)$. Subscripts such as $\partial_t \varphi$ and $\partial_z \Phi$ indicate the variable with respect to which the convex subdifferential is taken.

If X is a Hilbert space, $H : X \rightarrow \mathbb{R}_\infty$ is proper, convex, and lower semicontinuous, and $\tau > 0$, we use the convention

$$\operatorname{prox}_{\tau H}(w) := \arg \min_{v \in X} \left\{ \frac{1}{2\tau} \|v - w\|_X^2 + H(v) \right\} = (I + \tau \partial H)^{-1}(w).$$

For an Orlicz integrand φ , the symbol ρ_φ denotes the associated modular, L^φ the corresponding Orlicz space, and $\|\cdot\|_\varphi$ its Luxemburg norm. Precise definitions are recalled below. The pointwise Fenchel conjugate of the scalar integrand $t \mapsto \varphi(x, t)$ is denoted by $\varphi^*(x, \cdot)$. For the associated radial integrand

$$\Phi(x, z) := \varphi(x, |z|),$$

we write $\Phi^*(x, \cdot)$ for the corresponding vectorial conjugate. We emphasize the distinction between

$$\rho_\varphi^* \quad \text{and} \quad \rho_{\varphi^*} :$$

the former is the Fenchel conjugate of the functional ρ_φ , while the latter is the modular generated by the pointwise conjugate integrand φ^* .

Finally, for $r > 0$, we set

$$\log^+ r := \max\{0, \log r\}, \quad \log^+ 0 := 0.$$

For discrete images $u, v \in \mathbb{R}^{n \times n}$, we define the mean squared error by

$$\text{MSE}(u, v) := \frac{1}{n^2} \sum_{i,j=1}^n |u_{ij} - v_{ij}|^2.$$

The peak signal-to-noise ratio is

$$\text{PSNR}(u, v) := 10 \log_{10} \left(\frac{L^2}{\text{MSE}(u, v)} \right),$$

where L denotes the maximum possible image intensity. All images used in the numerical experiments are normalized to $[0, 1]$, and hence $L = 1$. PSNR values are reported in decibels (dB).

2.2 UNIFORM INTEGRABILITY AND DE LA VALLÉE POUSSIN'S THEOREM

We recall the notion of uniform integrability following Rao–Ren [15].

Definition 2.1 (uniform integrability). A sequence $\{f_n\}_{n \in \mathbb{N}} \subset L^0(\Omega)$ is called *uniformly integrable* if

$$\lim_{|A| \rightarrow 0} \int_A |f_n| \, dx = 0,$$

uniformly in $n \in \mathbb{N}$.

We also use the following characterization of uniform integrability; see [15, Theorem 2].

Theorem 2.2 (de La Vallée Poussin). Let $\{f_n\}_{n \in \mathbb{N}} \subset L^0(\Omega)$. Then the following are equivalent

- (i) $\{f_n\}_{n \in \mathbb{N}}$ is uniformly integrable.
- (ii) $\lim_{\lambda \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\{|f_n| > \lambda\}} |f_n| \, dx = 0$.
- (iii) There exists a convex function $\Phi : \mathbb{R} \rightarrow [0, \infty)$ such that $\Phi(0) = 0$, $\Phi(-x) = \Phi(x)$, and $\frac{\Phi(x)}{x} \rightarrow \infty$ as $x \rightarrow \infty$, in terms of which

$$C = \sup_{n \in \mathbb{N}} \int_{\Omega} \Phi(f_n) \, dx < \infty.$$

Finally, we recall the Dunford–Pettis theorem in the form stated in [10, Theorem 1.43].

Theorem 2.3 (Dunford–Pettis). Let $\{f_n\}_{n \in \mathbb{N}} \subset L^1(\Omega)$ satisfy

$$\sup_{n \in \mathbb{N}} \int_{\Omega} |f_n| \, dx < \infty$$

and suppose that

$$\lim_{\lambda \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\{|f_n| > \lambda\}} |f_n| \, dx = 0.$$

Then there exist a subsequence $\{f_{n_k}\}_{k \in \mathbb{N}}$ and $f \in L^1(\Omega)$ such that

$$f_{n_k} \rightharpoonup f \quad \text{in } L^1(\Omega).$$

2.3 CONVEX SUBDIFFERENTIALS AND FENCHEL DUALITY

For this section let X and Y be Banach spaces, let $F : X \rightarrow \mathbb{R}_\infty$ and $G : Y \rightarrow \mathbb{R}_\infty$ be proper, convex, and lower semicontinuous, and let $A : X \rightarrow Y$ be continuous and linear.

Definition 2.4 (convex subdifferential). The subdifferential of F at $x \in \text{dom } F$ is

$$\partial F(x) := \{x^* \in X^* \mid \langle x^*, \tilde{x} - x \rangle_X \leq F(\tilde{x}) - F(x) \text{ for all } \tilde{x} \in X\}.$$

Definition 2.5 (Fenchel conjugate). The Fenchel conjugate of $F : X \rightarrow \mathbb{R}_\infty$ is given by

$$F^* : X^* \rightarrow \mathbb{R}_\infty, \quad F^*(x^*) = \sup_{x \in X} \langle x^*, x \rangle_X - F(x).$$

Definition 2.6 (primal and dual problem). The primal problem is

$$(P) \quad \inf_{x \in X} (F(x) + G(Ax))$$

with the corresponding dual problem

$$(D) \quad \sup_{y^* \in Y^*} (-F^*(-A^*y^*) - G^*(y^*))$$

Theorem 2.7 (Fenchel–Rockafellar). Assume that

- (i) the primal problem (P) admits a solution $\bar{x} \in X$;
- (ii) there exists $x_0 \in \text{dom } F \cap \text{dom}(G \circ A)$ with $Ax_0 \in (\text{dom } G)^\circ$.

Then, the dual problem (D) admits a solution $\bar{y}^* \in Y^*$ and

$$\min_{x \in X} (F(x) + G(Ax)) = \max_{y^* \in Y^*} (-F^*(-A^*y^*) - G^*(y^*)).$$

Furthermore, $\bar{x}$ and $\bar{y}^*$ are solutions to (P) and (D) respectively, if and only if

$$(OC) \quad \begin{cases} -A^*\bar{y}^* \in \partial F(\bar{x}), \\ \bar{y}^* \in \partial G(A\bar{x}). \end{cases}$$

2.4 ORLICZ SPACES

We follow [11, Chapter 3] for ordinary Orlicz spaces, where φ depends only on the scalar variable. Section 3 uses an x -dependent modular, but only the basic growth, modular, and duality notions below are needed in the sequel.

Definition 2.8 (growth conditions). Let $f : (0, \infty) \rightarrow \mathbb{R}$ and $p, q > 0$. We say that f satisfies

(Inc) $_p$ if $\frac{f(t)}{t^p}$ is increasing;

(aInc) $_p$ if $\frac{f(t)}{t^p}$ is almost increasing;

(Dec) $_q$ if $\frac{f(t)}{t^q}$ is decreasing;

(aDec) $_q$ if $\frac{f(t)}{t^q}$ is almost decreasing;

We say that f satisfies (aInc), (Inc), (aDec) or (Dec) if there exist $p > 1$ or $q < \infty$ such that f satisfies (aInc) $_p$, (Inc) $_p$, (aDec) $_q$ or (Dec) $_q$, respectively.

Definition 2.9 (Δ_2 -growth condition). We say that a function $\varphi : [0, \infty) \rightarrow [0, \infty]$ satisfies Δ_2 , if there exists a constant $K \geq 2$ such that

$$\varphi(2t) \leq K\varphi(t) \quad \text{for all } t \geq 0.$$

Definition 2.10 (Φ -function). Let $\varphi : [0, \infty) \rightarrow [0, \infty]$ be increasing with $\varphi(0) = 0$, $\lim_{t \rightarrow 0^+} \varphi(t) = 0$ and $\lim_{t \rightarrow \infty} \varphi(t) = \infty$. We say that φ is a

- (i) (weak) Φ -function if it satisfies (aInc)₁ on $(0, \infty)$;
- (ii) convex Φ -function if it is left-continuous and convex;
- (iii) strong Φ -function if it is continuous in the topology of $[0, \infty]$ and convex.

The sets of weak, convex and strong Φ -functions are denoted by Φ_w , Φ_c and Φ_s , respectively.

Definition 2.11 (Orlicz space). Let $\varphi \in \Phi_w$. The modular associated with φ is

$$\rho_\varphi(f) := \int_\Omega \varphi(|f(x)|) dx.$$

The Orlicz space $L^\varphi(\Omega)$ is defined by

$$L^\varphi(\Omega) := \{f \in L^0(\Omega) \mid \rho_\varphi(\lambda f) < \infty \text{ for some } \lambda > 0\}.$$

Definition 2.12 (Luxemburg norm). For $f \in L^\varphi(\Omega)$, define

$$\|f\|_\varphi := \inf \left\{ \lambda > 0 \mid \rho_\varphi\left(\frac{f}{\lambda}\right) \leq 1 \right\}.$$

Remark 2.13. If φ satisfies the Δ_2 -condition, then

$$L^\varphi(\Omega) = \{f \in L^0(\Omega) \mid \rho_\varphi(f) < \infty\}.$$

Thus, under the standing Δ_2 -assumption used in Section 3, the modular class and the Luxemburg-norm Orlicz space coincide.

Definition 2.14 (Orlicz–Sobolev space). Let $\varphi \in \Phi_w$. The function $u \in L^\varphi(\Omega) \cap L^1_{loc}(\Omega)$ belongs to the Sobolev space $W^{1,\varphi}(\Omega)$ if its weak partial derivatives $\partial_i u$ exist and belong to $L^\varphi(\Omega)$ for all $i \in \{1, \dots, N\}$. We set

$$\rho_{W^{1,\varphi}(\Omega)}(u) := \rho_\varphi(u) + \sum_{i=1}^N \rho_\varphi(\partial_i u),$$

which induces a (quasi-)norm by

$$\|u\|_{W^{1,\varphi}(\Omega)} := \inf \left\{ \lambda > 0 \mid \rho_{W^{1,\varphi}(\Omega)}\left(\frac{u}{\lambda}\right) \leq 1 \right\}.$$

With this setup, we record some useful results for us from [11]:

- Hölder's inequality:

$$\int_\Omega |f||g| dx \leq 2\|f\|_\varphi \|g\|_{\varphi^*}$$

for all $f \in L^\varphi(\Omega)$ and $g \in L^{\varphi^*}(\Omega)$.

- If $\varphi \in \Phi_c$, then $L^\varphi(\Omega)$ and $W^{1,\varphi}(\Omega)$ are Banach spaces.
- If φ satisfies Δ_2 , then boundedness of $\rho_\varphi(f_n)$ implies boundedness of $\|f_n\|_\varphi$.
- If $\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t} = \infty$, then $L^\varphi(\Omega) \hookrightarrow L^1(\Omega)$ on bounded domains, and hence $W^{1,\varphi}(\Omega) \hookrightarrow W^{1,1}(\Omega)$.

2.5 THE ROF FUNCTIONAL

The Rudin–Osher–Fatemi (ROF) model is the standard variational model for total variation denoising [17] and is given by

$$\mathcal{F}_0(u) := \begin{cases} \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \alpha |Du|(\Omega), & u \in BV(\Omega) \cap L^2(\Omega), \\ +\infty, & u \in L^1(\Omega) \setminus (BV(\Omega) \cap L^2(\Omega)). \end{cases}$$

Its quadratic fidelity term enforces closeness to the noisy datum f , while the total variation term penalizes oscillations without smoothing across jump discontinuities as strongly as Sobolev-type regularizers. The natural domain is $BV(\Omega) \cap L^2(\Omega)$: for $u \in BV(\Omega)$, the distributional derivative Du is a finite Radon measure and need not be represented by an L^1 -function. We denote by $BV(\Omega)$ the space of functions of bounded variation and by $|Du|(\Omega)$ the total variation of Du ; for compactness and lower semicontinuity properties of BV we use [1, Chapter 3].

3 RESULTS IN CONVEX OPTIMIZATION FOR FUNCTIONALS ON ORLICZ SPACES

From now on let

$$\varphi : \Omega \times [0, \infty) \rightarrow [0, \infty)$$

be such that $x \mapsto \varphi(x, t)$ is measurable for every $t \geq 0$, and $t \mapsto \varphi(x, t)$ is convex, increasing, continuous, and satisfies $\varphi(x, 0) = 0$ for almost every $x \in \Omega$. We also assume that $\varphi(\cdot, t) \in L^1(\Omega)$ for every $t \geq 0$. As for growth we assume that the uniform Δ_2 -condition

$$\varphi(x, 2t) \leq C_\Delta \varphi(x, t) \quad \text{for a.e. } x \in \Omega, \ t \geq 0,$$

and the uniform superlinear growth condition

$$\lim_{t \rightarrow \infty} \operatorname{ess\,inf}_{x \in \Omega} \frac{\varphi(x, t)}{t} = \infty,$$

are both satisfied. For $z \in \mathbb{R}^N$, set

$$\Phi(x, z) := \varphi(x, |z|),$$

which is convex and lower semicontinuous for almost every x . We define the modular

$$\rho_\varphi(v) := \int_\Omega \varphi(x, |v(x)|) \, dx = \int_\Omega \Phi(x, v(x)) \, dx$$

for measurable vector fields $v : \Omega \rightarrow \mathbb{R}^N$, and write

$$L^\varphi(\Omega; \mathbb{R}^N) := \{v \in L^0(\Omega; \mathbb{R}^N) \mid \rho_\varphi(\lambda v) < \infty \text{ for some } \lambda > 0\}.$$

Since the standing assumptions include the uniform Δ_2 -condition, this space coincides with the finite-modular class defined by the condition

$$\rho_\varphi(v) < \infty.$$

We consider the standard Luxemburg norm

$$\|v\|_\varphi := \inf \left\{ \lambda > 0 \mid \rho_\varphi\left(\frac{v}{\lambda}\right) \leq 1 \right\}.$$

For $f \in L^2(\Omega)$, define

$$(3.1) \quad \mathcal{F} : X_\varphi \rightarrow \mathbb{R}, \quad \mathcal{F}(u) := \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \rho_\varphi(\nabla u).$$

The natural space for admissible functions is

$$X_\varphi := \{u \in L^2(\Omega) \cap W^{1,1}(\Omega) \mid \rho_\varphi(\nabla u) < \infty\},$$

equipped with the norm

$$\|u\|_{X_\varphi} := \|u\|_{L^2(\Omega)} + \|\nabla u\|_\varphi.$$

Since $L^\varphi(\Omega; \mathbb{R}^N)$ is a Banach space and the weak gradient is closed, X_φ equipped with this norm is a Banach space. Moreover, the operator

$$A : X_\varphi \rightarrow L^\varphi(\Omega; \mathbb{R}^N), \quad Au := \nabla u,$$

is continuous.

3.1 EXISTENCE OF MINIMIZERS

The following compactness result is the basic tool for the direct method.

Lemma 3.1 (compactness). *Let $\{u_n\}_n \subset X_\varphi$ satisfy*

$$\sup_{n \in \mathbb{N}} \mathcal{F}(u_n) \leq C.$$

Then there exist a subsequence, not relabelled, and $u \in L^2(\Omega) \cap W^{1,1}(\Omega)$ such that

$$\begin{aligned} u_n &\rightarrow u && \text{strongly in } L^1(\Omega), \\ u_n &\rightharpoonup u && \text{weakly in } L^2(\Omega), \end{aligned}$$

and

$$\nabla u_n \rightharpoonup \nabla u \quad \text{weakly in } L^1(\Omega; \mathbb{R}^N).$$

Proof. By the uniform superlinear growth condition, for every $M > 0$ there exists $t_M > 0$ such that

$$\varphi(x, t) \geq Mt \quad \text{for a.e. } x \in \Omega \text{ and all } t \geq t_M.$$

Therefore, for every n ,

$$\begin{aligned} \int_\Omega |\nabla u_n| dx &= \int_{\{|\nabla u_n| < t_M\}} |\nabla u_n| dx + \int_{\{|\nabla u_n| \geq t_M\}} |\nabla u_n| dx \\ &\leq t_M |\Omega| + \frac{1}{M} \int_\Omega \varphi(x, |\nabla u_n|) dx. \end{aligned}$$

Taking $M = 1$ and using the assumed modular bound gives

$$\sup_n \|\nabla u_n\|_{L^1(\Omega; \mathbb{R}^N)} < \infty.$$

Since $f \in L^2(\Omega)$, $\{u_n\}_n$ is bounded in $L^2(\Omega)$ and since Ω is bounded, it is also bounded in $L^1(\Omega)$. Hence $\{u_n\}_n$ is bounded in $W^{1,1}(\Omega)$.

By Rellich–Kondrachov, after passing to a subsequence,

$$u_n \rightarrow u \quad \text{strongly in } L^1(\Omega)$$

for some $u \in L^1(\Omega)$. The L^2 -bound gives, after passing to a further subsequence,

$$u_n \rightharpoonup \tilde{u} \quad \text{weakly in } L^2(\Omega).$$

The weak L^2 -limit and the strong L^1 -limit agree by uniqueness of distributional limits, so $\tilde{u} = u$. Thus

$$u_n \rightharpoonup u \quad \text{weakly in } L^2(\Omega).$$

It remains to prove weak L^1 -compactness of the gradients. We first show uniform integrability. Let $M > 0$, and choose $t_M > 0$ as above. For every $\lambda \geq t_M$,

$$\begin{aligned} \int_{\{|\nabla u_n| > \lambda\}} |\nabla u_n| \, dx &\leq \frac{1}{M} \int_{\{|\nabla u_n| > \lambda\}} \varphi(x, |\nabla u_n|) \, dx \\ &\leq \frac{1}{M} \int_{\Omega} \varphi(x, |\nabla u_n|) \, dx \leq \frac{C}{M}. \end{aligned}$$

Taking the supremum over n and then letting $\lambda \rightarrow \infty$ gives

$$\limsup_{\lambda \rightarrow \infty} \sup_n \int_{\{|\nabla u_n| > \lambda\}} |\nabla u_n| \, dx \leq \frac{C}{M}.$$

Since $M > 0$ was arbitrary,

$$\lim_{\lambda \rightarrow \infty} \sup_n \int_{\{|\nabla u_n| > \lambda\}} |\nabla u_n| \, dx = 0.$$

By [Theorem 2.2](#), equivalently by the tail characterization of uniform integrability, the family $\{|\nabla u_n|\}_n$ is uniformly integrable.

Since $|\partial_i u_n| \leq |\nabla u_n|$ for each component i , every family $\{\partial_i u_n\}_n$ is uniformly integrable. The previous L^1 -bound also gives

$$\sup_n \|\partial_i u_n\|_{L^1(\Omega)} < \infty.$$

By Dunford–Pettis, for each i there is a subsequence and a function $g_i \in L^1(\Omega)$ such that

$$\partial_i u_n \rightharpoonup g_i \quad \text{weakly in } L^1(\Omega).$$

Since there are only finitely many components, a diagonal extraction gives $g = (g_1, \dots, g_N) \in L^1(\Omega; \mathbb{R}^N)$ such that

$$\nabla u_n \rightharpoonup g \quad \text{weakly in } L^1(\Omega; \mathbb{R}^N).$$

Finally, for every $\eta \in C_c^\infty(\Omega)$ and every $i \in \{1, \dots, N\}$, the strong L^1 -convergence of u_n and the weak L^1 -convergence of $\partial_i u_n$ give

$$\int_{\Omega} u \, \partial_i \eta \, dx = \lim_{n \rightarrow \infty} \int_{\Omega} u_n \, \partial_i \eta \, dx = - \lim_{n \rightarrow \infty} \int_{\Omega} \partial_i u_n \, \eta \, dx = - \int_{\Omega} g_i \eta \, dx.$$

Hence $g_i = \partial_i u$ in the sense of distributions. Since $g_i \in L^1(\Omega)$ for every i , we have $u \in W^{1,1}(\Omega)$ and $\nabla u = g$. Therefore

$$\nabla u_n \rightharpoonup \nabla u \quad \text{weakly in } L^1(\Omega; \mathbb{R}^N),$$

which proves the claim. $\square$

Remark 3.2 (lower semicontinuity). The map

$$u \mapsto \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2$$

is lower semicontinuous with respect to weak convergence in $L^2(\Omega)$. For the modular part, first note that by the lower semicontinuity of $z \mapsto \Phi(x, z)$ and Fatou's lemma, the modular

$$v \mapsto \rho_\varphi(v) = \int_{\Omega} \Phi(x, v(x)) \, dx$$

is lower semicontinuous with respect to a.e. convergence along subsequences; see [11, Lemma 3.1.4]. Consequently, if $v_n \rightarrow v$ strongly in $L^1(\Omega; \mathbb{R}^N)$, then after passing to a subsequence realizing the liminf, $v_n(x) \rightarrow v(x)$ for a.e. x , and hence

$$\rho_\varphi(v) \leq \liminf_{n \rightarrow \infty} \rho_\varphi(v_n).$$

Thus ρ_φ is strongly lower semicontinuous on $L^1(\Omega; \mathbb{R}^N)$. Since ρ_φ is convex, its epigraph is convex and strongly closed, hence weakly closed. Therefore ρ_φ is weakly lower semicontinuous on $L^1(\Omega; \mathbb{R}^N)$.

Proposition 3.3 (existence of minimizers). *There exists a unique minimizer of $\mathcal{F}$ over X_φ .*

Proof. The direct method applies. Let $\{u_n\}_n \subset X_\varphi$ be a minimizing sequence. Since $\mathcal{F}$ is bounded from below, we may assume $\sup_n \mathcal{F}(u_n) \leq C$, and hence

$$\sup_n \|u_n - f\|_{L^2(\Omega)}^2 \leq C, \quad \sup_n \rho_\varphi(\nabla u_n) \leq C.$$

Thus $\{u_n\}_n$ is bounded in $L^2(\Omega)$, and Lemma 3.1 yields a subsequence and $\bar{u} \in L^2(\Omega) \cap W^{1,1}(\Omega)$ such that

$$u_n \rightharpoonup \bar{u} \quad \text{in } L^2(\Omega), \quad \nabla u_n \rightharpoonup \nabla \bar{u} \quad \text{in } L^1(\Omega; \mathbb{R}^N).$$

By Remark 3.2,

$$\rho_\varphi(\nabla \bar{u}) \leq \liminf_{n \rightarrow \infty} \rho_\varphi(\nabla u_n) < \infty,$$

so $\bar{u} \in X_\varphi$, and again by lower semicontinuity,

$$\mathcal{F}(\bar{u}) \leq \liminf_{n \rightarrow \infty} \mathcal{F}(u_n) = \inf_{u \in X_\varphi} \mathcal{F}(u).$$

Uniqueness follows from the strict convexity of the fidelity term on $L^2(\Omega)$ and the convexity of the integral term. $\square$

3.2 DUALITY IN ORLICZ SPACES

Definition 3.4 (pointwise conjugate integrand). For almost every $x \in \Omega$, define

$$\varphi^*(x, s) := \sup_{t \geq 0} \{st - \varphi(x, t)\}, \quad s \geq 0.$$

Equivalently, for $\Phi(x, z) = \varphi(x, |z|)$, define

$$\Phi^*(x, \xi) := \sup_{z \in \mathbb{R}^N} \{\xi \cdot z - \Phi(x, z)\}.$$

Notice that

$$\Phi^*(x, \xi) = \varphi^*(x, |\xi|).$$

since

$$\sup_{z \in \mathbb{R}^N} \{\xi \cdot z - \varphi(x, |z|)\} = \sup_{t \geq 0} \{t|\xi| - \varphi(x, t)\}.$$

Remark 3.5 (Young's inequality). The pointwise conjugate gives

$$st \leq \varphi(x, t) + \varphi^*(x, s) \quad \text{for a.e. } x \in \Omega \text{ and all } s, t \geq 0.$$

We write

$$\rho_{\varphi^*}(p) := \int_{\Omega} \varphi^*(x, |p(x)|) dx, \quad L^{\varphi^*}(\Omega; \mathbb{R}^N) := \{p \in L^0(\Omega; \mathbb{R}^N) \mid \rho_{\varphi^*}(\lambda p) < \infty \text{ for some } \lambda > 0\},$$

with the corresponding Luxemburg norm $\|p\|_{\varphi^*}$. Since φ^* need not satisfy a Δ_2 -condition, the Orlicz space L^{φ^*} must be distinguished from the finite-modular class

$$\text{dom } \rho_{\varphi^*} := \{p \in L^{\varphi^*}(\Omega; \mathbb{R}^N) \mid \rho_{\varphi^*}(p) < \infty\}.$$

The conjugate functional ρ_{φ}^* is finite precisely on $\text{dom } \rho_{\varphi^*}$ and equals $+\infty$ otherwise.

Lemma 3.6 (duality pairing and dual identification). *Assume that φ satisfies the standing assumptions of Section 3. Then the associate space of $L^{\varphi}(\Omega; \mathbb{R}^N)$, in the sense of [11, Definition 3.4.5], is $L^{\varphi^*}(\Omega; \mathbb{R}^N)$, with duality pairing*

$$\langle p, v \rangle := \int_{\Omega} p \cdot v dx.$$

Since φ satisfies the uniform Δ_2 -condition, $L^{\varphi}(\Omega; \mathbb{R}^N)$ has order-continuous norm. Consequently, the norm dual can be identified with its associate space, and hence

$$(L^{\varphi}(\Omega; \mathbb{R}^N))^* \simeq L^{\varphi^*}(\Omega; \mathbb{R}^N)$$

through the above pairing. In particular,

$$\int_{\Omega} |p \cdot v| dx \leq 2\|p\|_{\varphi^*}\|v\|_{\varphi}$$

for all $v \in L^{\varphi}(\Omega; \mathbb{R}^N)$ and $p \in L^{\varphi^*}(\Omega; \mathbb{R}^N)$.

Proof. By [11, Definition 3.4.5 and Theorem 3.4.6], the associate space of $L^{\varphi}(\Omega; \mathbb{R}^N)$ is $L^{\varphi^*}(\Omega; \mathbb{R}^N)$, with comparable norms. This gives the integral pairing and the Hölder inequality.

We first show that the Luxemburg norm is order continuous under the present assumptions. Let

$$0 \leq |v_k| \leq |v|, \quad v_k \rightarrow 0 \quad \text{a.e.},$$

with $v \in L^{\varphi}(\Omega; \mathbb{R}^N)$. Since $v \in L^{\varphi}$, there exists $\lambda_0 > 0$ such that

$$\rho_{\varphi}(\lambda_0 v) < \infty.$$

By the uniform Δ_2 -condition, this implies

$$\rho_{\varphi}(\lambda v) < \infty \quad \text{for every } \lambda > 0.$$

For fixed $\lambda > 0$, we have

$$0 \leq \varphi(x, \lambda |v_k(x)|) \leq \varphi(x, \lambda |v(x)|) \in L^1(\Omega),$$

and the integrand converges pointwise to zero. By dominated convergence,

$$\rho_{\varphi}(\lambda v_k) \rightarrow 0.$$

Taking $\lambda = 1/\varepsilon$, we find that for all sufficiently large k ,

$$\rho_{\varphi}(v_k/\varepsilon) \leq 1.$$

Thus $\|v_k\|_\varphi \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\|v_k\|_\varphi \rightarrow 0$. This proves order continuity.

It remains to exclude singular functionals. Let $\ell \in (L^\varphi(\Omega; \mathbb{R}^N))^*$. For $i = 1, \dots, N$, define

$$v_i(E) := \ell(\chi_E e_i)$$

for measurable $E \subset \Omega$. This is well-defined because $\varphi(\cdot, 1) \in L^1(\Omega)$, and therefore $\chi_E e_i \in L^\varphi(\Omega; \mathbb{R}^N)$. The linearity of ℓ implies finite additivity.

We next prove countable additivity. Let $(E_j)_{j \in \mathbb{N}}$ be pairwise disjoint and set

$$E := \bigcup_{j=1}^{\infty} E_j, \quad E^{(m)} := \bigcup_{j=1}^m E_j.$$

Then

$$\chi_{E^{(m)}} e_i \rightarrow \chi_E e_i \quad \text{a.e.}$$

and

$$|\chi_E e_i - \chi_{E^{(m)}} e_i| = \chi_{E \setminus E^{(m)}}.$$

Since $E \setminus E^{(m)} \downarrow \emptyset$, order continuity of the Luxemburg norm gives

$$\|\chi_E e_i - \chi_{E^{(m)}} e_i\|_\varphi \rightarrow 0.$$

By the continuity of ℓ ,

$$v_i(E) - v_i(E^{(m)}) = \ell((\chi_E - \chi_{E^{(m)}}) e_i) \rightarrow 0.$$

Using finite additivity for $E^{(m)}$, we obtain

$$v_i(E) = \lim_{m \rightarrow \infty} \sum_{j=1}^m v_i(E_j) = \sum_{j=1}^{\infty} v_i(E_j).$$

Thus v_i is countably additive.

Moreover, if $|E| = 0$, then $\chi_E e_i = 0$ in $L^\varphi(\Omega; \mathbb{R}^N)$, and hence $v_i(E) = 0$. Therefore $v_i \ll \mathcal{L}^N$.

Finally, we show finite variation explicitly. For every finite measurable partition $(E_j)_{j=1}^m$ of Ω , choose signs $\varepsilon_j \in \{-1, 1\}$ such that

$$\varepsilon_j v_i(E_j) = |v_i(E_j)|.$$

Then

$$\sum_{j=1}^m |v_i(E_j)| = \ell \left(\sum_{j=1}^m \varepsilon_j \chi_{E_j} e_i \right) \leq \|\ell\| \left\| \sum_{j=1}^m \varepsilon_j \chi_{E_j} e_i \right\|_\varphi.$$

Since

$$\left\| \sum_{j=1}^m \varepsilon_j \chi_{E_j} e_i \right\|_\varphi = \|\chi_\Omega e_i\|_\varphi,$$

the lattice property of the Luxemburg norm gives

$$\left\| \sum_{j=1}^m \varepsilon_j \chi_{E_j} e_i \right\|_\varphi = \|\chi_\Omega e_i\|_\varphi < \infty.$$

Hence

$$|v_i|(\Omega) \leq \|\ell\| \|\chi_\Omega e_i\|_\varphi < \infty.$$

Thus ν_i is a finite signed measure. Since $\nu_i \ll \mathcal{L}^N$, the Radon–Nikodym theorem gives $p_i \in L^1(\Omega)$ such that

$$\nu_i(E) = \int_E p_i \, dx$$

for every measurable $E \subset \Omega$.

Thus, for $p = (p_1, \dots, p_N)$,

$$\ell(v) = \int_{\Omega} p \cdot v \, dx$$

for all simple vector fields v . By [11, Theorem 3.5.1], simple functions in L^φ are dense in L^φ . Therefore the representation extends to all $v \in L^\varphi(\Omega; \mathbb{R}^N)$. Since ℓ is bounded, p belongs to the associate space of $L^\varphi(\Omega; \mathbb{R}^N)$, and hence, by [11, Theorem 3.4.6],

$$p \in L^{\varphi^*}(\Omega; \mathbb{R}^N).$$

This proves the norm-dual identification. $\square$

By definition, the Fenchel conjugate of ρ_φ on the associated space is

$$\rho_\varphi^*(p) := \sup_{v \in L^\varphi(\Omega; \mathbb{R}^N)} \left\{ \int_{\Omega} p \cdot v \, dx - \rho_\varphi(v) \right\}.$$

Just as in Lebesgue spaces, this conjugate can be expressed pointwise almost everywhere via the conjugate of the modular.

Proposition 3.7 (Fenchel conjugate of the modular). *For $p \in L^{\varphi^*}(\Omega; \mathbb{R}^N)$,*

$$\rho_\varphi^*(p) = \int_{\Omega} \Phi^*(x, p(x)) \, dx = \int_{\Omega} \varphi^*(x, |p(x)|) \, dx.$$

In particular, $\rho_\varphi^*(p) < \infty$ if and only if $p \in \text{dom } \rho_\varphi^*$.

Proof. The following argument is a direct radial specialization of the standard conjugacy formula for convex integral functionals with normal integrands in [16, Theorem 2].

We first compute the pointwise conjugate. For a.e. $x \in \Omega$ and every $\xi \in \mathbb{R}^N$, we have

$$\Phi^*(x, \xi) = \sup_{z \in \mathbb{R}^N} \{ \xi \cdot z - \varphi(x, |z|) \} = \sup_{t \geq 0} \{ |\xi| t - \varphi(x, t) \} = \varphi^*(x, |\xi|).$$

Indeed, the inequality “ $\leq$ ” follows from

$$\xi \cdot z \leq |\xi| |z|,$$

while the reverse inequality follows, for $\xi \neq 0$, by choosing $z = t\xi/|\xi|$. The case $\xi = 0$ is immediate.

Let now $p \in L^{\varphi^*}(\Omega; \mathbb{R}^N)$. For every $v \in L^\varphi(\Omega; \mathbb{R}^N)$, the pointwise Fenchel inequality gives

$$p(x) \cdot v(x) - \Phi(x, v(x)) \leq \Phi^*(x, p(x)) = \varphi^*(x, |p(x)|) \quad \text{for a.e. } x \in \Omega.$$

After integration and taking the supremum over v , we obtain

$$\rho_\varphi^*(p) \leq \int_{\Omega} \varphi^*(x, |p|) \, dx.$$

It remains to prove the reverse inequality. For $R > 0$, define

$$M_R(x) := \sup_{0 \leq t \leq R} \{ |p(x)| t - \varphi(x, t) \}.$$

For $m \in \mathbb{N}$, let

$$D_{R,m} := \left\{ \frac{jR}{2^m} : j = 0, \dots, 2^m \right\}$$

and set

$$M_{R,m}(x) := \max_{t \in D_{R,m}} \{|p(x)|t - \varphi(x, t)\}.$$

Since $D_{R,m}$ is finite and $x \mapsto \varphi(x, t)$ is measurable for every fixed $t \geq 0$, the function $M_{R,m}$ is measurable. Moreover, $0 \in D_{R,m}$, and hence

$$M_{R,m}(x) \geq 0.$$

Choose $t_{R,m}(x) \in D_{R,m}$ as the smallest grid point where the maximum defining $M_{R,m}(x)$ is attained. This gives a measurable function $t_{R,m}$. Indeed, if

$$D_{R,m} = \{s_0, \dots, s_{2^m}\},$$

then the sets on which $t_{R,m} = s_j$ are obtained by finite intersections and unions of measurable sets of the form

$$\{|p|s_j - \varphi(\cdot, s_j) \geq |p|s_k - \varphi(\cdot, s_k)\}.$$

Define

$$e_p(x) := \begin{cases} \frac{p(x)}{|p(x)|}, & p(x) \neq 0, \\ e_1, & p(x) = 0, \end{cases}$$

where e_1 is the first coordinate vector in $\mathbb{R}^N$, and set

$$v_{R,m}(x) := t_{R,m}(x)e_p(x).$$

Then $v_{R,m}$ is measurable and satisfies

$$|v_{R,m}(x)| = t_{R,m}(x) \leq R.$$

Therefore

$$\rho_\varphi(v_{R,m}) = \int_{\Omega} \varphi(x, |v_{R,m}|) dx \leq \int_{\Omega} \varphi(x, R) dx < \infty,$$

by the standing assumption that $\varphi(\cdot, R) \in L^1(\Omega)$. Hence

$$v_{R,m} \in L^\varphi(\Omega; \mathbb{R}^N).$$

For this admissible test field, we have pointwise

$$p(x) \cdot v_{R,m}(x) - \Phi(x, v_{R,m}(x)) = |p(x)|t_{R,m}(x) - \varphi(x, t_{R,m}(x)) = M_{R,m}(x).$$

Thus, by the definition of the conjugate functional,

$$\rho_\varphi^*(p) \geq \int_{\Omega} M_{R,m}(x) dx.$$

For fixed $R > 0$, the grids $D_{R,m}$ are increasing in m and their union is dense in $[0, R]$. Since

$$t \mapsto |p(x)|t - \varphi(x, t)$$

is continuous on $[0, R]$ for a.e. x , we have

$$M_{R,m}(x) \uparrow M_R(x) \quad \text{as } m \rightarrow \infty.$$

By the monotone convergence theorem,

$$\rho_\varphi^*(p) \geq \int_{\Omega} M_R(x) dx.$$

Finally,

$$M_R(x) \uparrow \sup_{t \geq 0} \{|p(x)|t - \varphi(x, t)\} = \varphi^*(x, |p(x)|) \quad \text{as } R \rightarrow \infty.$$

A second application of the monotone convergence theorem yields

$$\rho_\varphi^*(p) \geq \int_{\Omega} \varphi^*(x, |p|) dx.$$

Combining both inequalities gives

$$\rho_\varphi^*(p) = \int_{\Omega} \varphi^*(x, |p|) dx = \int_{\Omega} \Phi^*(x, p(x)) dx.$$

In particular, this identity is understood in the extended-valued sense: the conjugate is finite precisely on the finite-modular class

$$\text{dom } \rho_{\varphi^*} = \left\{ p \in L^{\varphi^*}(\Omega; \mathbb{R}^N) : \int_{\Omega} \varphi^*(x, |p|) dx < \infty \right\}. \quad \square$$

Corollary 3.8 (subdifferential of the modular). For $v \in L^\varphi(\Omega; \mathbb{R}^N)$,

$$\partial \rho_\varphi(v) = \left\{ p \in L^{\varphi^*}(\Omega; \mathbb{R}^N) \mid p(x) \in \partial_z \Phi(x, v(x)) \text{ for a.e. } x \in \Omega \right\}.$$

Proof. By the Fenchel equality for the modular, $p \in \partial \rho_\varphi(v)$ is equivalent to

$$\rho_\varphi(v) + \rho_\varphi^*(p) = \int_{\Omega} p \cdot v dx.$$

Using [Proposition 3.7](#), this identity can be written as

$$\int_{\Omega} (\Phi(x, v(x)) + \Phi^*(x, p(x)) - p(x) \cdot v(x)) dx = 0.$$

The integrand is nonnegative by the pointwise Fenchel inequality. Hence the integral vanishes if and only if

$$\Phi(x, v(x)) + \Phi^*(x, p(x)) = p(x) \cdot v(x)$$

for almost every $x \in \Omega$. This pointwise Fenchel equality is equivalent to

$$p(x) \in \partial_z \Phi(x, v(x)) \quad \text{for a.e. } x \in \Omega,$$

which gives the stated characterization. The converse follows by reversing the same argument. $\square$

3.3 OPTIMALITY CONDITIONS

Since [Proposition 3.3](#) yields a minimizer of $\mathcal{F}$, we can analyze the dual problem and optimality conditions from [Theorem 2.7](#). In the notation of [Section 2.3](#), we take

- $X = X_\varphi$ and $Y = L^\varphi(\Omega; \mathbb{R}^N)$,
- $A : X_\varphi \rightarrow Y, Au := \nabla u$,

- $F_0(u) := \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2$, and $G(v) := \rho_\varphi(v)$.

Before applying Fenchel–Rockafellar duality, we verify the qualification condition. [Proposition 3.3](#) gives existence of a primal minimizer. Moreover, $0 \in X_\varphi$,

$$F_0(0) = \frac{1}{2} \|f\|_{L^2(\Omega)}^2 < \infty, \quad G(A(0)) = \rho_\varphi(0) = 0 < \infty.$$

Since G is finite on $Y = L^\varphi(\Omega; \mathbb{R}^N)$, we have

$$\text{dom } G = Y \quad \text{and hence} \quad A(0) = 0 \in (\text{dom } G)^\circ.$$

Thus the Fenchel–Rockafellar theorem applies, and the dual problem admits a solution.

Lemma 3.9 (dual problem). *The Fenchel dual problem corresponding to minimization of $\mathcal{F} = F_0 + G \circ A$ is*

$$\sup_{p \in K_\varphi} \left\{ -\frac{1}{2} \|\text{div } p\|_{L^2(\Omega)}^2 - \int_\Omega f \text{div } p \, dx - \int_\Omega \varphi^*(x, |p(x)|) \, dx \right\},$$

where

$$K_\varphi := \left\{ p \in L^{\varphi^*}(\Omega; \mathbb{R}^N) \mid \text{div } p \in L^2(\Omega) \text{ and } \int_\Omega p \cdot \nabla u \, dx = - \int_\Omega (\text{div } p) u \, dx \text{ for all } u \in X_\varphi \right\}.$$

The integration-by-parts identity is the weak formulation of the homogeneous normal boundary condition $p \cdot \nu = 0$ on $\partial\Omega$. This problem admits a maximizer.

Proof. We take dual variables in the associate space $L^{\varphi^*}(\Omega; \mathbb{R}^N)$, using [Lemma 3.6](#). For $p \in L^{\varphi^*}(\Omega; \mathbb{R}^N)$, the functional $A^*p \in X_\varphi^*$ is given by

$$\langle A^*p, u \rangle = \int_\Omega p \cdot \nabla u \, dx.$$

We first note that X_φ is dense in $L^2(\Omega)$. Indeed,

$$W^{1,\infty}(\Omega) \subset X_\varphi,$$

because Ω is bounded, $\nabla u \in L^\infty(\Omega; \mathbb{R}^N)$, and $\varphi(\cdot, t) \in L^1(\Omega)$ for every fixed $t \geq 0$. Moreover, $W^{1,\infty}(\Omega)$ is dense in $L^2(\Omega)$. Hence X_φ is dense in $L^2(\Omega)$.

The quantity $F_0^*(-A^*p)$ is finite only if $-A^*p$ is represented by an element of $L^2(\Omega)$. If there exists $g \in L^2(\Omega)$ such that

$$\int_\Omega p \cdot \nabla u \, dx = - \int_\Omega g u \, dx \quad \text{for all } u \in X_\varphi,$$

then we write $g = \text{div } p$ in this weak sense. In this case $-A^*p$ is represented by $\text{div } p \in L^2(\Omega)$, and therefore

$$F_0^*(-A^*p) = \sup_{u \in X_\varphi} \left\{ \int_\Omega (\text{div } p) u \, dx - \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 \right\}.$$

Since X_φ is dense in $L^2(\Omega)$, this supremum equals the supremum over all $u \in L^2(\Omega)$. Setting $w = u - f$, one obtains

$$\begin{aligned} F_0^*(-A^*p) &= \sup_{w \in L^2(\Omega)} \left\{ \int_\Omega (\text{div } p)(w + f) \, dx - \frac{1}{2} \|w\|_{L^2(\Omega)}^2 \right\} \\ &= \frac{1}{2} \|\text{div } p\|_{L^2(\Omega)}^2 + \int_\Omega f \text{div } p \, dx. \end{aligned}$$

If no such L^2 -representative exists, then $F_0^*(-A^*p) = +\infty$. Thus $F_0^*(-A^*p)$ is finite exactly on the divergence and weak-boundary class appearing in K_φ .

Proposition 3.7 gives

$$G^*(p) = \int_{\Omega} \varphi^*(x, |p(x)|) dx,$$

with value $+\infty$ when $p \notin \text{dom } \rho_{\varphi^*}$. Combining the two conjugates proves the dual formula. $\square$

Proposition 3.10 (optimality conditions). *The dual problem in Lemma 3.9 admits a maximizer $p \in K_\varphi$. Let $u \in X_\varphi$ be the primal minimizer and let $p \in K_\varphi$ be a dual maximizer. Then*

$$\text{div } p = u - f \quad \text{in } L^2(\Omega),$$

and

$$p(x) \in \partial_z \Phi(x, \nabla u(x)) \quad \text{for a.e. } x \in \Omega.$$

If $\nabla u(x) \neq 0$, this can be written as

$$(3.2) \quad p(x) = a(x) \frac{\nabla u(x)}{|\nabla u(x)|}, \quad a(x) \in \partial_t \varphi(x, |\nabla u(x)|).$$

Proof. Dual attainment follows from Fenchel–Rockafellar and the qualification verified above. The Fenchel–Rockafellar optimality system gives

$$-A^*p \in \partial F_0(u), \quad p \in \partial G(\nabla u).$$

Since $\partial F_0(u) = u - f$, the first condition is $\text{div } p = u - f$ in $L^2(\Omega)$, with the boundary condition already encoded by $p \in K_\varphi$. By Corollary 3.8, the second condition is precisely

$$p(x) \in \partial_z \Phi(x, \nabla u(x)) \quad \text{for a.e. } x \in \Omega.$$

The radial description follows from the subdifferential of $z \mapsto \varphi(x, |z|)$. $\square$

4 THE SCALED LOGARITHMIC MODEL

We now specialize the abstract Orlicz framework from Section 3 to a spatially dependent, logarithmically corrected total-variation density. The idea is to obtain a continuous Orlicz–Sobolev relaxation that avoiding the BV-relaxation of the ROF model, while remaining close to total variation. Through a local scale function κ , the model can tune how strongly the logarithmic correction enters across the domain. We record resulting basic properties, dual formulation, primal-dual splitting structure, and the Γ -convergence to the ROF functional as the logarithmic correction vanishes.

4.1 DEFINITION AND BASIC PROPERTIES

Let $\alpha > 0$ be fixed. We prescribe a local gradient scale

$$(4.1) \quad \kappa \in L^\infty(\Omega), \quad 0 < \kappa_- \leq \kappa(x) \leq \kappa_+ < \infty \quad \text{for a.e. } x \in \Omega.$$

For $\gamma > 0$, define the scaled logarithmic integrand

$$(4.2) \quad \varphi_\gamma(x, t) := \alpha t \left(1 + \gamma \log^+ \frac{t}{\kappa(x)} \right), \quad t \geq 0.$$

The scale function κ remains fixed throughout this section and is therefore suppressed in the notation. Equivalently,

$$(4.3) \quad \varphi_Y(x, t) = \begin{cases} \alpha t, & 0 \leq t \leq \kappa(x), \\ \alpha t \left(1 + \gamma \log \frac{t}{\kappa(x)}\right), & t \geq \kappa(x). \end{cases}$$

Remark 4.1 (smooth logarithmic variant). As a smooth alternative to (4.3), one may consider

$$\tilde{\varphi}_Y(x, t) := \alpha t \left(1 + \gamma \log \left(1 + \frac{t}{\kappa(x)}\right)\right), \quad t \geq 0.$$

Both φ_Y and $\tilde{\varphi}_Y$ generate the same generalized Orlicz and Orlicz–Sobolev spaces with equivalent Luxemburg norms. However, the pointwise dual proximal map related to $\tilde{\varphi}_Y$ requires the inversion of a scalar nonlinear function rather than admitting the explicit piecewise formula derived in Section 4.3.

Lemma 4.2 (properties of the scaled logarithmic integrand). For every $\gamma > 0$, the function

$$\varphi_Y : \Omega \times [0, \infty) \rightarrow [0, \infty)$$

satisfies the standing assumptions of Section 3. More precisely:

1. For almost every $x \in \Omega$, the map $t \mapsto \varphi_Y(x, t)$ is convex, continuous, increasing, satisfies $\varphi_Y(x, 0) = 0$, and

$$\lim_{t \rightarrow \infty} \frac{\varphi_Y(x, t)}{t} = \infty.$$

2. The superlinear growth is uniform in x :

$$\lim_{t \rightarrow \infty} \operatorname{ess\,inf}_{x \in \Omega} \frac{\varphi_Y(x, t)}{t} = \infty.$$

3. The uniform Δ_2 -condition holds:

$$\varphi_Y(x, 2t) \leq 2(1 + \gamma \log 2) \varphi_Y(x, t) \quad \text{for a.e. } x \in \Omega, \quad t \geq 0.$$

4. For almost every $x \in \Omega$, the function $t \mapsto \varphi_Y(x, t)$ satisfies $(\operatorname{Inc})_1$, but it does not satisfy $(\operatorname{aInc})_p$ for any $p > 1$.

5. For every $q \geq 1 + \gamma$, the function $t \mapsto \varphi_Y(x, t)$ satisfies $(\operatorname{Dec})_q$ uniformly in x .

6. As $\gamma \downarrow 0$,

$$\varphi_Y(x, t) \rightarrow \alpha t$$

locally uniformly in $t \in [0, \infty)$, uniformly in x .

Proof. Since κ is measurable, $x \mapsto \varphi_Y(x, t)$ is measurable for every $t \geq 0$; since $\kappa \geq \kappa_- > 0$ and Ω is bounded, $\varphi_Y(\cdot, t) \in L^1(\Omega)$ for every $t \geq 0$. The piecewise representation (4.3) shows continuity, monotonicity, and $\varphi_Y(x, 0) = 0$. For $0 < t < \kappa(x)$, the derivative with respect to t equals α , while for $t > \kappa(x)$ it equals

$$\alpha \left(1 + \gamma + \gamma \log \frac{t}{\kappa(x)}\right),$$

which is increasing in t . Hence $t \mapsto \varphi_Y(x, t)$ is convex.

The lower bound $\varphi_Y(x, t) \geq \alpha t$ is immediate from $\log^+ \geq 0$. Moreover,

$$\frac{\varphi_Y(x, t)}{t} \geq \alpha \left(1 + \gamma \log^+ \frac{t}{\kappa_+} \right) \quad \text{for } t \geq \kappa_+,$$

which proves uniform superlinear growth.

To prove the Δ_2 -condition, note that $\log^+(2s) \leq \log 2 + \log^+ s$ for $s \geq 0$. Therefore

$$\varphi_Y(x, 2t) \leq 2\alpha t \left(1 + \gamma \log 2 + \gamma \log^+ \frac{t}{\kappa(x)} \right) \leq 2(1 + \gamma \log 2) \varphi_Y(x, t).$$

The $(\text{Inc})_1$ property follows from the monotonicity of

$$\frac{\varphi_Y(x, t)}{t} = \alpha \left(1 + \gamma \log^+ \frac{t}{\kappa(x)} \right).$$

Let $p > 1$. Suppose that $\varphi_Y(x, \cdot)$ satisfied $(\text{aInc})_p$ with constant L . Then for $0 < s < t < \kappa(x)$,

$$\frac{\varphi_Y(x, s)}{s^p} \leq L \frac{\varphi_Y(x, t)}{t^p}.$$

Since $\varphi_Y(x, r) = \alpha r$ for $0 < r < \kappa(x)$, this becomes

$$\alpha s^{1-p} \leq L \alpha t^{1-p}.$$

Fixing $t \in (0, \kappa(x))$ and letting $s \downarrow 0$ gives a contradiction. Hence $(\text{aInc})_p$ fails for every $p > 1$.

For the almost decreasing property, set

$$h_q(t) := \frac{\varphi_Y(x, t)}{t^q}.$$

On $(0, \kappa(x))$, this is αt^{1-q} and is decreasing for $q \geq 1$. On $(\kappa(x), \infty)$,

$$h'_q(t) = \alpha t^{-q} \left(\gamma - (q-1) \left(1 + \gamma \log^+ \frac{t}{\kappa(x)} \right) \right),$$

which is nonpositive for $q \geq 1 + \gamma$. Thus $(\text{Dec})_q$ holds uniformly in x . Finally, on each interval $0 \leq t \leq R$,

$$0 \leq \varphi_Y(x, t) - \alpha t \leq \alpha \gamma R \log^+ \frac{R}{\kappa_-},$$

which gives local uniform convergence to αt . $\square$

For $\gamma > 0$, set

$$X_Y := \left\{ u \in L^2(\Omega) \cap W^{1,1}(\Omega) \mid \int_{\Omega} \varphi_Y(x, |\nabla u|) \, dx < \infty \right\}.$$

We consider

$$(4.4) \quad \mathcal{F}_Y(u) := \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \int_{\Omega} \varphi_Y(x, |\nabla u|) \, dx, \quad u \in X_Y.$$

Proposition 4.3 (existence and uniqueness). *For every $\gamma > 0$, the functional $\mathcal{F}_Y$ admits a unique minimizer in X_Y .*

Proof. By Lemma 4.2, the integrand φ_Y satisfies the standing assumptions of Section 3. Hence Proposition 3.3 yields existence of a minimizer in X_Y . Uniqueness follows from the strict convexity of the L^2 -fidelity term and the convexity of the regularization term. $\square$

Remark 4.4. Alternatively to φ_Y ,

4.2 FENCHEL DUALITY AND OPTIMALITY SYSTEM

Lemma 4.5 (dual problem for the scaled logarithmic regularizer). *The Fenchel dual problem corresponding to (4.4) is*

$$\sup_{p \in K_Y} \left\{ -\frac{1}{2} \|\operatorname{div} p\|_{L^2(\Omega)}^2 - \int_{\Omega} f \operatorname{div} p \, dx - \int_{\Omega} \varphi_Y^*(x, |p(x)|) \, dx \right\},$$

where

$$K_Y := \left\{ p \in L^{\varphi_Y^*}(\Omega; \mathbb{R}^N) \mid \operatorname{div} p \in L^2(\Omega) \text{ and } \int_{\Omega} p \cdot \nabla u \, dx = - \int_{\Omega} (\operatorname{div} p) u \, dx \text{ for all } u \in X_Y \right\}.$$

The integration-by-parts identity is the weak homogeneous normal boundary condition.

Proof. Apply Lemma 3.9 with $\varphi = \varphi_Y$. The assumptions of Section 3 hold by Lemma 4.2. $\square$

Lemma 4.6 (scalar subdifferentials and conjugate). *For almost every $x \in \Omega$, regarding $t \mapsto \varphi_Y(x, t)$ as a convex function on $[0, \infty)$, there holds*

$$\partial_t \varphi_Y(x, t) = \begin{cases} (-\infty, \alpha], & t = 0, \\ \{\alpha\}, & 0 < t < \kappa(x), \\ [\alpha, \alpha(1 + \gamma)], & t = \kappa(x), \\ \left\{ \alpha \left(1 + \gamma + \gamma \log \frac{t}{\kappa(x)} \right) \right\}, & t > \kappa(x). \end{cases}$$

For clarity, we regard $\varphi_Y^*(x, \cdot)$ as the Fenchel conjugate on $\mathbb{R}$ of the function $\varphi_Y(x, \cdot)$ defined on $[0, \infty)$, that is,

$$\varphi_Y^*(x, s) = \sup_{t \geq 0} \{st - \varphi_Y(x, t)\}, \quad s \in \mathbb{R}.$$

Then

$$\varphi_Y^*(x, s) = \begin{cases} 0, & s \leq \alpha, \\ \kappa(x)(s - \alpha), & \alpha < s \leq \alpha(1 + \gamma), \\ \alpha\gamma\kappa(x) \exp\left(\frac{s/\alpha - 1 - \gamma}{\gamma}\right), & s > \alpha(1 + \gamma). \end{cases}$$

The formula is continuous at $s = \alpha$ and $s = \alpha(1 + \gamma)$. Moreover,

$$(4.5) \quad \partial_s \varphi_Y^*(x, s) = \begin{cases} \{0\}, & s < \alpha, \\ [0, \kappa(x)], & s = \alpha, \\ \{\kappa(x)\}, & \alpha < s \leq \alpha(1 + \gamma), \\ \left\{ \kappa(x) \exp\left(\frac{s/\alpha - 1 - \gamma}{\gamma}\right) \right\}, & s > \alpha(1 + \gamma). \end{cases}$$

Finally, for

$$\Phi_Y(x, z) := \varphi_Y(x, |z|),$$

the corresponding vectorial subdifferential is radial. For $z \neq 0$,

$$\partial_z \Phi_Y(x, z) = \partial_t \varphi_Y(x, |z|) \frac{z}{|z|},$$

and hence

$$\partial_z \Phi_Y(x, z) = \begin{cases} \left\{ \alpha \frac{z}{|z|} \right\}, & 0 < |z| < \kappa(x), \\ \left\{ \lambda \frac{z}{|z|} : \lambda \in [\alpha, \alpha(1 + \gamma)] \right\}, & |z| = \kappa(x), \\ \left\{ \alpha \left(1 + \gamma + \gamma \log \frac{|z|}{\kappa(x)} \right) \frac{z}{|z|} \right\}, & |z| > \kappa(x). \end{cases}$$

At the origin,

$$\partial_z \Phi_Y(x, 0) = \{ \xi \in \mathbb{R}^N : |\xi| \leq \alpha \}.$$

Proof. Fix x outside a null set and write $\kappa = \kappa(x)$. Then

$$\varphi_Y(x, t) = \begin{cases} \alpha t, & 0 \leq t \leq \kappa, \\ \alpha t \left(1 + \gamma \log \frac{t}{\kappa} \right), & t \geq \kappa. \end{cases}$$

We first compute the scalar subdifferential of $\varphi_Y(x, \cdot)$. For $0 < t < \kappa$,

$$\varphi'_Y(x, t) = \alpha,$$

while for $t > \kappa$,

$$\varphi'_Y(x, t) = \alpha \left(1 + \gamma + \gamma \log \frac{t}{\kappa} \right).$$

At $t = \kappa$, the left derivative is α and the right derivative is $\alpha(1 + \gamma)$, hence

$$\partial_t \varphi_Y(x, \kappa) = [\alpha, \alpha(1 + \gamma)].$$

At $t = 0$, the subgradient inequality over the domain $[0, \infty)$ is

$$\xi t \leq \varphi_Y(x, t) \quad \text{for all } t \geq 0.$$

If $\xi \leq \alpha$, then $\xi t \leq \alpha t \leq \varphi_Y(x, t)$ for every $t \geq 0$. If $\xi > \alpha$, the inequality fails for every sufficiently small positive $t \leq \kappa$. Thus it holds precisely when $\xi \leq \alpha$, and

$$\partial_t \varphi_Y(x, 0) = (-\infty, \alpha].$$

We now compute the conjugate. On the linear branch $0 \leq t \leq \kappa$,

$$st - \varphi_Y(x, t) = (s - \alpha)t.$$

Therefore the supremum over this interval is 0 if $s \leq \alpha$, and $\kappa(s - \alpha)$ if $s > \alpha$.

On the logarithmic branch $t \geq \kappa$, set

$$\Psi_s(t) := st - \alpha t \left(1 + \gamma \log \frac{t}{\kappa} \right).$$

Then

$$\Psi'_s(t) = s - \alpha(1 + \gamma) - \alpha\gamma \log \frac{t}{\kappa}, \quad \Psi''_s(t) = -\frac{\alpha\gamma}{t} < 0.$$

Thus a critical point is the unique maximizer on the logarithmic branch. Solving $\Psi'_s(t) = 0$ gives

$$t_s = \kappa \exp \left(\frac{s/\alpha - 1 - \gamma}{\gamma} \right).$$

This satisfies $t_s \geq \kappa$ exactly when $s \geq \alpha(1 + \gamma)$. If $s \leq \alpha(1 + \gamma)$, then $\Psi'_s(\kappa) \leq 0$, so concavity implies that the logarithmic branch is maximized at the endpoint $t = \kappa$, giving the same value $\kappa(s - \alpha)$ as the right endpoint of the linear branch. Combining this with the linear-branch computation gives the first two cases of the conjugate.

For $s > \alpha(1 + \gamma)$, the unique maximizer is $t_s > \kappa$. The stationarity identity can be rewritten as

$$s - \alpha - \alpha\gamma \log \frac{t_s}{\kappa} = \alpha\gamma.$$

Consequently,

$$\begin{aligned} \varphi_Y^*(x, s) &= st_s - \alpha t_s \left(1 + \gamma \log \frac{t_s}{\kappa}\right) \\ &= t_s \left(s - \alpha - \alpha\gamma \log \frac{t_s}{\kappa}\right) \\ &= \alpha\gamma t_s \\ &= \alpha\gamma\kappa \exp\left(\frac{s/\alpha - 1 - \gamma}{\gamma}\right), \end{aligned}$$

which is the exponential branch.

The endpoint values agree. At $s = \alpha$,

$$0 = \kappa(\alpha - \alpha).$$

At $s = \alpha(1 + \gamma)$,

$$\kappa(\alpha(1 + \gamma) - \alpha) = \alpha\gamma\kappa$$

and

$$\alpha\gamma\kappa \exp\left(\frac{1 + \gamma - 1 - \gamma}{\gamma}\right) = \alpha\gamma\kappa.$$

The formula for $\partial_s \varphi_Y^*$ follows by differentiating the smooth pieces and taking the interval between the one-sided slopes at $s = \alpha$. At $s = \alpha(1 + \gamma)$, the left and right derivatives both equal κ , so the subdifferential there is the singleton $\{\kappa\}$.

It remains to justify the vectorial subdifferential. For $z \neq 0$, the stated formula follows from the chain rule for the convex radial map $z \mapsto \varphi_Y(x, |z|)$. At $z = 0$, the subgradient condition is

$$\xi \cdot w \leq \varphi_Y(x, |w|) \quad \text{for all } w \in \mathbb{R}^N.$$

If $|\xi| \leq \alpha$, then $\xi \cdot w \leq \alpha|w| \leq \varphi_Y(x, |w|)$ for all w . Conversely, if $|\xi| > \alpha$, choosing $w = \varepsilon\xi/|\xi|$ with $0 < \varepsilon \leq \kappa$ gives

$$\xi \cdot w = \varepsilon|\xi| > \alpha\varepsilon = \varphi_Y(x, \varepsilon),$$

so the subgradient inequality fails. Hence $\partial_z \Phi_Y(x, 0) = \{\xi \in \mathbb{R}^N : |\xi| \leq \alpha\}$. $\square$

Proposition 4.7 (optimality conditions). *Let $u_Y \in X_Y$ be the primal minimizer and let $p_Y \in K_Y$ be a dual maximizer. Then*

$$\begin{aligned} \operatorname{div} p_Y &= u_Y - f && \text{in } L^2(\Omega), \\ p_Y(x) &\in \partial_z \Phi_Y(x, \nabla u_Y(x)) && \text{for a.e. } x \in \Omega, \end{aligned}$$

where

$$\Phi_Y(x, z) := \varphi_Y(x, |z|).$$

Equivalently, if $\nabla u_Y(x) \neq 0$, then

$$p_Y(x) = a(x) \frac{\nabla u_Y(x)}{|\nabla u_Y(x)|},$$

$$a(x) \in \partial_t \varphi_Y(x, |\nabla u_Y(x)|),$$

pointwise a.e. in Ω , and if $\nabla u_Y(x) = 0$, then

$$|p_Y(x)| \leq \alpha.$$

Proof. This is (3.2) applied to $\varphi = \varphi_Y$. The radial form follows from the chain rule for the convex radial integrand $\Phi_Y(x, z) = \varphi_Y(x, |z|)$ and Lemma 4.6. $\square$

4.3 PRIMAL-DUAL SPLITTING STRUCTURE

We now isolate the continuous composite structure of the scaled logarithmic model. Throughout this subsection, $\gamma > 0$ and the scale function κ are fixed. The model can be written as

$$\min_u \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \int_{\Omega} \varphi_Y(x, |\nabla u|) dx.$$

With

$$F(u) := \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2, \quad G(v) := \int_{\Omega} \varphi_Y(x, |v|) dx, \quad A := \nabla,$$

this is the composite convex problem

$$\min_u F(u) + G(Au) = \min_u F(u) + G(\nabla u).$$

This is exactly the form used in the Fenchel-duality framework of Section 2.3 and in the primal-dual splitting framework recalled in [6, Chapter 7].

The conjugate of the integral term is, formally and consistently with Lemma 4.5,

$$G^*(p) = \int_{\Omega} \varphi_Y^*(x, |p(x)|) dx.$$

The corresponding saddle formulation is

$$\inf_u \sup_p \left[\frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \int_{\Omega} p \cdot \nabla u dx - \int_{\Omega} \varphi_Y^*(x, |p|) dx \right].$$

When integrating by parts, the homogeneous normal boundary condition is understood in the weak sense encoded in the admissible dual class K_Y .

Recall that the Fenchel extremality conditions for $F + G \circ A$ are

$$-\nabla^* p \in \partial F(u), \quad p \in \partial G(\nabla u).$$

Equivalently, using $\nabla^* = -\operatorname{div}$ and the Fenchel relation $p \in \partial G(\nabla u)$ if and only if $\nabla u \in \partial G^*(p)$, they read

$$\operatorname{div} p = u - f, \quad p(x) \in \partial_z \Phi_Y(x, \nabla u(x)) \quad \text{for a.e. } x \in \Omega.$$

The formal proximal fixed-point form is obtained from the standard resolvent identity. For any $\tau > 0$ and $\sigma > 0$,

$$-\nabla^* p \in \partial F(u) \quad \Leftrightarrow \quad u = \operatorname{prox}_{\tau F}(u - \tau \nabla^* p),$$

and

$$\nabla u \in \partial G^*(p) \quad \Leftrightarrow \quad p = \text{prox}_{\sigma G^*}(p + \sigma \nabla u).$$

Since $\nabla^* = -\text{div}$, this becomes

$$u = \text{prox}_{\tau F}(u + \tau \text{div } p), \quad p = \text{prox}_{\sigma G^*}(p + \sigma \nabla u).$$

The associated formal primal-dual iteration is

$$(4.6) \quad \begin{cases} p^{k+1} = \text{prox}_{\sigma G^*}(p^k + \sigma \nabla \bar{u}^k), \\ u^{k+1} = \text{prox}_{\tau F}(u^k + \tau \text{div } p^{k+1}), \\ \bar{u}^{k+1} = u^{k+1} + \theta(u^{k+1} - u^k). \end{cases}$$

The proximal map of the fidelity term is explicit. Indeed, for $w \in L^2(\Omega)$,

$$\text{prox}_{\tau F}(w) = \frac{w + \tau f}{1 + \tau}.$$

The dual proximal map is radial and also has an explicit pointwise form.

Lemma 4.8 (explicit pointwise dual proximal map). *Let $\sigma > 0$. Formally, with respect to the $L^2(\Omega; \mathbb{R}^N)$ -metric, the proximal map of G^* is pointwise radial. For $q \in L^2(\Omega; \mathbb{R}^N)$,*

$$\text{prox}_{\sigma G^*}(q)(x) = \begin{cases} 0, & q(x) = 0, \\ \rho_\sigma(x, |q(x)|) \frac{q(x)}{|q(x)|}, & q(x) \neq 0, \end{cases}$$

where, for $r \geq 0$,

$$(4.7) \quad \rho_\sigma(x, r) = \begin{cases} r, & 0 \leq r \leq \alpha, \\ \alpha, & \alpha < r \leq \alpha + \sigma \kappa(x), \\ r - \sigma \kappa(x), & \alpha + \sigma \kappa(x) < r \leq \alpha(1 + \gamma) + \sigma \kappa(x), \\ r - \alpha \gamma W\left(\frac{\sigma \kappa(x)}{\alpha \gamma} \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right)\right), & r > \alpha(1 + \gamma) + \sigma \kappa(x). \end{cases}$$

Here W denotes the principal branch of the Lambert W -function (see [7]).

Proof. Since

$$G^*(p) = \int_{\Omega} \phi_\gamma^*(x, |p(x)|) dx,$$

the formal L^2 -proximal problem is

$$\text{prox}_{\sigma G^*}(q) = \arg \min_p \int_{\Omega} \left[\frac{1}{2\sigma} |p(x) - q(x)|^2 + \phi_\gamma^*(x, |p(x)|) \right] dx.$$

The integrand is pointwise and radial. Therefore the minimizer at each x is parallel to $q(x)$. If $q(x) = 0$, the minimizer is 0. If $q(x) \neq 0$, write

$$r := |q(x)|$$

and search for the minimizer in the form

$$p(x) = \rho \frac{q(x)}{|q(x)|}.$$

Then $\rho \geq 0$ minimizes

$$\min_{\rho \geq 0} \frac{1}{2\sigma} (\rho - r)^2 + \varphi_Y^*(x, \rho).$$

Fix x and write $\kappa = \kappa(x)$. Since $\rho \geq 0$, [Lemma 4.6](#) gives

$$\varphi_Y^*(x, \rho) = \begin{cases} 0, & 0 \leq \rho \leq \alpha, \\ \kappa(\rho - \alpha), & \alpha < \rho \leq \alpha(1 + \gamma), \\ \alpha\gamma\kappa \exp\left(\frac{\rho/\alpha - 1 - \gamma}{\gamma}\right), & \rho > \alpha(1 + \gamma). \end{cases}$$

We distinguish four regimes.

If $0 \leq \rho \leq \alpha$, then the scalar objective is

$$J(\rho) = \frac{1}{2\sigma} (\rho - r)^2.$$

The minimizer is $\rho = r$, which is admissible exactly when $0 \leq r \leq \alpha$. Hence $\rho_\sigma(x, r) = r$ in the first case.

At the boundary $\rho = \alpha$, the scalar optimality condition is

$$0 \in \frac{1}{\sigma} (\rho - r) + \partial_\rho \varphi_Y^*(x, \rho).$$

Since $\partial_\rho \varphi_Y^*(x, \alpha) = [0, \kappa]$, the point $\rho = \alpha$ is optimal if and only if

$$0 \in \frac{1}{\sigma} (\alpha - r) + [0, \kappa],$$

or equivalently

$$\frac{r - \alpha}{\sigma} \in [0, \kappa].$$

Thus $\alpha \leq r \leq \alpha + \sigma\kappa$. The endpoint $r = \alpha$ is already covered by the first regime, so the second case is $\rho_\sigma(x, r) = \alpha$ for $\alpha < r \leq \alpha + \sigma\kappa$.

On the linear branch $\alpha < \rho \leq \alpha(1 + \gamma)$, one has

$$\varphi_Y^*(x, \rho) = \kappa(\rho - \alpha), \quad \frac{d}{d\rho} \varphi_Y^*(x, \rho) = \kappa.$$

The optimality condition is

$$0 = \frac{1}{\sigma} (\rho - r) + \kappa,$$

so $\rho = r - \sigma\kappa$. This belongs to the linear branch precisely when

$$\alpha < r - \sigma\kappa \leq \alpha(1 + \gamma),$$

equivalently

$$\alpha + \sigma\kappa < r \leq \alpha(1 + \gamma) + \sigma\kappa.$$

This gives the third case.

On the exponential branch $\rho > \alpha(1 + \gamma)$,

$$\varphi_Y^*(x, \rho) = \alpha\gamma\kappa \exp\left(\frac{\rho/\alpha - 1 - \gamma}{\gamma}\right),$$

and

$$\frac{d}{d\rho}\varphi_Y^*(x, \rho) = \kappa \exp\left(\frac{\rho/\alpha - 1 - \gamma}{\gamma}\right).$$

The scalar optimality condition becomes

$$0 = \frac{1}{\sigma}(\rho - r) + \kappa \exp\left(\frac{\rho/\alpha - 1 - \gamma}{\gamma}\right),$$

or

$$r - \rho = \sigma\kappa \exp\left(\frac{\rho/\alpha - 1 - \gamma}{\gamma}\right).$$

Set $\delta := r - \rho$. Then

$$\delta = \sigma\kappa \exp\left(\frac{(r - \delta)/\alpha - 1 - \gamma}{\gamma}\right) = \sigma\kappa \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right) \exp\left(-\frac{\delta}{\alpha\gamma}\right).$$

Hence

$$\delta \exp\left(\frac{\delta}{\alpha\gamma}\right) = \sigma\kappa \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right).$$

Dividing by $\alpha\gamma$ gives

$$\frac{\delta}{\alpha\gamma} \exp\left(\frac{\delta}{\alpha\gamma}\right) = \frac{\sigma\kappa}{\alpha\gamma} \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right).$$

By the defining identity $W(z)e^{W(z)} = z$,

$$\frac{\delta}{\alpha\gamma} = W\left(\frac{\sigma\kappa}{\alpha\gamma} \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right)\right).$$

Therefore

$$\rho = r - \delta = r - \alpha\gamma W\left(\frac{\sigma\kappa}{\alpha\gamma} \exp\left(\frac{r/\alpha - 1 - \gamma}{\gamma}\right)\right).$$

This branch applies exactly when $r > \alpha(1 + \gamma) + \sigma\kappa$. Putting the four regimes together proves the formula. $\square$

The scaled logarithmic model therefore retains the same primal–dual splitting architecture as ROF denoising. The primal proximal map is the explicit averaging step associated with the quadratic fidelity term, whereas the dual proximal map remains pointwise and radial. More precisely, Lemma 4.8 shows that, for $q \neq 0$, it can be written in the form

$$\text{prox}_{\sigma G_Y^*}(q) = \rho_{\sigma, \gamma}(|q|) \frac{q}{|q|},$$

with the natural value 0 at $q = 0$. In contrast to the ROF case, the dual update is no longer the projection onto the ball $\{|p| \leq \alpha\}$. Instead, its radial profile $\rho_{\sigma, \gamma}$ consists of an identity branch, a constant sticking branch, a linear transition branch, and a nonlinear Lambert- W branch.

Figure 1 illustrates these four regimes for $\alpha = \kappa = \sigma = 1$. For $0 \leq r \leq \alpha$, the proximal map acts as the identity, while for $\alpha < r \leq \alpha + \sigma\kappa$ its radial profile is constant and coincides with the ROF projection onto the dual ball. The scaled logarithmic model therefore agrees exactly with the ROF dual update

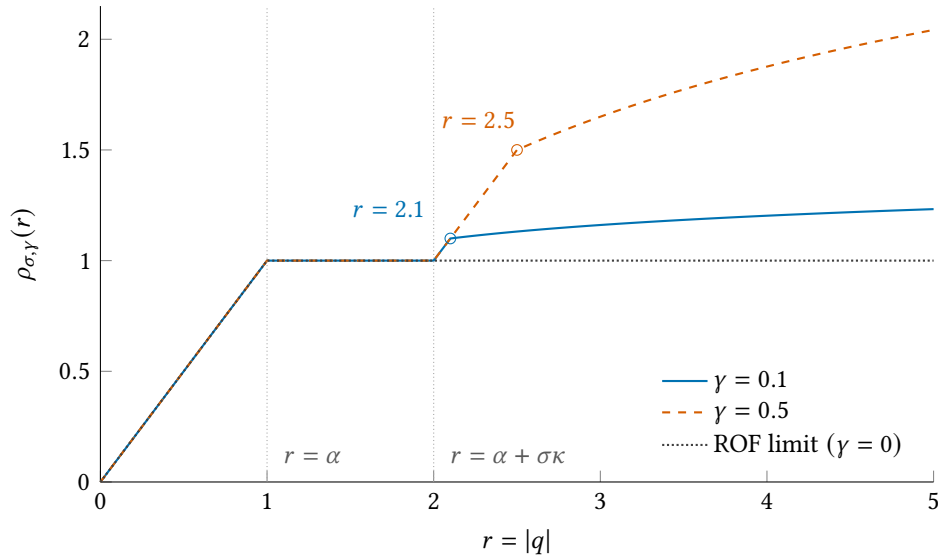


Figure 1: Radial profile $r \mapsto \rho_{\sigma,\gamma}(r)$ of the pointwise dual proximal map for $\alpha = \kappa = \sigma = 1$ and $\gamma \in \{0.1, 0.5\}$. The dotted black curve is the limiting ROF projection $\rho_{\sigma,0}(r) = \min\{r, \alpha\}$. The vertical dotted lines indicate the common transition points $r = \alpha$ and $r = \alpha + \sigma\kappa$, while the marked points indicate the transition from the linear branch to the nonlinear Lambert-W branch.

throughout these first two regimes. The difference appears only for $r > \alpha + \sigma\kappa$, where the profile leaves the constraint boundary and first follows the linear branch $\rho_{\sigma,\gamma}(r) = r - \sigma\kappa$. This branch terminates at the γ -dependent transition point $r = \alpha(1 + \gamma) + \sigma\kappa$, after which the nonlinear Lambert-W expression from Lemma 4.8 applies.

The length of the linear transition interval is $\alpha\gamma$, and hence increases with γ . Accordingly, larger values of γ delay the onset of the nonlinear branch and lead to a less pronounced saturation of the dual variable for large input magnitudes. Conversely, as $\gamma \downarrow 0$, the transition interval collapses and the radial profile converges to the hard ROF saturation $\rho_{\sigma,0}(r) = \min\{r, \alpha\}$. One can observe both the explicit piecewise structure of the dual proximal map and its limiting relation to the ROF projection.

4.4 Γ -CONVERGENCE AS $\gamma \rightarrow 0$

We extend $\mathcal{F}_\gamma$ to $L^1(\Omega)$ by setting

$$(4.8) \quad \mathcal{F}_\gamma(u) := \begin{cases} \frac{1}{2}\|u - f\|_{L^2(\Omega)}^2 + \int_{\Omega} \varphi_\gamma(x, |\nabla u|) dx, & u \in X_\gamma, \\ +\infty, & u \in L^1(\Omega) \setminus X_\gamma. \end{cases}$$

The limiting functional is the ROF functional

$$(4.9) \quad \mathcal{F}_0(u) := \begin{cases} \frac{1}{2}\|u - f\|_{L^2(\Omega)}^2 + \alpha|Du|(\Omega), & u \in BV(\Omega) \cap L^2(\Omega), \\ +\infty, & u \in L^1(\Omega) \setminus (BV(\Omega) \cap L^2(\Omega)). \end{cases}$$

Proposition 4.9 (compactness). *Let $\gamma_n \downarrow 0$ and let $\{u_n\}_n \subset L^1(\Omega)$ satisfy*

$$\sup_n \mathcal{F}_{\gamma_n}(u_n) < \infty.$$

Then $\{u_n\}_n$ is bounded in $L^2(\Omega)$ and in $BV(\Omega)$. In particular, there exist a subsequence, not relabelled, and $u \in BV(\Omega) \cap L^2(\Omega)$ such that

$$u_n \rightarrow u \quad \text{strongly in } L^1(\Omega),$$

and

$$u_n \rightharpoonup u \quad \text{weakly in } L^2(\Omega).$$

Proof. The energy bound yields

$$\sup_n \|u_n - f\|_{L^2(\Omega)} < \infty,$$

and hence $\{u_n\}_n$ is bounded in $L^2(\Omega)$. Since

$$\varphi_{\gamma_n}(x, t) \geq \alpha t \quad \text{for a.e. } x \in \Omega, \quad t \geq 0,$$

we also have

$$\alpha \int_{\Omega} |\nabla u_n| dx \leq \int_{\Omega} \varphi_{\gamma_n}(x, |\nabla u_n|) dx \leq C.$$

Together with $\|u_n\|_{L^1(\Omega)} \leq |\Omega|^{1/2} \|u_n\|_{L^2(\Omega)}$, this gives boundedness in $BV(\Omega)$. By compactness of bounded sets in $BV(\Omega)$ with respect to strong $L^1(\Omega)$ -convergence, the compact embedding $BV(\Omega) \hookrightarrow L^1(\Omega)$ yields, after passing to a subsequence, strong convergence in $L^1(\Omega)$; see [1, Chapter 3]. The L^2 -boundedness yields a further subsequence converging weakly in $L^2(\Omega)$. The L^1 -limit and the weak L^2 -limit coincide distributionally. $\square$

Theorem 4.10 (Γ -convergence to total variation). *As $\gamma \downarrow 0$, the functionals $\mathcal{F}_{\gamma} : L^1(\Omega) \rightarrow [0, \infty]$ Γ -converge to $\mathcal{F}_0$ with respect to strong convergence in $L^1(\Omega)$. Moreover, the family is equicoercive in $L^1(\Omega)$: if $\gamma_n \downarrow 0$ and $\sup_n \mathcal{F}_{\gamma_n}(u_n) < \infty$, then $(u_n)_n$ is precompact in $L^1(\Omega)$, and every L^1 -cluster point belongs to $BV(\Omega) \cap L^2(\Omega)$.*

Proof. We first prove the liminf inequality. Let $\gamma_n \downarrow 0$ and suppose that $u_n \rightarrow u$ strongly in $L^1(\Omega)$. If

$$\liminf_{n \rightarrow \infty} \mathcal{F}_{\gamma_n}(u_n) = +\infty,$$

there is nothing to prove. Otherwise, we pass to a subsequence realizing the liminf and with bounded energy. By Proposition 4.9, the limit satisfies $u \in BV(\Omega) \cap L^2(\Omega)$, and, after passing to a further subsequence,

$$u_n \rightharpoonup u \quad \text{weakly in } L^2(\Omega).$$

Hence

$$\frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 \leq \liminf_{n \rightarrow \infty} \frac{1}{2} \|u_n - f\|_{L^2(\Omega)}^2.$$

Moreover, $\varphi_{\gamma_n}(x, t) \geq \alpha t$, and the lower semicontinuity of total variation with respect to strong L^1 -convergence gives

$$\alpha |Du|(\Omega) \leq \liminf_{n \rightarrow \infty} \alpha \int_{\Omega} |\nabla u_n| dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} \varphi_{\gamma_n}(x, |\nabla u_n|) dx;$$

see [1, Chapter 3]. Combining the two estimates gives

$$\mathcal{F}_0(u) \leq \liminf_{n \rightarrow \infty} \mathcal{F}_{\gamma_n}(u_n).$$

We now prove the limsup inequality. Let $u \in L^1(\Omega)$. If $\mathcal{F}_0(u) = +\infty$, any sequence converging to u in $L^1(\Omega)$ is a recovery sequence. We therefore assume $u \in BV(\Omega) \cap L^2(\Omega)$. By the strict approximation

theorem for BV -functions on bounded Lipschitz domains, see [1, Theorem 3.9], and after the usual truncation argument if necessary, we may choose

$$v_k \in W^{1,\infty}(\Omega) \cap L^2(\Omega)$$

such that

$$v_k \rightarrow u \quad \text{in } L^1(\Omega), \quad v_k \rightarrow u \quad \text{in } L^2(\Omega),$$

and

$$\int_{\Omega} |\nabla v_k| \, dx \rightarrow |Du|(\Omega).$$

For fixed k , we then have $\nabla v_k \in L^\infty(\Omega; \mathbb{R}^N)$. Since $\kappa(x) \geq \kappa_- > 0$,

$$\begin{aligned} 0 &\leq \varphi_\gamma(x, |\nabla v_k|) - \alpha |\nabla v_k| \\ &= \alpha \gamma |\nabla v_k| \log^+ \left(\frac{|\nabla v_k|}{\kappa(x)} \right), \end{aligned}$$

and hence

$$0 \leq \varphi_\gamma(x, |\nabla v_k|) - \alpha |\nabla v_k| \leq \alpha \gamma \|\nabla v_k\|_{L^\infty(\Omega)} \log^+ \left(\frac{\|\nabla v_k\|_{L^\infty(\Omega)}}{\kappa_-} \right).$$

The right-hand side is a constant depending on k , multiplied by γ . Therefore

$$\lim_{\gamma \downarrow 0} \int_{\Omega} \varphi_\gamma(x, |\nabla v_k|) \, dx = \alpha \int_{\Omega} |\nabla v_k| \, dx.$$

Consequently,

$$\lim_{k \rightarrow \infty} \lim_{\gamma \downarrow 0} \mathcal{F}_\gamma(v_k) = \mathcal{F}_0(u).$$

A standard diagonal argument for recovery sequences, see [8, Chapter 1], gives $u_\gamma = v_{k(\gamma)}$ such that $u_\gamma \rightarrow u$ strongly in $L^1(\Omega)$ and

$$\limsup_{\gamma \downarrow 0} \mathcal{F}_\gamma(u_\gamma) \leq \mathcal{F}_0(u).$$

This proves the Γ -convergence. □

Remark 4.11 (convergence of minimizers). The equicoercivity statement in [Theorem 4.10](#) is the compactness property needed to pass from Γ -convergence to convergence of minimizers. Let u_γ be the unique minimizer of $\mathcal{F}_\gamma$. Since

$$\mathcal{F}_\gamma(u_\gamma) \leq \mathcal{F}_\gamma(0) = \frac{1}{2} \|f\|_{L^2(\Omega)}^2,$$

the family $(u_\gamma)_\gamma$ has L^1 -convergent subsequences as $\gamma \downarrow 0$. By the fundamental theorem of Γ -convergence, every cluster point is a minimizer of $\mathcal{F}_0$. Since $\mathcal{F}_0$ is strictly convex because of the quadratic fidelity term, its minimizer is unique. Therefore the whole family u_γ converges in $L^1(\Omega)$ to the unique minimizer of $\mathcal{F}_0$. The same argument applies to asymptotic minimizers satisfying

$$\mathcal{F}_\gamma(u_\gamma) \leq \inf \mathcal{F}_\gamma + o(1).$$

Remark 4.12 (role of the scale function). The scale function κ affects the finite- γ behavior of the regularizer by determining where the logarithmic correction becomes active. It does not appear in the $\gamma \downarrow 0$ limit, since $\varphi_\gamma(x, t) \rightarrow \alpha t$ locally uniformly in t , uniformly in x . Thus κ should be interpreted as a local gradient scale for the relaxed model, not as a weight in the limiting total variation. A weighted-TV limit would instead require an integrand whose leading term is $w(x)t$, producing a limit of the form $\alpha \int_{\Omega} w \, d|Du|$.

5 NUMERICAL ILLUSTRATIONS

We conclude with two numerical illustrations of the finite- γ behavior of the scaled logarithmic regularizer. The purpose is to demonstrate the effect of the logarithmic correction in a controlled example and to compare it with standard denoising models.

5.1 DISCRETIZATION AND EXPERIMENTAL SETUP

All images consist of $n \times n$ pixels with $n = 256$. We denote by D the unscaled forward-difference gradient, with outward differences set to zero at the image boundary, and use the Euclidean norm on the two gradient components. The discrete Orlicz problem is

$$(5.1) \quad \min_{u \in \mathbb{R}^{n \times n}} \frac{1}{2} \sum_{i,j} (u_{ij} - f_{ij})^2 + \alpha \sum_{i,j} |(Du)_{ij}| \left(1 + \gamma \log^+ \frac{|(Du)_{ij}|}{\kappa} \right).$$

We compare (5.1) with the discrete ROF model, obtained by setting $\gamma = 0$, and with quadratic H^1 regularization, whose regularization term is

$$\frac{\alpha}{2} \sum_{i,j} |(Du)_{ij}|^2.$$

The ROF and Orlicz problems are solved by the primal–dual algorithm of [4]. In both cases we use extrapolation parameter $\theta = 1$, primal and dual step sizes $\tau = \sigma = 0.99/\sqrt{8}$, and 5000 iterations. Thus the two implementations differ only in the pointwise dual proximal map: the ROF projection is replaced by the radial map derived in Section 4.3. The quadratic H^1 problem is solved by conjugate gradients.

Although the model permits a spatially varying scale, we use a scalar κ in all experiments. It is estimated from the noisy datum, without access to the ground truth, by

$$(5.2) \quad \kappa = \frac{1}{2} Q_{0.95} \left((|D(G_1 * f)|_{ij})_{(i,j): |D(G_1 * f)|_{ij} > 0} \right),$$

where G_1 denotes Gaussian smoothing with standard deviation one pixel and $Q_{0.95}$ denotes the empirical 95th percentile of the positive pixelwise gradient magnitudes. Since the clean reference is known in both examples, the remaining parameters are chosen independently for each method by minimizing the mean squared error over finite parameter grids. This oracle choice is used to compare the attainable reconstructions and is not intended as a practical parameter-selection rule.

The source code, notebooks, and numerical outputs are available in the archived software release [18].

5.2 BLURRED-DISK EXPERIMENT

On the square $[-1, 1]^2$, consider the radially symmetric phantom

$$u_\eta(x) = \frac{1}{1 + \exp\left(\frac{|x| - 0.35}{\eta}\right)}, \quad f_\eta = u_\eta + 0.2 \xi,$$

where ξ consists of independent standard Gaussian random variables with zero mean and standard deviation 0.2. The same noise realization is used for every transition width η . The scale estimates from (5.2) range from 0.04695 to 0.05037.

For each value of η , we performed an oracle parameter sweep over finite search grids. For the ROF and quadratic H^1 models, we selected the value of α attaining the lowest MSE, while for the Orlicz

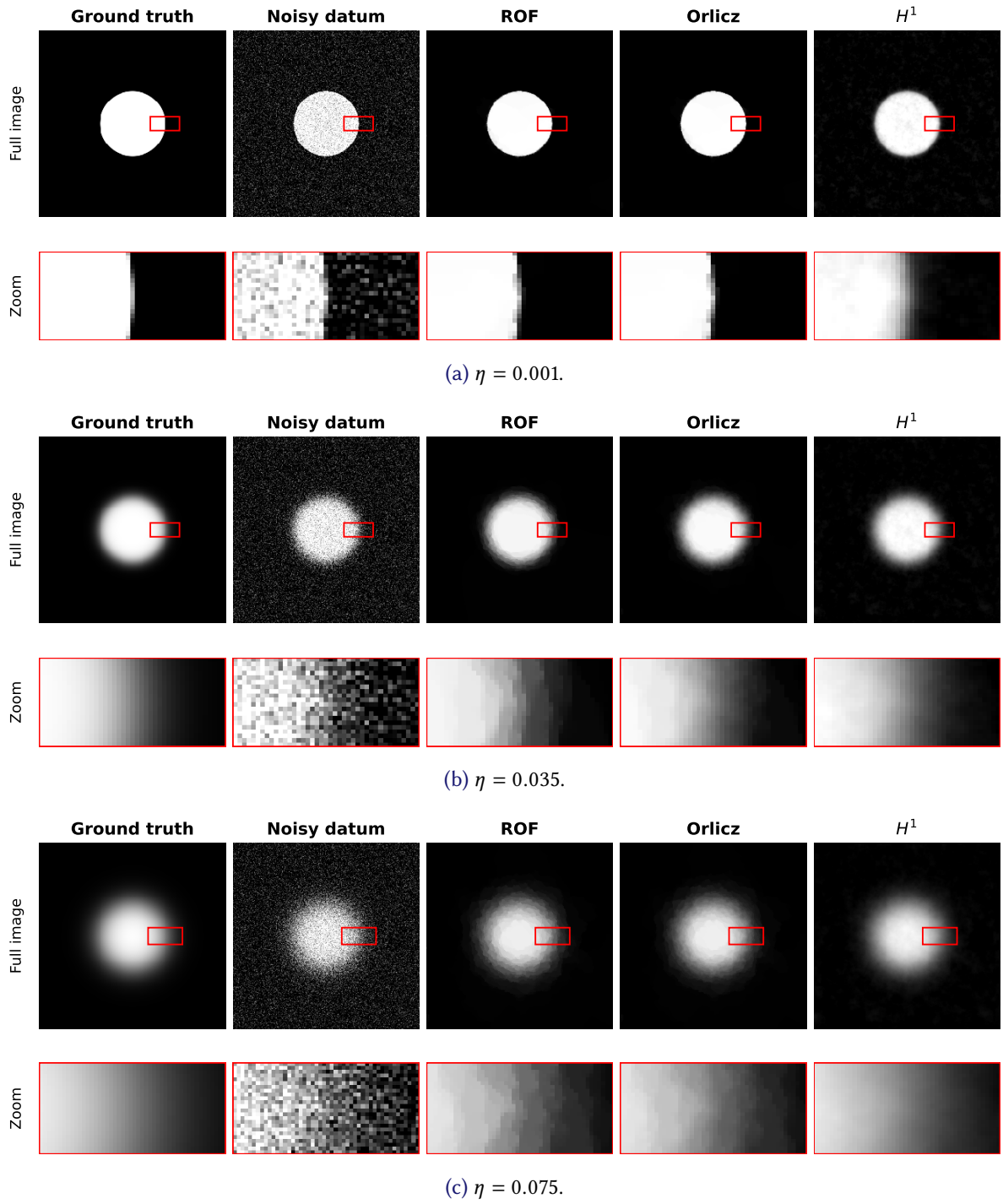


Figure 2: MSE-selected reconstructions of blurred disks with increasing transition width. In each panel, the columns show the ground truth, noisy datum, ROF, Orlicz, and quadratic H^1 reconstructions. The upper row contains the full images and the lower row enlarges the red transition region.

η	ROF		Orlicz		H^1	
	α	10^4MSE	(α, γ)	10^4MSE	α	10^4MSE
0.001	0.3	1.6961	(0.3, 0.005)	1.7279	5	24.9600
0.010	0.4	1.5427	(0.4, 0.10)	1.3412	5	10.6559
0.035	0.5	1.8969	(0.3, 0.60)	1.0796	10	4.2646
0.050	0.3	1.9827	(0.3, 1.00)	1.1580	20	3.0074
0.075	0.3	2.0650	(0.3, 1.25)	1.4624	20	2.2531

Table 1: MSE-selected parameters and corresponding mean squared errors for the blurred-disk experiment.

model we jointly selected the pair (α, γ) attaining the lowest MSE. Only these minimizing parameter choices are reported in Table 1; the corresponding reconstructions are those displayed in Figure 2.

For the nearly sharp interface $\eta = 0.001$, the selected value $\gamma = 0.005$ is close to the ROF limit, and the two reconstructions in Figure 2a are correspondingly very similar. As the transition becomes more diffuse, the oracle selection favours a stronger logarithmic correction. For $\eta = 0.035$ and $\eta = 0.075$, ROF exhibits visible concentric staircasing, whereas the Orlicz model produces a smoother transition and attains a lower MSE. Taken together, these observations suggest that the logarithmic correction can improve the reconstruction of genuinely diffuse transitions, while yielding results comparable to ROF in the sharp-interface regime.

5.3 NATURAL-IMAGE COMPARISON

Finally, we corrupt the standard Cameraman image, normalized to $[0, 1]$, by additive Gaussian noise with standard deviation 0.12. Figure 3 shows the oracle reconstructions. ROF selects $\alpha = 0.1$ and attains $\text{MSE } 1.491 \cdot 10^{-3}$ (PSNR 28.27 dB), while the Orlicz model selects $(\alpha, \gamma) = (0.1, 0.01)$ with $\kappa = 0.05425$ and attains $\text{MSE } 1.499 \cdot 10^{-3}$ (PSNR 28.24 dB). The quadratic H^1 model, with $\alpha = 1$, attains $\text{MSE } 2.586 \cdot 10^{-3}$ (PSNR 25.87 dB). The Orlicz and ROF reconstructions are therefore essentially tied on this example; consistently, the selected value of γ again places the Orlicz model close to its ROF limit.

To illustrate the influence of the logarithmic correction independently of the oracle parameter choice, we fix $\alpha = 0.1$ and retain the same noisy datum and estimated scale $\kappa \approx 0.05425$. Figure 4 compares the ROF reconstruction with the Orlicz reconstructions for $\gamma \in \{0.01, 0.1, 0.5, 1\}$. For $\gamma = 0.01$, the reconstruction is visually very close to ROF, consistent with the limiting behavior as $\gamma \downarrow 0$. Increasing γ produces progressively stronger smoothing, with a visible loss of edge sharpness and fine detail relative to ROF.

To compare the stronger smoothing at $\gamma = 1$ with quadratic H^1 regularization, we numerically select



Figure 3: Cameraman denoising with additive Gaussian noise of standard deviation 0.12. From left to right: ground truth, noisy datum, and the MSE-selected ROF, Orlicz, and quadratic H^1 reconstructions. $[0, 1]$.

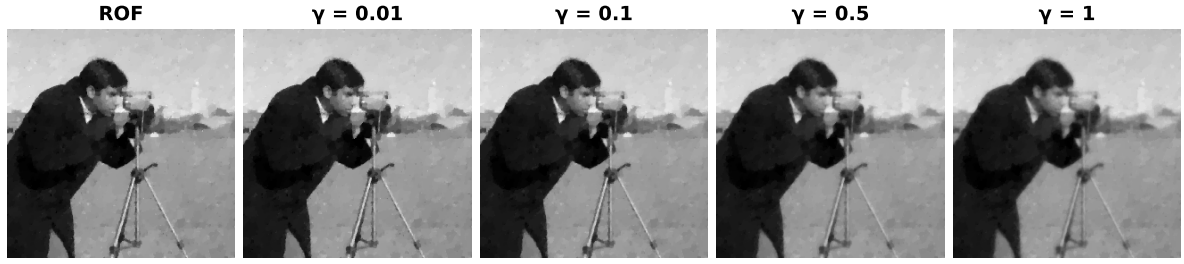


Figure 4: Influence of γ on Cameraman denoising at fixed $\alpha = 0.1$ and $\kappa \approx 0.05425$. From left to right: ROF and Orlicz reconstructions with $\gamma = 0.01, 0.1, 0.5$, and 1 .

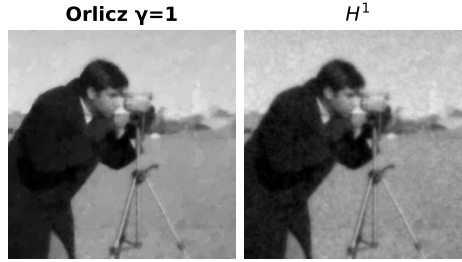


Figure 5: Cameraman denoising with Orlicz regularization at $\alpha = 0.1$, $\gamma = 1$, and $\kappa \approx 0.05425$ (left), and matched quadratic H^1 regularization at $\alpha \approx 3.15435$ (right).

the H^1 weight that minimizes the mean squared difference to the fixed Orlicz reconstruction, using a logarithmic parameter search followed by local refinement. The clean reference image is not used in this matching step. Figure 5 shows similar overall smoothing but distinct reconstruction characteristics. In this example, the Orlicz reconstruction appears more uniform in approximately homogeneous regions, such as the sky, whereas the matched H^1 reconstruction retains more visible small-scale intensity fluctuations. Differences are also apparent around the boundaries of the cameraman. The root mean squared difference between the reconstructions is approximately $1.295 \cdot 10^{-2}$.

6 CONCLUSION

The scaled $L \log L$ model provides a variationally consistent approximation of total variation denoising with additional control over the penalization of large gradients. The Orlicz framework establishes well-posedness and Fenchel duality despite the possible lack of reflexivity, while the explicit dual proximal map makes the model accessible to primal–dual splitting methods. The Γ -convergence result and convergence of minimizers give a rigorous connection to ROF as the logarithmic correction vanishes.

The numerical examples illustrate both the potential and the limitations of this correction. It can reduce staircasing in diffuse transitions, whereas stronger corrections at fixed α and κ can smooth out edges and fine detail in a natural image. The choice of the regularization parameters is therefore central to exploiting the model’s flexibility.

The theoretical framework permits spatially dependent integrands beyond the particular logarithmic density studied here. In particular, it already accommodates a spatially varying scale κ , whose numerical potential remains to be explored: local scale choices could adapt the onset of the logarithmic correction to different image regions. Further directions include extending the analysis and numerical realization to inverse problems such as deblurring, and developing parameter-choice rules that do not require access to a clean reference image.

DECLARATION OF AI ASSISTANCE

ChatGPT models GPT-5.4 and GPT-5.5 were used during the preparation of this manuscript as interactive assistants for drafting and revising text, improving the structure and exposition of mathematical arguments, suggesting alternative formulations and intermediate derivations. This assistance concerned in particular the introduction and conclusion, parts of the mathematical analysis in Sections 3 and 4, and the numerical presentation and implementation in Section 5. The mathematical ideas, formulation of the problem, main results, and underlying proof strategies were developed by the authors. In the mathematical sections, ChatGPT was used primarily to refine and streamline proofs, test alternative formulations of arguments, and improve their presentation. Suggestions generated by the model were checked against the relevant mathematical literature and independently verified, modified, or discarded by the authors as appropriate.

Codex was used for developing and debugging parts of the numerical code.

REFERENCES

- [1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, Oxford University Press, 2000, [doi:10.1093/oso/9780198502456.001.0001](https://doi.org/10.1093/oso/9780198502456.001.0001).
- [2] K. Bredies, K. Kunisch, and T. Pock, Total generalized variation, *SIAM Journal on Imaging Sciences* 3 (2010), 492–526, [doi:10.1137/090769521](https://doi.org/10.1137/090769521).
- [3] A. Chambolle and P. L. Lions, Image recovery via total variation minimization and related problems, *Numerische Mathematik* 76 (1997), 167–188, [doi:10.1007/s002110050258](https://doi.org/10.1007/s002110050258).
- [4] A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *Journal of Mathematical Imaging and Vision* 40 (2011), 120–145, [doi:10.1007/s10851-010-0251-1](https://doi.org/10.1007/s10851-010-0251-1).
- [5] Y. Chen, S. Levine, and M. Rao, Variable exponent, linear growth functionals in image restoration, *SIAM Journal on Applied Mathematics* 66 (2006), 1383–1406, [doi:10.1137/050624522](https://doi.org/10.1137/050624522).
- [6] C. Clason and T. Valkonen, *Introduction to Nonsmooth Analysis and Optimization*, volume 38 of MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics, 2026, [doi:10.1137/1.9781611978995](https://doi.org/10.1137/1.9781611978995).
- [7] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert W function, *Advances in Computational Mathematics* 5 (1996), 329–359, [doi:10.1007/bf02124750](https://doi.org/10.1007/bf02124750).
- [8] G. Dal Maso, *An Introduction to Γ -Convergence*, volume 8 of Progress in Nonlinear Differential Equations and their Applications, Birkhäuser Boston, Inc., Boston, MA, 1993, [doi:10.1007/978-1-4612-0327-8](https://doi.org/10.1007/978-1-4612-0327-8).
- [9] M. Eleuteri, P. Harjulehto, and P. Hästö, Bounded variation spaces with generalized Orlicz growth related to image denoising, *Mathematische Zeitschrift* 310 (2025), 26, [doi:10.1007/s00209-025-03731-9](https://doi.org/10.1007/s00209-025-03731-9).
- [10] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, revised edition, 2015.
- [11] P. Harjulehto and P. Hästö, *Orlicz Spaces and Generalized Orlicz Spaces*, volume 2236 of Lecture Notes in Mathematics, Springer, Cham, 2019, [doi:10.1007/978-3-030-15100-3](https://doi.org/10.1007/978-3-030-15100-3).

- [12] P. Harjulehto and P. Hästö, Double phase image restoration, *Journal of Mathematical Analysis and Applications* 501 (2021), 123832, [doi:10.1016/j.jmaa.2019.123832](https://doi.org/10.1016/j.jmaa.2019.123832).
- [13] K. Jalalzai, Some remarks on the staircasing phenomenon in total variation-based image denoising, *Journal of Mathematical Imaging and Vision* 54 (2016), 256–268, [doi:10.1007/s10851-015-0600-1](https://doi.org/10.1007/s10851-015-0600-1).
- [14] A. Kozek, Convex integral functionals on Orlicz spaces, *Comment. Math. Prace Mat.* 21 (1980), 109–135.
- [15] M. M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, volume 146 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, Inc., New York, 1991.
- [16] R. T. Rockafellar, Integrals which are convex functionals, *Pacific Journal of Mathematics* 24 (1968), 525–539, [doi:10.2140/pjm.1968.24.525](https://doi.org/10.2140/pjm.1968.24.525).
- [17] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* 60 (1992), 259–268, [doi:10.1016/0167-2789\(92\)90242-f](https://doi.org/10.1016/0167-2789(92)90242-f).
- [18] T. Unterberger, CU Orlicz numerics, Zenodo, 2026, [doi:10.5281/zenodo.22799164](https://doi.org/10.5281/zenodo.22799164), <https://doi.org/10.5281/zenodo.22799164>. Version 0.1.1.