Jonathan Chirinos-Rodríguez (jonathan-eduardo.chirinos-rodriguez@irit.fr)
Emmanuel Soubies (emmanuel.soubies@irit.fr)
Luca Calatroni (luca.calatroni@unige.it)
Silvia Villa (silvia.villa@unige.it)
October 14, 2025

Master internship proposal
—
Learning Optimal Regularization Parameters for Inverse Problems

Context

An inverse problem consists in recovering a solution u^* from a set of (possibly noisy) linear measurements Au^*, where A is an operator modeling the measurements acquisition process. This task appears in a broad range of practical problems in engineering, signal processing, medical imaging or computer vision. In mathematical terms, an inverse problem can be modeled as

$$x = Au^* + \varepsilon,$$

where x denotes the available measurements and ε is a deterministic quantity modeling the possible presence of noise in the observations. The task of finding u^* from the knowledge of x becomes hard when the problem is ill-posed. Regularization theory [1, 2] offers a systematic way to address ill-posedness by providing stable approximations of the inverse. A classical approach for restoring well-posedness is based on finding solutions of the following variational problem

$$\min_u \ell(Au, x) + \lambda R(u), \tag{1}$$

for some $\lambda \in (0, +\infty)$. The term $\ell(A\cdot, x)$ is known as the *data-fitting* term, and constraints the solution to remain close to the available measurements. The function R, referred to as the *regularization function*, incorporates prior knowledge about the solution into the problem formulation. Finally, the scalar $\lambda \in (0, +\infty)$ is known as the *regularization parameter*. This parameter allows to choose the relative importance of the data-fitting term and the regularization function, thereby influencing the quality of the recovery results. Consequently, a proper selection of the regularization parameter is essential for achieving optimal reconstruction outcomes. To this day, the task of selecting a suitable $\lambda > 0$ remains a challenging problem.

The above strategy for solving inverse problems can be viewed as a *model-based* technique, relying on a mathematical model with well-established properties. For instance, variational methods (1) have for a long time achieved state-of-the-art results [3] in imaging problems. Notwithstanding this, *data-driven* methodologies have gained significant attention in recent years, since they demonstrate improved performance in various practical scenarios while overcoming some challenges of classical methods (see [4] and references therein). The starting point of data-driven approaches consists in assuming that a finite set of pairs of measurements and exact solutions $(\bar{x}_1, \bar{u}_1), \ldots, (\bar{x}_n, \bar{u}_n)$, $n \in \mathbb{N}$, is available. This *training set* is then used to define, or refine, a regularization strategy to be applied to any future observation $\bar{x}_{\text{new}}$, for which an exact solution is not known.

Objectives

The main objective of the internship is to extend, both from a theoretical and a practical point of view, the work [5], which analyzes the following data-driven approach: in the variational model (1), we fix the regularizer R and we aim to learn the regularization parameter $\lambda \in (0, +\infty)$ from the given training set. This approach is based on the following bilevel optimization problem (see, for instance, [6, 7]). Given a set $\Lambda \subset (0, +\infty)$, we select the regularization parameter as

$$\widehat{\lambda} \in \arg\min_{\lambda \in \Lambda} \frac{1}{n} \sum_{i=1}^{n} \|u_\lambda^i - \bar{u}_i\|^2,$$

where $u_\lambda^i := u_\lambda(\bar{x}_i)$ is such that

$$u_\lambda(\bar{x}_i) \in \arg\min_{u \in \mathcal{U}} \ell(Au, \bar{x}_i) + \lambda R(u),$$

see [8, Chapter 3] and references therein. We then utilize $\widehat{\lambda}$ as regularization parameter for subsequent instances of the same inverse problem: given $\bar{x}_{\text{new}}$, we consider $u_{\widehat{\lambda}}(\bar{x}_{\text{new}})$ as an approximation of $\bar{u}_{\text{new}}$.

To start, the student is expected to study the necessary tools for understanding the aforementioned work, which range between supervised learning theory, inverse problems, and convex optimization. Most of these tools can be found in [9]. Next, we mention some of the directions that could be followed:

- **Parametrized regularizers.** The framework studied in [5] can be generalized by fixing instead regularizers R that are parametrized by a vector $\theta = (\theta_1, ..., \theta_k) \in \Theta \subseteq \mathbb{R}^k$, $k \in \mathbb{N}$, and so of the form $R = R(\cdot, \theta)$. The variational model in this cases reads as

$$\min_{u \in \mathcal{U}} \ell(Au, x) + R(u, \theta).$$

  Here, the objective is to learn the vector of parameters $\boldsymbol{\theta}$. Artificial neural networks that are parametrized by a large set of scalars have been considered, e.g. the Total Deep Variation [10]. A similar approach to the one defined above can be used to find the optimal $\boldsymbol{\theta}$ given a finite training set of input/output pairs.
- **Space-varying Total Variation regularization.** Let Ω represent the domain of the image u^* and let $\Omega_k \subseteq \mathcal{U}$, $k = 1, \ldots, m$, $1 < m < +\infty$, denote bounded areas of the image u^* such that $\Omega_i \cap \Omega_j = \emptyset$ and $\cup_k \Omega_k = \Omega$. With this, we may consider, for every $k = 1, \ldots, m$, the Total Variation regularizer (see, e.g., [11, 12]) restricted to the set Ω_k, informally defined as

$$\mathrm{TV}_{\Omega_k}(u) := \sum_{(i,j) \in \Omega_k} \|(Du)_{i,j}\|_{1,2},$$

  where Du denotes the vector of discrete horizontal and vertical derivatives, and consider the following problem

$$\min_u \frac{1}{2}\|u - x\|_2^2 + \sum_{k=1}^{m} \lambda_k \mathrm{TV}_{\Omega_k}(u).$$

  The above formulations allows for different regularization parameters λ_k in different areas of the image u^*, and can be seen as a weighted total variation regularizer. In this vein, several works have been produced, see e.g. [13]. Yet, few theoretical guarantees have been provided from a statistical learning viewpoint.

Practical aspects

We are looking for a highly motivated student, with a background in applied mathematics (optimization, probability and statistics, geometry) and/or computer science (signal/image processing). The theoretical modeling of the learning schemes will be complemented by extensive numerical validation performed using the DeepInv Python library (`https://deepinv.github.io/deepinv/`).

The internship will be co-supervised by Jonathan Chirinos-Rodríguez (Post-doctoral researcher, IRIT), Luca Calatroni (Associate Professor, University of Genova) and Emmanuel Soubies (CNRS researcher, IRIT), at the IRIT laboratory in Toulouse, France. The intern will be granted a research scholarship of around ∼670 euros/month.

Upon successful completion of the internship and subject to securing funding, PhD opportunities may be available. Do not hesitate to contact us for more information.

Bibliography

[1] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111, 2018.

[2] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[3] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, 2009.

[4] Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.

[5] Jonathan Chirinos-Rodríguez, Ernesto De Vito, Cesare Molinari, Lorenzo Rosasco, and Silvia Villa. On learning the optimal regularization parameter in inverse problems. *Inverse Problems*, 40(12):125004, nov 2024.

[6] J. C. De los Reyes, C.-B. Schönlieb, and T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *Journal of Mathematical Imaging and Vision*, 57(1):1–25, 2017.

[7] Karl Kunisch and Thomas Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM Journal on Imaging Sciences*, 6(2):938–983, 2013.

[8] Caroline Crockett and Jeffrey A. Fessler. Bilevel methods for image reconstruction. *Foundations and Trends® in Signal Processing*, 15(2-3):121–289, 2022.

[9] Jonathan Chirinos-Rodríguez. *Machine Learning Techniques for Inverse Problems*. Phd thesis, Università degli studi di Genova, May 2024.

[10] Erich Kobler, Alexander Effland, Karl Kunisch, and Thomas Pock. Total deep variation for linear inverse problems. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7546–7555, 2020.

[11] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[12] Antonin Chambolle and Pierre-Louis Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.

[13] Monica Pragliola, Luca Calatroni, Alessandro Lanza, and Fiorella Sgallari. On and beyond total variation regularization in imaging: The role of space variance. *SIAM Review*, 65(3):601–685, 2023.