



**HAL**  
open science

# Volume-Preserving Geometric Shape Optimization of Dirichlet Energy Using Variational Neural Networks

Amaury Bélières–Frendo, Victor Michel-Dansac, Emmanuel Franck, Yannick Privat

► **To cite this version:**

Amaury Bélières–Frendo, Victor Michel-Dansac, Emmanuel Franck, Yannick Privat. Volume-Preserving Geometric Shape Optimization of Dirichlet Energy Using Variational Neural Networks. 2024. hal-04663197

**HAL Id: hal-04663197**

**<https://hal.science/hal-04663197>**

Preprint submitted on 26 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

# Learning-based geometric shape optimization of the Dirichlet energy

A. Bélières Frendo\*    E. Franck†    Victor Michel-Dansac‡    Yannick Privat§¶

July 26, 2024

## Abstract

In this work, we explore the numerical solution of geometric shape optimization problems using neural network-based approaches. This involves minimizing a numerical criterion that includes solving a partial differential equation with respect to a domain, often under geometric constraints like constant volume. Our goal is to develop a proof of concept using a flexible and parallelizable methodology to tackle these problems. We focus on a prototypical problem: minimizing the so-called Dirichlet energy with respect to the domain under a volume constraint, involving a Poisson equation in  $\mathbb{R}^2$ . We use physics-informed neural networks (PINN) to approximate the Poisson equation's solution on a given domain and represent the shape through a neural network that approximates a volume-preserving transformation from an initial shape to an optimal one. These processes are combined in a single optimization algorithm that minimizes the Dirichlet energy. One of the significant advantages of this approach is its parallelizable nature, which makes it easy to handle the addition of parameters. Additionally, it does not rely on shape derivative or adjoint calculations. Our approach is tested on Dirichlet and Robin boundary conditions, parametric right-hand sides, and extended to Bernoulli-type free boundary problems. The source code for solving the shape optimization problem is open-source and freely available.

**Keywords:** Shape optimization; Dirichlet energy; volume constraint; physics-informed neural networks; symplectic neural networks

**AMS Classification:** 49M41; 49Q10; 65K10; 68T07

## 1 Introduction

Throughout this article, let  $V_0 > 0$  be a fixed real number and  $n \in \mathbb{N}^*$ . We will denote by  $\mathbb{B}^n \subset \mathbb{R}^n$  the  $n$ -dimensional open ball with volume  $V_0$  and by  $\mathbb{S}^{n-1}$  its boundary. For simplicity, we will omit explicit notation of the dependency on  $V_0$ , as all the considered open sets share the same volume.

---

\*IRMA, Université de Strasbourg, CNRS UMR 7501, Inria, 7 rue René Descartes, 67084 Strasbourg, France. (amaury.belieres@math.unistra.fr)

†Inria, IRMA, Université de Strasbourg, CNRS UMR 7501, 7 rue René Descartes, 67084 Strasbourg, France. (emmanuel.franck@inria.fr)

‡Inria, IRMA, Université de Strasbourg, CNRS UMR 7501, 7 rue René Descartes, 67084 Strasbourg, France. (victor.michel-dansac@inria.fr)

§Université de Lorraine, CNRS, Institut Elie Cartan de Lorraine, Inria, BP 70239 54506 Vandœuvre-lès-Nancy Cedex, France. (yannick.privat@univ-lorraine.fr).

¶Institut Universitaire de France (IUF)

## 1.1 Shape Optimization using Learning Techniques

Shape optimization is a field of mathematics in which we seek to determine, if it exists, the shape of a domain that minimizes some numerical criterion. There are many such problems, ranging from the most fundamental to the most applied [11, 18]. For example, on the one hand, optimizing the eigenvalues of the Laplacian or Schrödinger operators has a long history, dating back to at least as far as the work of Lord Rayleigh or Kac [23]. The aim is to understand the very implicit links between the eigenmodes of these operators and the geometry. On the other hand, from a practical point of view, shape optimization is crucial for engineering applications. Let us mention, for instance, geometry optimization to increase the performance of a mechanical component, typically the optimal design of an aircraft wing [1, 26].

In this article, we focus on the numerical implementation of shape optimization problems. There are many tools and algorithms available for determining (at least local) minimizers of such problems. Many approaches are based on the use of an ad-hoc notion of differential, either called shape derivative or derivative in the sense of Hadamard, enabling the implementation of gradient methods (see e.g. [2]).

However, such numerical methods often suffer from several drawbacks. Indeed, they are based on a derivative calculation that can become highly complex, for example when the criterion involves the solution of a multiphysics model. This approach often requires the determination of an adjoint problem and a descent step, which can make it very costly in terms of computation time and memory allocation. Moreover, since shape optimization problems frequently have numerous local minimizers, these approaches tend to be highly local. A major drawback of these methods is their general lack of parallelizability. This makes them particularly time-consuming and poorly suited for scenarios where model parameters need to be varied. In contrast, the method we introduce is highly parallelizable, offering a significant advantage in efficiency and adaptability. Finally, such approaches are not suitable for all physical models. One example is models that are not “well-posed”, such as the turbulent Navier-Stokes equations.

Our goal is to introduce a methodology based on the use of neural networks to overcome these difficulties. Let us mention [28], in which a first step is taken in this direction: the equation of state (hyper-elastic problem) is solved using the feed-forward neural network and the so-called *adjoint problem* is obtained through the backpropagation of the network. A few authors have proposed shape optimization algorithms based on neural networks, such as PINNTOs (Physics-Informed Neural Network-based Topology Optimization) [20] based on the Solid Isotropic Material with Penalization method, where the linear elasticity equations are solved with a PINN (Physics-Informed Neural Network), or [33] trained PINNs to perform shape optimization for Maxwell’s equations using coordinate projection. We can also mention [10] that solves the Bernoulli free boundary problem with a PINNs, rephrasing it like an overdetermined PDE constrained by the measure of the positive part of its solution.

The aim of this article is to propose a proof of concept, highlighting the potential of this approach. We will validate it on a very simple model: Dirichlet energy minimization with homogeneous Dirichlet or Robin boundary condition. The aim is to optimize, with respect to the domain, the “natural” energy associated with a Poisson equation of the form

$$-\Delta u = f \quad \text{in } \Omega,$$

where  $\Omega \in \mathbb{R}^n$  is an open bounded connected set. We add a classical volume constraint, which models a manufacturing cost:  $|\Omega| \leq V_0$ , with  $V_0 > 0$ , a fixed parameter. We deliberately focus on this prototypical problem, which has been the subject of much work and is now very well understood.

Note that there are various types of shape optimization problems, including parametric, geometric, and topological problems. The latter allow for topological changes (such as the number of holes). In this article, we focus exclusively on *geometric optimization problems*, where the desired shape must be homeomorphic to a given initial shape, such as a ball.

To describe our strategy, note that our shape optimization problem consists in minimizing the energy associated to a PDE while also optimizing the shape of the space domain. To that end, we introduce two neural networks. The first one, a physics-informed neural network (PINN, see [30]), approximates the PDE solution. The second one, a symplectic neural network (SympNet, see [22]), takes care of the domain. Indeed, symplectic maps are examples of volume-preserving diffeomorphisms. Hence, they are suitable candidates for representing the optimal shape, which will be defined as the image, by the SympNet, of a given initial domain. As a consequence, in our strategy, the constant volume constraint is automatically satisfied, and it does not need to be added to the loss function. Moreover, regarding the PDE approximation, boundary conditions will be directly imposed in the network rather than in the loss function, in the spirit of [25]. Both of these remarks mean that we will only have to solve an optimization problem on the trainable parameters of the PINN and of the SympNet, with a single loss function, corresponding to the Dirichlet energy of the problem.

Conventional methods necessitate running a new simulation for each set of physical parameters or source terms. In contrast, a neural network can learn a parametric family of optimal shapes for a corresponding family of physical parameters or source terms, all within a time frame comparable to solving a single non-parametric problem.

Alongside this paper, we provide a turnkey open source code to solve such shape optimization problems. The code comes as a ready-to-use Python library that the user can freely download on GitHub<sup>1</sup>, with a detailed documentation. The code architecture is inspired by the AvaFrame python framework [32]. The code is divided into 3 computational modules: `com1PINNs`, for PDE resolution with PINNs, `com2SympNets` for learning a given shape with SympNets, and `com3DeepShape` for learning-based shape optimization. To run each computational module, the user will find appropriate run scripts in the folder `examples`. In each run script, the user will find parameters that can be freely modified: number of layers and neurons in the network, shape of the initial domain, name of backup file, boundary conditions, source term. Once these settings have been made, the code can be launched.

The paper is organized as follows: First, Section 1.3 introduces the shape optimization problem and reviews some theoretical results. Next, Section 2 presents the neural networks and the joint optimization algorithm based on the use of appropriate symplectomorphisms. Section 3 focuses on the numerical validation of the code with Dirichlet boundary conditions, while Section 4 addresses Robin boundary conditions.

## 1.2 Source code

The source code for solving the shape optimization problem is open and available at the following link:

<https://github.com/belieresfrendo/GeSONN>.

A comprehensive documentation explaining how to use it is also provided.

## 1.3 Presentation of the shape optimization problem

We first introduce a problem involving a PDE with Dirichlet conditions, but our approach generalizes to other conditions. We will also address the case of Robin conditions subsequently.

Let  $\Omega$  be an open bounded connected set in  $\mathbb{R}^n$  and  $f \in H^{-1}(\Omega)$ . In the whole article, we will denote by  $u_{\Omega}^f$  the unique solution in  $H_0^1(\Omega)$  of the Poisson problem

$$\begin{cases} -\Delta u_{\Omega}^f = f & \text{in } \Omega, \\ u_{\Omega}^f = 0 & \text{on } \partial\Omega. \end{cases} \quad (1)$$

---

<sup>1</sup>See Section 1.2



For simplicity, we will only deal with homogeneous Dirichlet boundary conditions on  $\partial\Omega$ . It is well-known that  $u_\Omega^f$  can also be defined as the unique solution of the variational problem, find  $u_\Omega^f \in H_0^1(\Omega)$ , such that

$$\mathcal{J}(\Omega, u_\Omega^f) = \inf\{\mathcal{J}(\Omega, u), u \in H_0^1(\Omega)\}, \quad (2)$$

with

$$\mathcal{J}(\Omega, u) = \frac{1}{2} \int_\Omega |\nabla u|^2 - \langle f, u \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}, \quad \forall u \in H_0^1(\Omega), \quad (3)$$

where  $\langle \cdot, \cdot \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}$  denotes the duality bracket between  $H^{-1}(\Omega)$  and  $H_0^1(\Omega)$ . Recall that the minimization problem above can be interpreted as an energetic formulation of the PDE (1). Furthermore, if  $f$  enjoys additional regularity (say  $f \in L^2(\Omega)$ ), the PDE (1) is satisfied at least pointwisely. Let us now introduce the so-called Dirichlet energy  $\mathcal{E}$ , a shape functional we will deal with throughout this article. It is given by

$$\mathcal{E}(\Omega) := \inf_{u \in H_0^1(\Omega)} \mathcal{J}(\Omega, u) \quad (4)$$

and it is notable that  $\mathcal{E}(\Omega) = \mathcal{J}(\Omega, u_\Omega^f)$ . Minimizing the Dirichlet energy within sets of given volume is a prototypical problem in shape optimization. It reads:

$$\boxed{\inf\{\mathcal{E}(\Omega), \Omega \text{ bounded open set of } \mathbb{R}^n, \text{ such that } |\Omega| = V_0\}}. \quad (5)$$

The analysis of such a problem is a long story. We refer for instance to [18] for a detailed review of results, related to the existence of an optimal shape, but also to geometric properties of minimizers.

**Remark 1** (Some comments about existence and regularity of optimal shapes). Existence issues in shape optimization are generally difficult. It is standard to introduce a relaxed formulation of the initial problem posed among open sets. This often leads to considering a set of admissible forms living in the set of quasi-open sets. Recall that a set  $A$  is said to be quasi-open if there exist sets  $\omega$  of capacity as small as desired, such that  $A \setminus \omega$  can be extended into an open by adding points of  $\omega$ .

In the case we are interested in, it is easy, by ad-hoc extension of the definition of the functional space  $H_0^1(\Omega)$  to quasi-open  $\Omega$ , to establish the existence of an optimal solution for the problem

$$\inf\{\mathcal{E}(\Omega), \Omega \text{ quasi open set of } D, \text{ such that } |\Omega| = V_0\},$$

where  $D$  denotes a given compact set of  $\mathbb{R}^n$ . We refer to [18, Chapter 4] for detailed explanations about such notions.

It is worth noting that in some cases (e.g. when considering shape optimization problems involving Dirichlet eigenvalues), the box constraint on  $D$  can be removed, using concentration-compactness arguments originally established by Lions. Once such an existence result has been established, an a posteriori analysis can sometimes be carried out to establish the existence of regular openings. We refer, for example, to [17, Chapters 2, 3].

In this article, the question of existence is not central. From the above-mentioned references, it is well known that, even if we consider families of domains included in a fixed compact  $D$ , there exists an optimal shape among the quasi-open sets. In the following, we will no longer mention this question, as our main problem here concerns the numerical determination of solutions to this problem, using techniques based on the use of neural networks.

Another very useful result is the following characterization of optimal shapes.

**Theorem 1.** ([18, Section 6.1.3]) Let  $f \in L^2_{loc}(\mathbb{R}^n)$ . If  $\Omega$  is a solution to (5) with a  $C^2$  boundary, then there exists  $c > 0$  such that

$$|\nabla u^f_\Omega| = c \text{ on } \partial\Omega. \tag{6}$$

It is noteworthy that this result provides another angle of attack for shape optimization problems. Indeed, rather than minimizing the shape functional  $\mathcal{E}$  given by (4) with respect to shapes  $\Omega$  of prescribed volume, we can seek to numerically solve the overdetermined PDE :

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \\ |\nabla u| = c & \text{on } \partial\Omega, \end{cases} \tag{7}$$

where  $c > 0$  is a given parameter implicitly encoding the volume constraint.

In what follows, we will use the optimality condition provided by **Theorem 1** as a criterion for validating the shapes obtained. Indeed, if the algorithm we present leads to a domain  $\Omega_0$ , we will consider it a good candidate for solving Problem (5) whenever Condition (6) is satisfied.

The overdetermined equation (7) will also serve as the starting point for solving Bernoulli problems, which are addressed in section 3.4.

## 2 Solving the shape optimization problem with Neural Networks

As outlined in **Section 1.3**, we will focus here on the Poisson problem (1). To identify areas for improvement, we briefly describe the main loop of classical shape optimization algorithms.

Initially, several PDEs are numerically solved to compute the state, the adjoint state if necessary, the shape gradient, and the gradient-descent step. Next, the computational domain is deformed according to the direction of the shape gradient. This is done through meshing, which requires a procedure known as *extension-regularization*, whose implementation is complex. This process is iterated until convergence, which can take hours, days, or even weeks for industrial applications [3, 1]. Essentially, this method is not parallelizable. However, these challenges are not unavoidable and depend on the problem’s formulation.

In what follows, we tackle these issues using neural networks (NNs), which present several advantages over traditional numerical methods for PDEs. For instance, by combining neural networks with the Monte Carlo algorithm [8] for approximating integrals in the loss function, we can efficiently handle parameter-dependent problems on complex domains. Moreover, NNs allow for joint gradient descent on multiple interdependent networks. The shape derivative computation is effectively hidden behind an automatic differentiation process of the network. This is significant because calculating the shape derivative is complex and sometimes impossible for certain multiphysics problems. This advancement opens up new prospects for interdisciplinary research, especially in the life sciences. This means we can simultaneously train one network to represent the PDE solution and another to represent the computational domain, resulting in a parallelizable algorithm for solving shape optimization problems.

We now present a methodology for developing a neural network algorithm to solve the shape optimization problem (5). This approach leverages DeepRitz [12] to represent the solution to the Poisson problem (1), and SympNets [22] to represent the computational domain.

### 2.1 PINNs and DeepRitz

Let us outline the architecture of the neural network representing the solution to the PDE within the computational domain. We will first describe a fully connected neural network and then explain how to

adapt it for solving PDEs, specifically the Poisson problem (1). For a comprehensive introduction to fully connected neural networks, PINNs, and DeepRitz, see [14, 30, 12]. In this work, the neural networks are tailored to minimize the Dirichlet energy (4). A distinctive feature of these networks is the strictly physical nature of their loss function; no data is used for training the networks.

### 2.1.1 Fully connected NNs

Consider again the problem (1), defined on an open set  $\Omega$ . The main idea behind fully-connected neural networks (NNs) is to represent the solution  $u \in H_0^1(\Omega)$  of the forward problem (1) by  $u_\theta \in C^\infty(\mathbb{R}^n)$ , a composition of nonlinear parametric functions that takes  $x \in \Omega$  as input and returns an approximation of  $u(x)$  as output, see [30]. The network and associated notation are depicted in Figure 1. The parameters  $\theta = \{W^k, b^k\}_{k=1}^\ell$  of the NN, called trainable weights, are then optimized by minimizing a loss function.

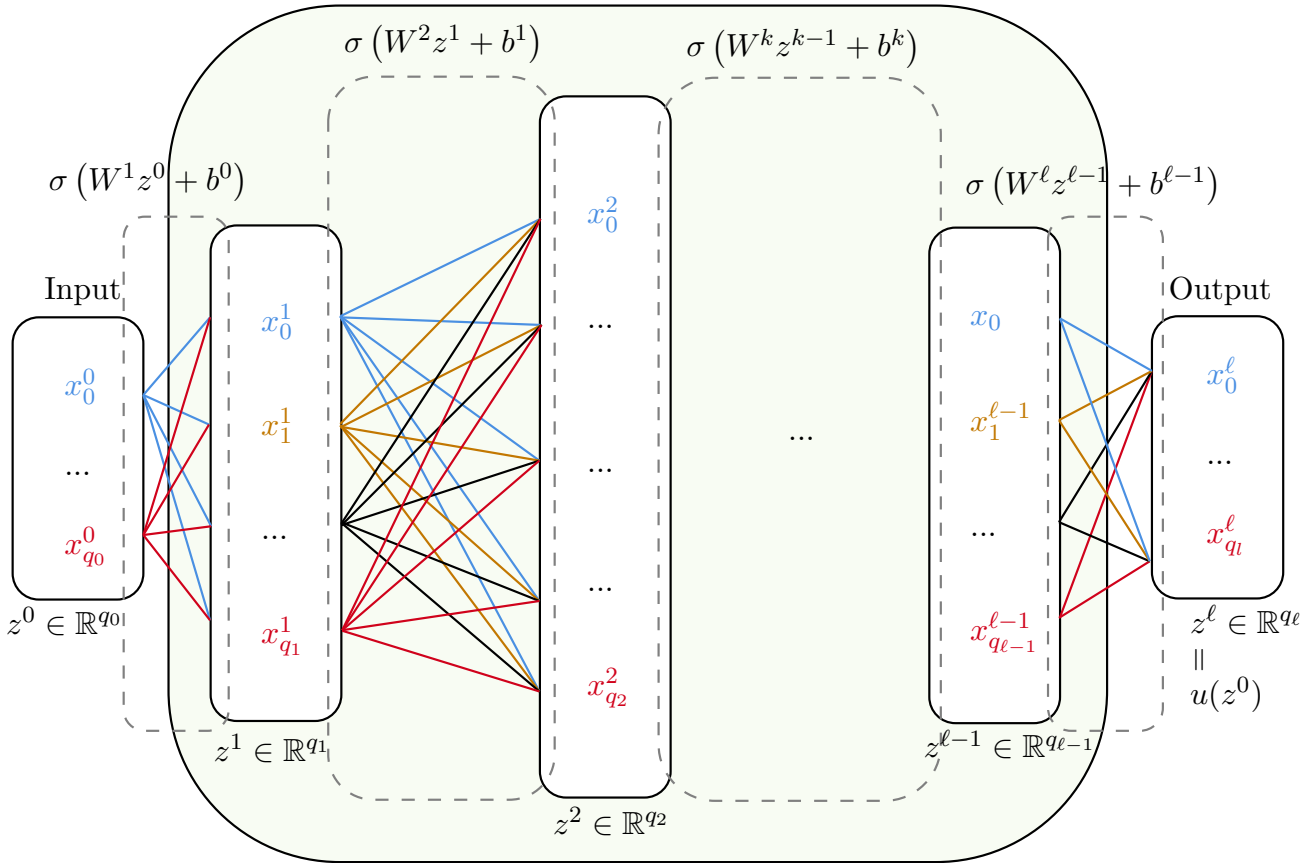


Figure 1: Diagram of a fully connected NN, where the output of each layer is a vector  $z_k \in \mathbb{R}^{q_k}$ , with  $q_k$  the number of neurons of the layer  $k$ . Computing  $z_k$  involves a weight matrix  $W^k \in \mathcal{M}_{q_k, q_{k-1}}(\mathbb{R})$ , a bias vector  $b^k \in \mathbb{R}^{q_k}$  and a nonlinear activation function  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  applied componentwise.

To obtain a solution respecting the boundary conditions, we introduce two functions  $\alpha, \beta \in C^1(\mathbb{R}^n, \mathbb{R})$  such that  $\alpha$  vanishes on  $\partial\Omega$ , and  $\beta$  is equal, on  $\partial\Omega$ , to the boundary condition of the Poisson problem (1).

Of course, for homogeneous Dirichlet boundary conditions,  $\beta = 0$  is suitable. Then, for all  $x \in \mathbb{R}^n$ ,

$$v_\theta(x) = \alpha(x) u_\theta(x) + \beta(x) \tag{8}$$

satisfies the approximation of the PDE given by the PINN [25, 13]. For example, if  $v_\theta$  has to satisfy Dirichlet homogeneous conditions on  $\mathbb{S}^1$ , the unit sphere of  $\mathbb{R}^2$ , it is sufficient to define, for  $x = (x_1, x_2) \in \mathbb{B}^2$ ,

$$\begin{cases} \alpha(x) = 1 - x_1^2 - x_2^2, \\ \beta(x) = 0. \end{cases} \tag{9}$$

This can of course be generalized for other domains and boundary conditions.

The network we train is still  $u_\theta$ , but the solution is represented by  $v_\theta$ , which is forced by construction to satisfy the boundary conditions on  $\partial\Omega$ . Imposed that way, the boundary conditions are easier to implement, and they are less reliant on parameter adjustment than if we had added a penalization term to the loss function.

### 2.1.2 The Dirichlet energy, an appropriate loss function

To obtain the optimal network parameters  $\theta^*$ , we minimize a loss function during the training process. Let us design this loss function. Given that the neural network will be trained without any data, the loss function must incorporate the Dirichlet energy. This is because minimizing the Dirichlet energy yields the solution to the underlying PDE, as indicated by (2).

**Remark 2.** The loss function  $\mathcal{J}$  defined in (3), coupled to a fully connected NN, is a particular case of a DeepRitz network introduced in [12]. If we had chosen the residual of the PDE as the loss function, it would have been a PINN, as defined in [30]. Nevertheless, for simplicity, we will call our NN representing the solution of the Poisson problem (1) a PINN, because of its essential physics-informed character.

As the computational domain for the Poisson problem (1) can have a complex topology, meshing it can be cumbersome and time-consuming. Therefore, the integral in the loss function  $\mathcal{J}$  is approximated with the Monte-Carlo method [8], rather than with a classical quadrature method. This leads us to defining  $\mathcal{J}_N$ , a discrete quadrature of  $\mathcal{J}$ , given by

$$\mathcal{J}_N(\theta; \{x^i\}_{i=1}^N) = \frac{V_0}{N} \sum_{i=1}^N \left\{ \frac{1}{2} |\nabla v_\theta(x^i)|^2 - f(x^i) v_\theta(x^i) \right\}, \tag{10}$$

where  $\{x^i\}_{i=1}^N \in \Omega^N$  are  $N$  collocation points on which to evaluate the loss function (10), and where  $u_\theta$  and  $v_\theta$  are defined in Section 2.1.1. Recall from [8] that the convergence speed of the Monte-Carlo algorithm does not depend on the dimension of the problem, but only on  $N$ . Since  $u_\Omega^f$  solving the Poisson problem (1) is the unique minimizer of  $\mathcal{J}$ , it is expected that  $v_\theta$  is a fair approximation of the minimizer of  $\mathcal{J}_N$ . The gradient of the loss function (10) is subsequently computed using the PyTorch library [29].

Starting with randomly initialized parameters  $\theta$ , we use a gradient descent method to find optimal parameters  $\theta^*$ , i.e., ones that minimize (10). At iteration  $k$ , the parameters are updated with the Adam Optimizer [24], a stochastic gradient descent algorithm. Once the training procedure is complete, the trained NN  $v_{\theta^*}$  provides an approximation of the solution to the forward problem (1).

### 2.1.3 Parametric problems

Note that the source term of the Poisson problem can be selected as a parametric function. Upon completing the training procedure, and with a slightly increased but comparable computation time, the Poisson problem can be solved for various parameter values within a given set.

Let  $n_\mu$  be the number of parameters in the problem. We denote by  $\mathbb{M} \subset \mathbb{R}^{n_\mu}$  the parameter space. The parametric neural network then takes as input both a position within the computational domain  $\Omega$  and parameters from the parameter space  $\mathbb{M}$ . Let  $f^p : \Omega \times \mathbb{M} \rightarrow \mathbb{R}$  represent the parametric source term.

The solution  $u^p : \Omega \times \mathbb{M} \rightarrow \mathbb{R}$  of the parametric Poisson problem satisfies

$$\begin{cases} -\Delta u^p(x; \mu) = f^p(x; \mu), & \text{for } (x; \mu) \in \Omega \times \mathbb{M}; \\ u^p(x; \mu) = 0, & \text{for } (x; \mu) \in \partial\Omega \times \mathbb{M}. \end{cases} \quad (11)$$

As an example, the space of parameters could be  $\mathbb{M} = \mathbb{R} \times \mathbb{R}$ , and  $f^p$  could be given by

$$f^p : (x = (x_1, x_2); \mu = (\mu_1, \mu_2)) \mapsto \exp\left(1 - \left(\frac{x_1}{\mu_1}\right)^2 - \left(\frac{x_2}{\mu_2}\right)^2\right).$$

At the end of the training procedure, the NN will be such that  $(x; \mu) \in \Omega \times \mathbb{M} \mapsto v_\theta(x; \mu)$  approximates the parametric Poisson problem (11) a.e. in  $\Omega \times \mathbb{M}$ . The parametric loss function  $\mathcal{J}_N^p$  then becomes

$$\mathcal{J}_N^p(\theta; \{x^i, \mu^i\}_{i=1}^N) = \frac{V_0}{N} \sum_{i=1}^N \left\{ \frac{1}{2} |\nabla v_\theta^p(x^i; \mu^i)|^2 - f^p(x^i; \mu^i) v_\theta^p(x^i; \mu^i) \right\}, \quad (12)$$

where  $\{x^i, \mu^i\}_{i=1}^N$  are  $N$  collocation points in  $\Omega \times \mathbb{M}$ . As defined in [Section 2.1.1](#), we set  $v_\theta^p = u_\theta^p \alpha^p + \beta^p$ , where the functions  $\alpha^p, \beta^p \in C^1(\mathbb{R}^2 \times \mathbb{M}, \mathbb{R})$  are such that  $\alpha^p$  vanishes on  $\partial\Omega$ , and  $\beta^p$  is equal, on  $\partial\Omega$ , to the boundary condition of the Poisson problem (1). Note that  $\mathcal{J}_N^p(\theta; \{x^i, \mu^i\}_{i=1}^N)$ , for  $\{x^i, \mu^i\}_{i=1}^N \in \Omega \times \mathbb{M}$ , is a quadrature of the limit functional  $\mathcal{J}^p$ , defined for  $v^p = u^p \alpha^p + \beta^p$ , with  $u^p : \Omega \times \mathbb{M} \mapsto \mathbb{R}$ , by

$$\mathcal{J}^p(u^p) = \int_{\Omega \times \mathbb{M}} \left( \frac{1}{2} |\nabla v^p(x; \mu)|^2 - f^p(x; \mu) v^p(x; \mu) \right) dx d\mu.$$

## 2.2 SympNets

In the previous section, we addressed solving a PDE in a given domain  $\Omega$ . Recall that our goal is to minimize the Dirichlet energy with respect to both the solution  $u$  and the domain  $\Omega$ . This means minimizing also  $\mathcal{J}(\Omega, u)$  with respect to  $\Omega$  among the open sets in  $\mathbb{R}^n$  with volume  $V_0$ . To achieve this, we must parameterize the set of open subsets of  $\mathbb{R}^n$ . For practical and simplicity reasons, we assume the optimal set is connected and enjoys several regularity properties. This will lead us to introduce appropriate invertible differentiable transformations of  $\mathbb{S}^{n-1}$ .

### 2.2.1 Using symplectic maps for shape optimization

Our objective is to develop a neural network representation for a shape. To achieve this, we seek an optimal volume-preserving diffeomorphism that maps  $\mathbb{S}^{n-1}$  to a shape that minimizes the shape criterion. By focusing on finding the optimal diffeomorphism, this method aligns with "geometric optimization" as it allows us to refine the shape's geometry while preserving its topological characteristics (in particular its gender).

Rather than considering generic diffeomorphisms, we focus on the special case of symplectic maps<sup>2</sup>, which are  $C^1$ -diffeomorphisms which also preserve volume.

This property is particularly relevant in our context since the shape optimization problem involves a volume constraint. However, symplectic maps are only defined for even-dimensional spaces. This is consistent with our focus of optimizing shapes in  $\mathbb{R}^2$ . Note that, in the specific case of  $\mathbb{R}^2$ , the symplectic form coincides with the volume form. This implies that, in  $\mathbb{R}^2$ , a diffeomorphism is volume-preserving if and only if it is a symplectic map.

The following section is dedicated to presenting the architecture of the NN that represents the shape of the domain and detailing the training process to approximate any symplectic map.

### 2.2.2 Symplectic maps and NNs

A SympNet is a NN with a symplectic structure capable of approximating any symplectic map. SympNets were first introduced in [22] to approximate the flows of Hamiltonian systems.

Our objective is now to train a SympNet to replicate any continuous and differentiable transformation  $\mathcal{T}$  of  $\mathbb{R}^{2d}$ . This section does not cover shape optimization; instead, it focuses on defining SympNets. Specifically, we aim to explain how to train a SympNet to learn the transformation of the  $2d - 1$ -dimensional sphere into a given shape in  $\mathbb{R}^{2d}$ , generated by a given symplectic map  $\mathcal{T}$ . Section 2.3 is dedicated to solving a PDE in a domain generated with a symplectic map.

To properly define SympNets, we first need to define shear maps, which can be seen as building blocks of symplectic maps.

**Definition 1** (Shear maps [4]). *One of the simplest families of symplectic transformations from  $\mathbb{R}^{2d}$  into  $\mathbb{R}^{2d}$  is called “shear maps”, and is defined by*

$$f_{\text{up}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \nabla V_{\text{up}}(x_2) \\ x_2 \end{pmatrix} \quad \text{and} \quad f_{\text{down}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + \nabla V_{\text{down}}(x_1) \end{pmatrix} \quad (13)$$

for  $x = (x_1, x_2)^\top$  with  $x_1 \in \mathbb{R}^d$  and  $x_2 \in \mathbb{R}^d$ , where  $V_{\text{up/down}} \in C^2(\mathbb{R}^d, \mathbb{R}^d)$ , and  $\nabla V_{\text{up/down}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the gradient of  $V_{\text{up/down}}$ .

Using the previous definition, the architecture of SympNets is based on the following lemma.

**Lemma 1** ([22]). *Any symplectic map can be approximated by composing of several shear maps, and the composition of several symplectic maps still remains symplectic.*

In practice, we could directly approximate the smooth functions  $V_{\text{up/down}}$  using a fully connected NN. However, this approach would require the computation of the gradient of  $V_{\text{up/down}}$  at each iteration. SympNets are designed to bypass this computational constraint, with the advantage of requiring very few trainable parameters compared to fully connected NNs. To this aim, [22] introduces gradient modules, the principles of which are summarized below.

---

<sup>2</sup>For those unfamiliar with symplectic maps, they are used to model the evolution of dynamical systems while preserving certain geometric properties. For instance, the flow of a Hamiltonian system is a symplectic map, preserving volume in the phase space of the dynamical system. For more details on symplectic maps, the reader is referred to [4, 27, 15]. In a nutshell, in  $\mathbb{R}^n$  with  $n = 2d$ , a  $C^1$ -diffeomorphism  $\mathcal{S}$  is called symplectic if its Jacobian matrix  $DS$  satisfies  $(DS)^\top J(DS) = J$ , where  $J$  is a block matrix called the standard symplectic form on  $\mathbb{R}^n$ , with  $J_{11} = J_{22} = 0_d$  and  $J_{12} = -J_{21} = I_d$ . This definition is enough to see that  $|\det DS| = 1$ , and therefore that  $\mathcal{S}$  is volume-preserving.

Let  $q > 0$  be the depth of the NN. In practice, we set  $q > 2d$ . We define  $\hat{\sigma}_{K,a,b}$  the approximation of  $\nabla V_{\text{up/down}}$  in terms of an activation function  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ , two vectors  $a, b \in \mathbb{R}^q$ , a matrix  $K \in \mathcal{M}_{q,d}(\mathbb{R})$ , and  $\text{diag}(a) = (a_i \delta_{ij})_{1 \leq i, j \leq q}$ , as follows, for all  $x \in \mathbb{R}^d$ ,

$$\hat{\sigma}_{K,a,b}(x) = K^\top \text{diag}(a) \sigma(Kx + b), \quad (14)$$

where it is understood that  $\sigma$  is applied to a vector componentwise. Then, gradient modules  $\mathcal{G}_{\text{up}}$  and  $\mathcal{G}_{\text{down}}$  are defined to approximate  $f_{\text{up}}$  and  $f_{\text{down}}$ , by

$$\mathcal{G}_{\text{up}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \hat{\sigma}_{K,a,b}(x_2) \\ x_2 \end{pmatrix} \quad \text{and} \quad \mathcal{G}_{\text{down}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + \hat{\sigma}_{K,a,b}(x_1) \end{pmatrix}. \quad (15)$$

These functions are called gradient modules because  $\hat{\sigma}_{K,a,b}$  is able to approximate any Jacobian of smooth vector fields, see [22, Appendix A].

Moreover, it is also possible to define a parametric gradient module to approximate a family of symplectic maps indexed by  $n_\mu$  parameters  $\mu \in \mathbb{M} \subset \mathbb{R}^{n_\mu}$ . To that end, we define a second matrix  $K_\mu \in \mathcal{M}_{q,n_\mu}(\mathbb{R})$ , and replace  $\hat{\sigma}_{K,a,b}(x)$  with

$$\tilde{\sigma}_{K,K_\mu,a,b}(x; \mu) = K^\top \sigma(Kx + b + K_\mu \mu). \quad (16)$$

One can show that  $\tilde{\sigma}_{K,K_\mu,a,b}$  is a gradient module [21], and that the whole network remains symplectic with respect to  $x \in \mathbb{R}^{2d}$ , for each parameter  $\mu \in \mathbb{M}$ .

Finally, the architecture of SympNets is made by composing several gradient modules, as shown in Figure 2.

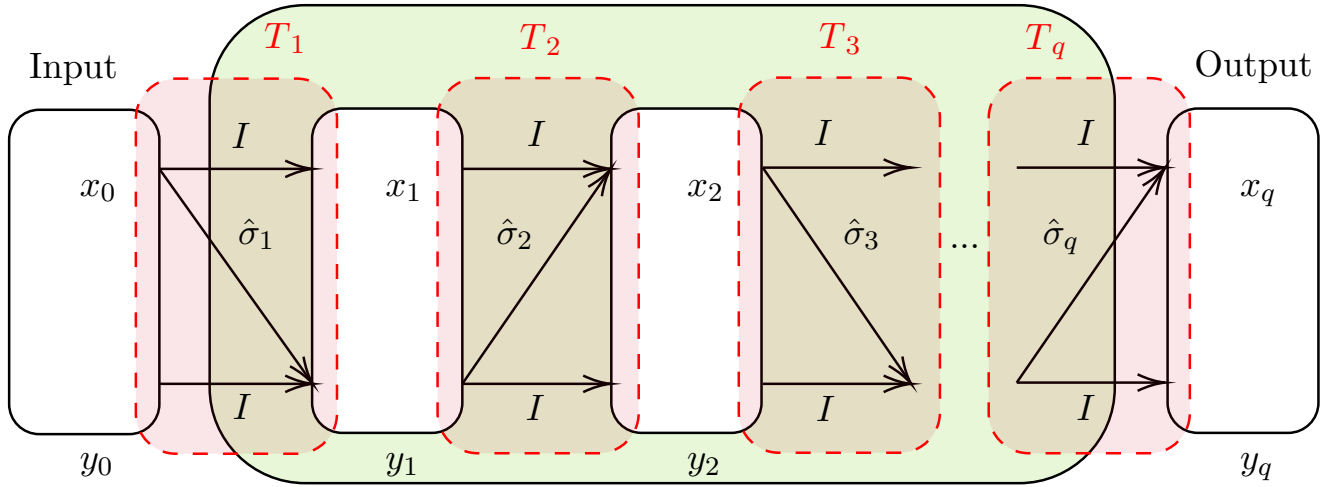


Figure 2: SympNet architecture. Here the  $(T_i)_{1 \leq i \leq q}$  are shear symplectic maps, while the  $(\hat{\sigma}_i)_{1 \leq i \leq q}$  denote their associated gradient modules.

Once the SympNet has been constructed, a loss function has to be designed to optimize all the trainable parameters contained in gradient modules. To that end, the starting space is assumed to be included in a compact set  $K$ , here the sphere  $\mathbb{S}^1$ . First, we draw points  $\{x^i\}_{i=1}^{N_b}$  on the boundary of  $K$ , and evaluate their image by  $\mathcal{T}$ , a given symplectic map that the SympNet  $T_\omega$  with trainable weights  $\omega$  will be trained to learn.



The symplectic NN  $T_\omega$  is therefore trained by minimizing the cost function

$$\mathcal{J}_S(\omega; \{x^i\}_{i=1}^N) = \sum_{i=1}^N |T_\omega(x^i) - \mathcal{T}(x^i)|^2, \quad (17)$$

with  $|\cdot|$  the euclidean norm on  $\mathbb{R}^2$ .

Equipped with PINNs (or rather DeepRitz networks) introduced in [Section 2.1](#) as well as SympNets, we now combine them in the next section to solve the shape optimization problem [\(5\)](#).

We will now focus on solving the shape optimization problem in two dimensions. From this point forward, we will assume that  $n = 2$ .

### 2.3 Solving a PDE with PINNs in a domain generated with a symplectic map

Now that we know how to learn a symplectic transformation with a SympNet, we now wish to combine this approach with PINNs to solve the Poisson problem [\(1\)](#) in this shape. For simplicity, we consider the 2D case, i.e., we take  $d = 1$  and transform the unit sphere  $\mathbb{B}^2$  into a shape  $\mathcal{T}\mathbb{B}^2$  homeomorphic to the ball with volume  $V_0$ . However, directly adapting the approach described in [Section 2.1](#) leads to difficulties in properly defining the boundary conditions. Indeed, the solution of the PDE should be expressed as in [\(8\)](#), where the trained network  $u_\theta$  is multiplied by a function  $\alpha$  and summed to a function  $\beta$ . In this case, we would have to devise new expressions of  $\alpha$  and  $\beta$  in  $\mathcal{T}\mathbb{B}^2$ . Faced with this difficulty, we choose to bypass that problem, by solving the PDE satisfied by  $w : \mathbb{B}^2 \rightarrow \mathbb{R}$ , defined for a.e.  $x \in \mathbb{B}^2$  and  $y \in \mathcal{T}\mathbb{B}^2$  s.t.  $y = \mathcal{T}x$  by

$$w(x) = (u_{\mathcal{T}} \circ \mathcal{T})(x) = u_{\mathcal{T}}(y), \quad (18)$$

with  $u_{\mathcal{T}}$  the solution of the Poisson problem [\(1\)](#) in  $\mathcal{T}\mathbb{B}^2$ . Since the inputs of the function  $v$  lie in the unit sphere, it is again very easy to find appropriate functions  $\alpha$  and  $\beta$ . [Figure 3](#) recaps this notation.

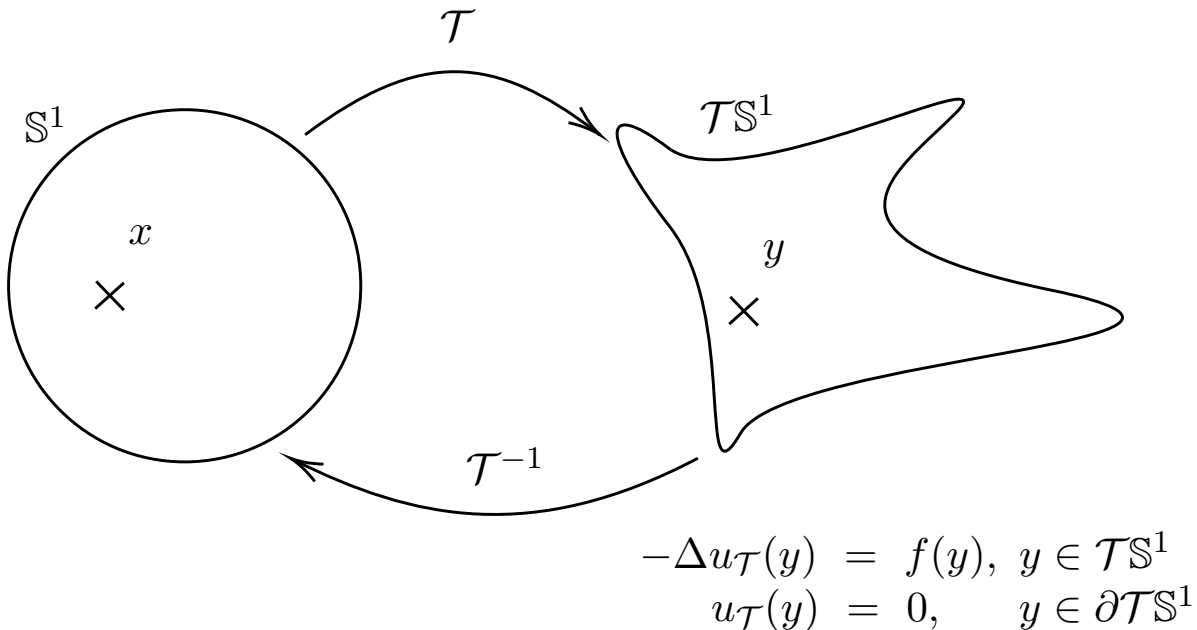


Figure 3: Transformation of the ball  $\mathbb{B}^2$  by a symplectic map  $\mathcal{T}$ .



Equipped with  $w$ , we now have to determine the PDE it is a solution to. The answer lies in the following lemma (that remains true in higher dimensions).

**Lemma 2.** *Let  $\mathcal{T}$  be a symplectic map on  $\mathbb{R}^2$  and  $u_{\mathcal{T}} \in H_0^1(\mathcal{T}\mathbb{B}^2)$ , such that  $w = u_{\mathcal{T}} \circ \mathcal{T} \in H_0^1(\mathbb{B}^2)$ . If  $u_{\mathcal{T}} \in H_0^1(\mathcal{T}\mathbb{B}^2)$  is the solution of the Poisson problem (1), then  $w \in H_0^1(\mathbb{B}^2)$  is solution of*

$$\begin{cases} -\operatorname{div}(A\nabla w) = \tilde{f}, & \text{in } \mathbb{B}^2; \\ w = 0, & \text{on } \mathbb{S}^1, \end{cases} \quad (19)$$

with  $A : \mathbb{B}^2 \rightarrow \mathcal{M}_2(\mathbb{R})$  is defined by  $A = J_{\mathcal{T}}^{-1} \cdot J_{\mathcal{T}}^{-\top}$ , with  $\tilde{f} = f \circ \mathcal{T}$  and with  $J_{\mathcal{T}}$  the jacobian matrix of  $\mathcal{T}$ . Moreover, the following problem can be formulated in a weaker sense, as an optimization problem

$$\inf \left\{ \frac{1}{2} \int_{\mathbb{B}^2} A\nabla w \cdot \nabla w - \int_{\mathbb{B}^2} \tilde{f}v, \quad \exists u_{\mathcal{T}} \in H_0^1(\mathcal{T}\mathbb{B}^2), \quad w = u_{\mathcal{T}} \circ \mathcal{T} \in H_0^1(\mathbb{B}^2) \right\}. \quad (20)$$

*Proof.* Assume  $u_{\mathcal{T}}$  solves the Poisson problem (1). Then, for all  $\varphi \in H_0^1(\mathcal{T}\mathbb{B}^2)$ , the following variational formulation holds:

$$\int_{\mathcal{T}\mathbb{B}^2} \nabla u_{\mathcal{T}}(y) \cdot \nabla \varphi(y) dy = \int_{\mathcal{T}\mathbb{B}^2} f(y) \varphi(y) dy.$$

Proceeding with the change of variable  $y = \mathcal{T}x$ , we get

$$\int_{\mathbb{B}^2} |J_{\mathcal{T}}(x)| \nabla u_{\mathcal{T}}(\mathcal{T}x) \cdot \nabla \varphi(\mathcal{T}x) dx = \int_{\mathbb{B}^2} |J_{\mathcal{T}}(x)| f(\mathcal{T}x) \varphi(\mathcal{T}x) dx,$$

with  $|J_{\mathcal{T}}|$  the determinant of the Jacobian matrix of  $\mathcal{T}$ . We now introduce  $w = u_{\mathcal{T}} \circ \mathcal{T}$ ,  $\tilde{f} = f \circ \mathcal{T}$  and  $\tilde{\varphi} = \varphi \circ \mathcal{T}$ . By using the chain rule, we get for a.e.  $x \in \mathbb{B}^2$ ,  $\nabla u_{\mathcal{T}}(\mathcal{T}x) = J_{\mathcal{T}}^{-\top}(x) \nabla w(x)$  and  $\nabla \varphi(\mathcal{T}x) = J_{\mathcal{T}}^{-1}(x) \nabla \tilde{\varphi}(x)$ . Since  $\mathcal{T}$  is a symplectic map and therefore is volume-preserving,  $|J_{\mathcal{T}}(x)| = 1$ , and we obtain

$$\int_{\mathbb{B}^2} \left[ J_{\mathcal{T}}^{-\top}(x) \nabla w(x) \right] \cdot \left[ J_{\mathcal{T}}^{-1}(x) \nabla \tilde{\varphi}(x) \right] dx = \int_{\mathbb{B}^2} \tilde{f}(x) \tilde{\varphi}(x) dx. \quad (21)$$

Then, (21) directly leads us to the variational problem of finding  $v \in H_0^1(\mathbb{B}^2)$ , such that, for all  $\tilde{\varphi} \in H_0^1(\mathbb{B}^2)$ ,

$$\int_{\mathbb{B}^2} \left[ J_{\mathcal{T}}^{-1}(x) J_{\mathcal{T}}^{-\top}(x) \right] \nabla w(x) \cdot \nabla \tilde{\varphi}(x) dx = \int_{\mathbb{B}^2} \tilde{f}(x) \tilde{\varphi}(x) dx. \quad (22)$$

The proof is thus concluded.  $\square$

**Remark 3.** Note that the proof of Lemma 2 follows a similar reasoning as in the proof of [18, Theorem 5.3.2]. Indeed, Lemma 2 can formally be seen as computing a shape derivative, but here the shape derivative is simply represented through the derivative of the symplectic map.

## 2.4 Shape optimization with PINNs and SympNets

To propose a shape optimization algorithm, we now just have to minimize a loss function depending on both PINNs and SympNets. It is given by

$$\mathcal{J}_{P/S}(\theta, \omega; \{x^i, \mu^i\}_{i=1}^N) = \frac{V_0}{N} \sum_{i=1}^N \left\{ \frac{1}{2} A_{\omega} \nabla v_{\theta, \omega} \cdot \nabla v_{\theta, \omega} - \tilde{f}_{\omega} v_{\theta, \omega} \right\} (x^i; \mu^i), \quad (23)$$

with

- $\theta$  the trainable weights of the PINN,  $\omega$  the trainable weights of the SympNet;
- $v_{\theta,\omega} : x \in \mathbb{B}^2 \mapsto \alpha(x)u_\theta(T_\omega x) + \beta(x) \in \mathbb{R}$  the solution of the Poisson problem set in  $T_\omega \mathbb{B}^2$ ;
- $T_\omega : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$  the SympNet;
- $A_\omega$  the diffusion matrix defined by  $A_\omega = J_{T_\omega}^{-1} J_{T_\omega}^{-\top}$ ;
- $u_\theta : T_\omega \mathbb{B}^2 \rightarrow \mathbb{R}$  the PINN;
- $\alpha : \mathbb{B}^2 \mapsto \mathbb{R}$  a  $C^\infty$  function that vanishes on  $\mathbb{S}^1$ ;
- $\beta : \mathbb{B}^2 \mapsto \mathbb{R}$  a  $C^\infty$  function that satisfies the boundary condition of the Poisson problem (here,  $\beta = 0$  on  $\mathbb{S}^1$ );
- $\tilde{f}_\omega : x \in \mathbb{B}^2 \mapsto (f \circ T_\omega)(x) \in \mathbb{R}$ ;
- $N$  the number of collocation points;
- $\{x^i, \mu^i\}_{i=1}^N$  a set of random collocation points and values of the parameters.

**Remark 4.** If we uniformly sample points  $\{x^i\}_i$  in the unit ball  $\mathbb{B}^2$  and apply a symplectic transformation  $\mathcal{T}$  to obtain a shape  $\mathcal{T}\mathbb{B}^2$ , then the points  $\{\mathcal{T}x^i\}_i$  are again uniformly distributed in  $\mathcal{T}\mathbb{B}^2$ . This is immediate, since  $\mathcal{T}$  is volume-preserving.

**Remark 5.** One of the exciting aspects of this formulation of the Dirichlet energy optimization problem is that we iterate the gradient descent on the SympNet and PINN at the same time. There is no need to wait for the PDE to be solved before iterating on the shape. The gradient descent is now parallelizable.

**Remark 6.** The major advantage of this formulation is that the shape derivative can be calculated using automatic differentiation techniques.

Let us summarize below the main steps of the shape optimization algorithm implemented.

**Remark 7.** At each iteration, contrary to classical shape derivative methods the SympNet representing the shape and the PINN representing the PDE are updated, even if the PINN is still not the solution of the PDE in the shape. The PINN is expected to solve the Poisson equation (1) in the shape only when the algorithm has converged. This allows us to parallelize the entire loop and go beyond conventional methods, which are intrinsically sequential.

### 3 Numerical results

Equipped with Section 3.2, we are now ready to provide a comprehensive numerical study. As a sanity check, we first present the results of the DeepRitz method in Section 3.1 and of the SympNets in Section 3.2. When no exact solution is available, the results of the code will be compared to results obtained via a fixed point algorithm implemented on FreeFem++ [9, 16]. To test our SympNets, we introduce the symplectic map  $\mathcal{T}_\lambda = \mathcal{S}_\lambda^1 \circ \mathcal{S}_\lambda^2 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , with

$$\begin{cases} \mathcal{S}_\lambda^1 : (x_1, x_2) \mapsto (x_1 - \lambda x_2^2 + 0.3 \sin(\frac{x_2}{\lambda}) - 0.2 \sin(8x_2), x_2), \\ \mathcal{S}_\lambda^2 : (x_1, x_2) \mapsto (x_1, x_2 + 0.2\lambda x_1 + 0.12 \cos(x_1)). \end{cases} \quad (24)$$

---

**Algorithm 1** Shape optimization algorithm

---

**Require:**

- *collocation domain*:  $\mathbb{B}^2 \times \mathbb{M}$ ;
- *PINN and SympNet hyperparameters*: see [Appendix B](#);
- *algorithm parameters*: gradient descent step, number of epochs, number of collocation points  $N$ , see [Appendix B](#);

**while** epoch < total number of epochs **do**

- ↪ randomly draw  $N$  collocation points  $\{x^i, \mu^i\}_{i=1}^N \in (\mathbb{B}^2 \times \mathbb{M})^N$ ;
- ↪ for each trainable weights  $\theta$  of the PINN, and each trainable weights  $\omega$  of the SympNet:
  - determination of the transformation  $T_\omega$  from the parameter  $\omega$ ;
  - knowing  $T_\omega$ , computation of the diffusion matrix  $A_\omega$  and of the source term  $\tilde{f}_\omega$  involved in [\(19\)](#);
  - computation of  $v_{\theta, \omega}$  the composition of the PINN by the SympNet at each point  $(x^i, \mu^i)$ ;
- ↪ calculation of the loss given by [Equation \(23\)](#);
- ↪ calculation by automatic differentiation of the loss derivative with respect to PINN and SympNet trainable weights  $(\theta, \omega)$ ;
- ↪ implementation of a gradient method step and update of the PINN and SympNet trainable weights;

**end while**

---

Note that  $\mathcal{T}_\lambda$  is a nonlinear function of  $\lambda$ . This map will be the support of several numerical experiments. Then, we combine both approaches in [Section 3.3](#) to solve the shape optimization problem [\(5\)](#). Finally, we solve a more intricate problem, namely the exterior Bernoulli free-boundary problem, in [Section 3.4](#). The hyperparameters of the networks are summarized in [Appendix B](#).

### 3.1 Solving the Poisson problem with the DeepRitz method

This section is dedicated to checking that the energetic formulation [\(20\)](#) introduced in [Lemma 2](#) is indeed able to solve the Poisson problem [\(1\)](#) in a given shape  $\Omega$ . For this purpose, we choose a complex shape  $\Omega$  given by the symplectic map [\(24\)](#) with  $\lambda = 0.5$  applied to the annulus with inner radius 0.2 and outer radius 1. This leads to a shape with a hole. More specifically, we highlight the capability of our neural network-based approach to tackle a parametric Poisson problem, with a parametric source term  $f$ , given by

$$f(x_1, x_2; \mu) = \exp\left(1 - \left(\frac{x_1}{\mu}\right)^2 - (\mu x_2)^2\right), \quad (25)$$

where the parameter  $\mu \in \mathbb{M} = (0.5, 1.5)$  controls the elongation of the ellipse described by the source term. Therefore, the neural network is a function of the two space variables  $x_1$  and  $x_2$ , as well as of the parameter  $\mu$ . The results are displayed on [Figure 4](#).

To quantify the approximation quality, we cannot directly use the Dirichlet energy [\(4\)](#), since we do not know its optimal value for this specific problem. Instead, we elect to use the variational formulation [\(22\)](#). Indeed, we know that the solution to the Poisson problem satisfies, for all  $\tilde{\varphi} \in H_0^1(\mathbb{B}^2)$ ,

$$\int_{0.5}^{1.5} \int_{\mathbb{B}^2} \left( \left[ J_{\mathcal{T}}^{-1}(x) J_{\mathcal{T}}^{-\top}(x) \right] \nabla w(x; \mu) \cdot \nabla \tilde{\varphi}(x) - \tilde{f}(x; \mu) \tilde{\varphi}(x) \right) dx d\mu = 0. \quad (26)$$

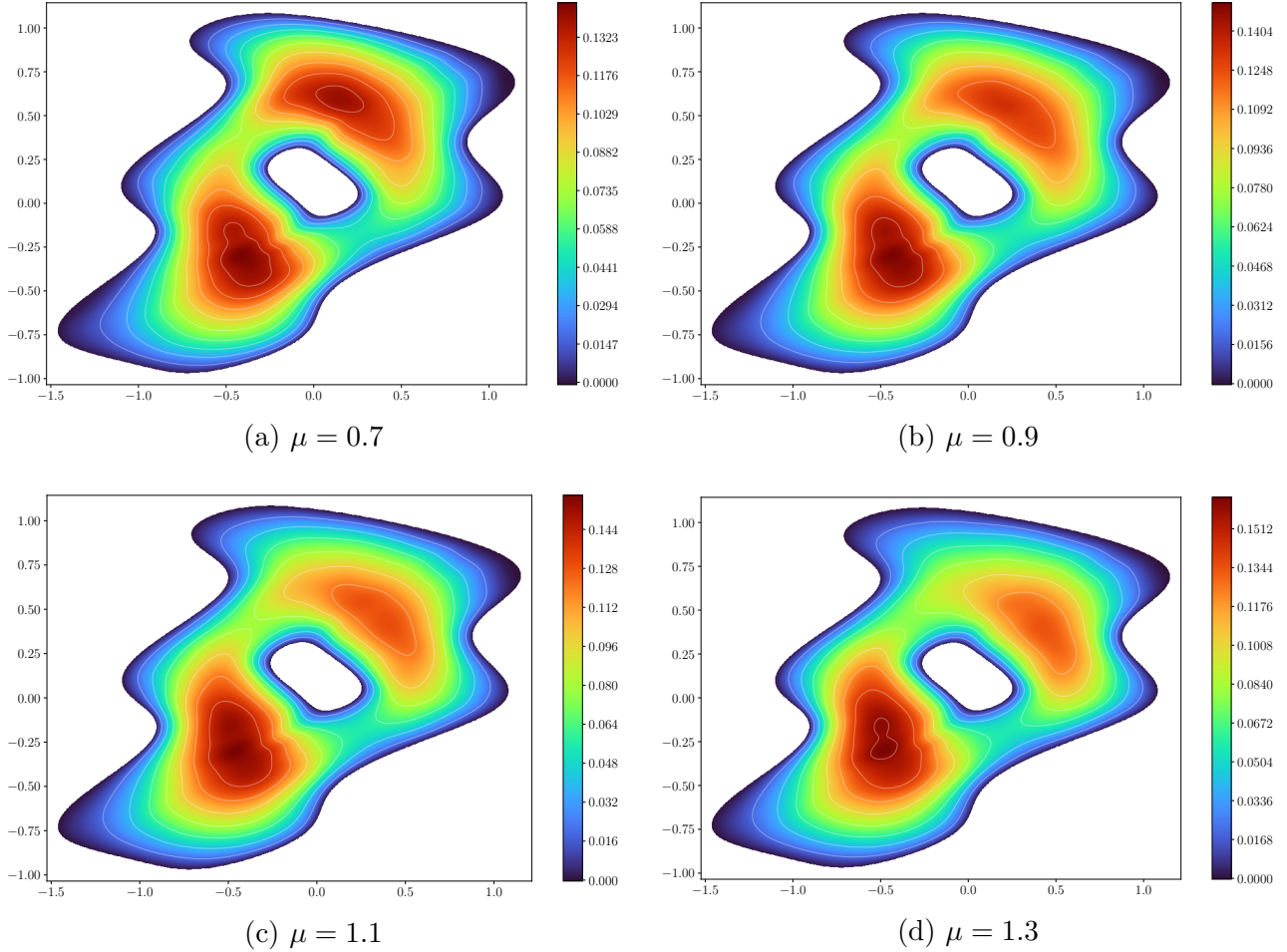


Figure 4: Results of the PINN applied to the Poisson problem (1) with the source term (25), on the domain given by applying the symplectic map (24) with  $\lambda = 0.5$  to the annulus with inner radius 0.2 and outer radius  $\mu$ . Each subfigure corresponds to a different value of the parameter  $\mu$ .

To that end, we generate a set of  $10^3$  functions  $\varphi \in H_0^1(\mathbb{B}^2)$ , defined by  $\varphi(x_1, x_2) = \alpha(x_1, x_2)P(x_1, x_2)$ , where  $\alpha$  is the previously defined function that vanishes on  $\mathbb{S}^1$ , and where  $P$  is a two-dimensional polynomial of degree 2 whose six coefficients are uniformly sampled in  $(0, 1)$ . Then, the integral in (26) is approximated using 20 000 collocation points for  $x_1, x_2$  and  $\mu$ . The results are reported in Table 1, where we observe that, on average, the error is of the order of  $10^{-2}$ . Therefore, our approach indeed provides a good approximation of the solution to the parametric Poisson problem on a complex shape.

Table 1: Statistics on the absolute value of the integral in (26) for the approximation of the parametric Poisson problem in Section 3.1.

Mean value	Maximal value	Minimal value	Standard deviation
$8.43 \times 10^{-3}$	$4.53 \times 10^{-2}$	$1.10 \times 10^{-5}$	$7.08 \times 10^{-3}$

### 3.2 Learning a simply connected parameterized shape with a SympNet

The goal of this section is to show that our parametric SympNet introduced in Section 2.2.2 is indeed able to learn a given symplectic map as a deformation of the unit sphere of  $\mathbb{R}^2$ . To that end, we train a SympNet to learn the shape given by the symplectic map (24), with  $\lambda \in \mathbb{M} = (0.5, 2)$ . Therefore, the SympNet takes as input the parameter  $\lambda$ , as well as the two space variables  $x_1$  and  $x_2$ . To measure the error made by the SympNet, we use the Hausdorff distance, computed with the algorithm from [31], between the true shape and the learned shape.

The results are displayed on Figure 5. We observe a good agreement between the reference shape and the learned shape, for all considered values of  $\lambda$ . Hence, our parametric SympNet is able to learn this rather complex parametric symplectic map.

To quantify the quality of the approximation, we compute the Hausdorff distance between the reference shape and the learned shape. To that end, we randomly sample 1000 points in the parameter space  $\mathbb{M} = (0.5, 2)$ , and compute some statistics on the Hausdorff distance. They are collected in Table 2, where we observe that the mean value of the Hausdorff distance is of the order of  $10^{-2}$ , which makes up for a satisfactory approximation.

Table 2: Statistics on the Hausdorff distance between the reference shape and the learned shape, on the parameter set  $\mathbb{M} = (0.5, 2)$ .

Mean value	Maximal value	Minimal value	Standard deviation
$9.45 \times 10^{-3}$	$2.10 \times 10^{-2}$	$6.05 \times 10^{-3}$	$2.12 \times 10^{-3}$

### 3.3 Solving the geometric Dirichlet energy problem with PINNs and SympNets

The goal of this section is to show that our joint optimization procedure, involving both PINNs and SympNets and based on the loss function (23), is able to learn the optimal shape for the parametric Poisson problem (11). To that end, we first consider non-parametric problems to validate our approach, before moving on to parametric ones. According to Theorem 1, we will look at the first order optimality conditions and for a given shape  $\Omega$  and  $u$  the associated solution of the PDE, we will consider the optimality error, an optimality criterion involving the standard deviation of Equation (6) on  $\partial\Omega$ :

$$\text{optimality error} := \sqrt{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} (\partial_n u(x) - \overline{\partial_n u})^2 \, d\sigma} = 0,$$

where  $\overline{\partial_n u}$  denotes the average value of  $\partial_n u$  on  $\partial\Omega$ .

#### 3.3.1 Non-parametric problem: results with $f = 1$

The results with  $f = 1$  are displayed on Figure 6. In this case, the optimal shape is known to be the unit sphere [18]). However, we do not impose any constraint on the center of the unit sphere, and therefore the optimal shape is not necessarily centered at the origin. We observe a good agreement between the learned optimal shape and the true one (top left panel), as well as a good value of the deviation from the average of the optimality condition (top right panel). Moreover, the error between the approximate PDE solution

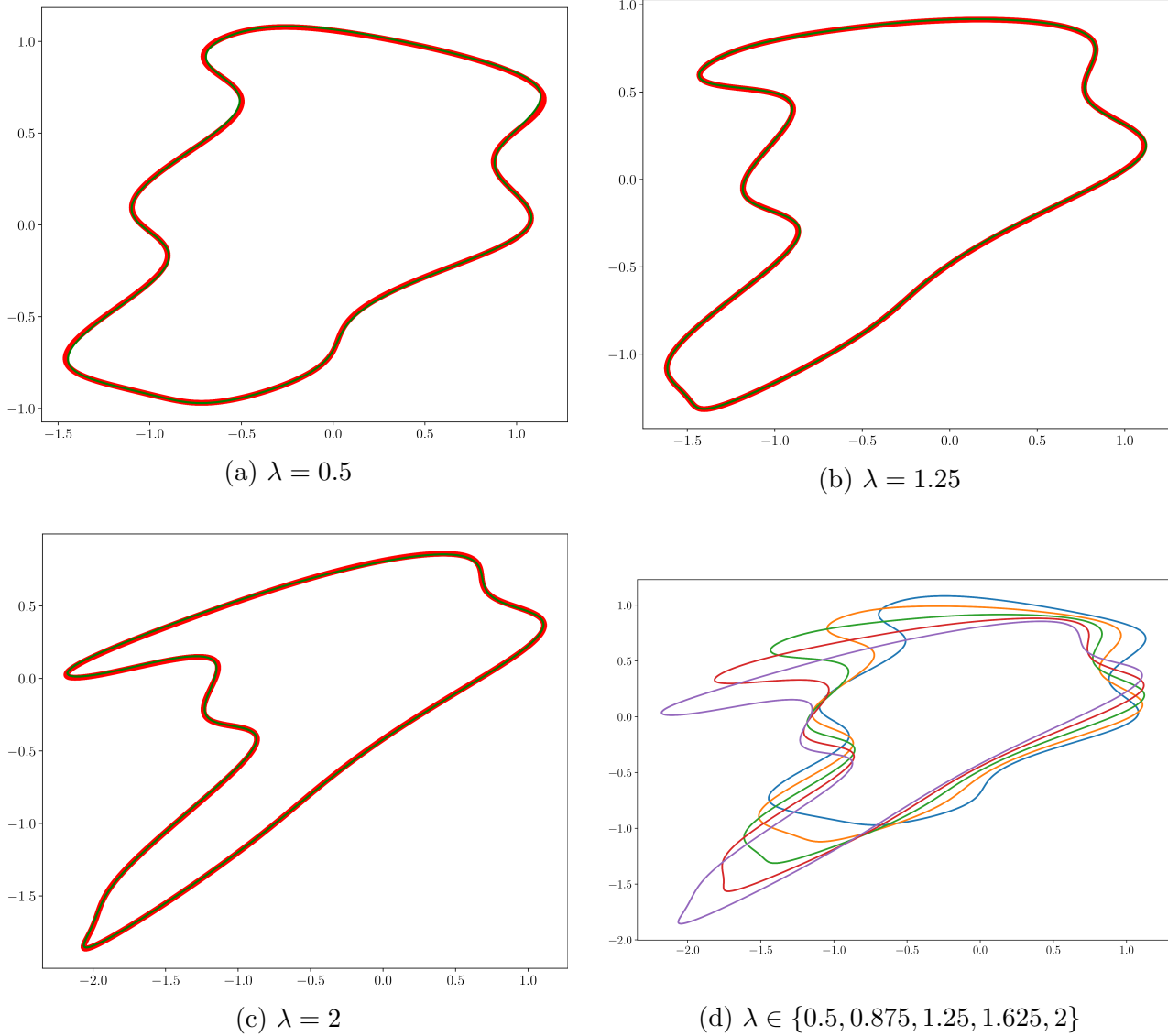


Figure 5: SympNets for  $\mathcal{T}_\lambda$ , for  $\lambda = 0.5$  (upper left panel),  $\lambda = 1.25$  (upper right panel) and  $\lambda = 2$  (lower left panel). The red lines correspond to the reference shape, while the green ones correspond to the learned shape. The lower right picture is a superposition of learned shapes for  $\lambda \in \{0.5, 0.875, 1.25, 1.625, 2\}$ .

and the exact one is low (bottom right panel). In addition, we report some metrics in [Table 3](#) (namely, the Hausdorff distance between the learned and true shapes, the optimality error, and the  $L^2$  error between the exact and approximate solutions). Remark that the Hausdorff distance reported in [Table 3](#) is of the same order as the one obtained in [Table 2](#) when using a SympNet to approximate a shape. This is a good indicator that the SympNet in our approach is fully working as expected.

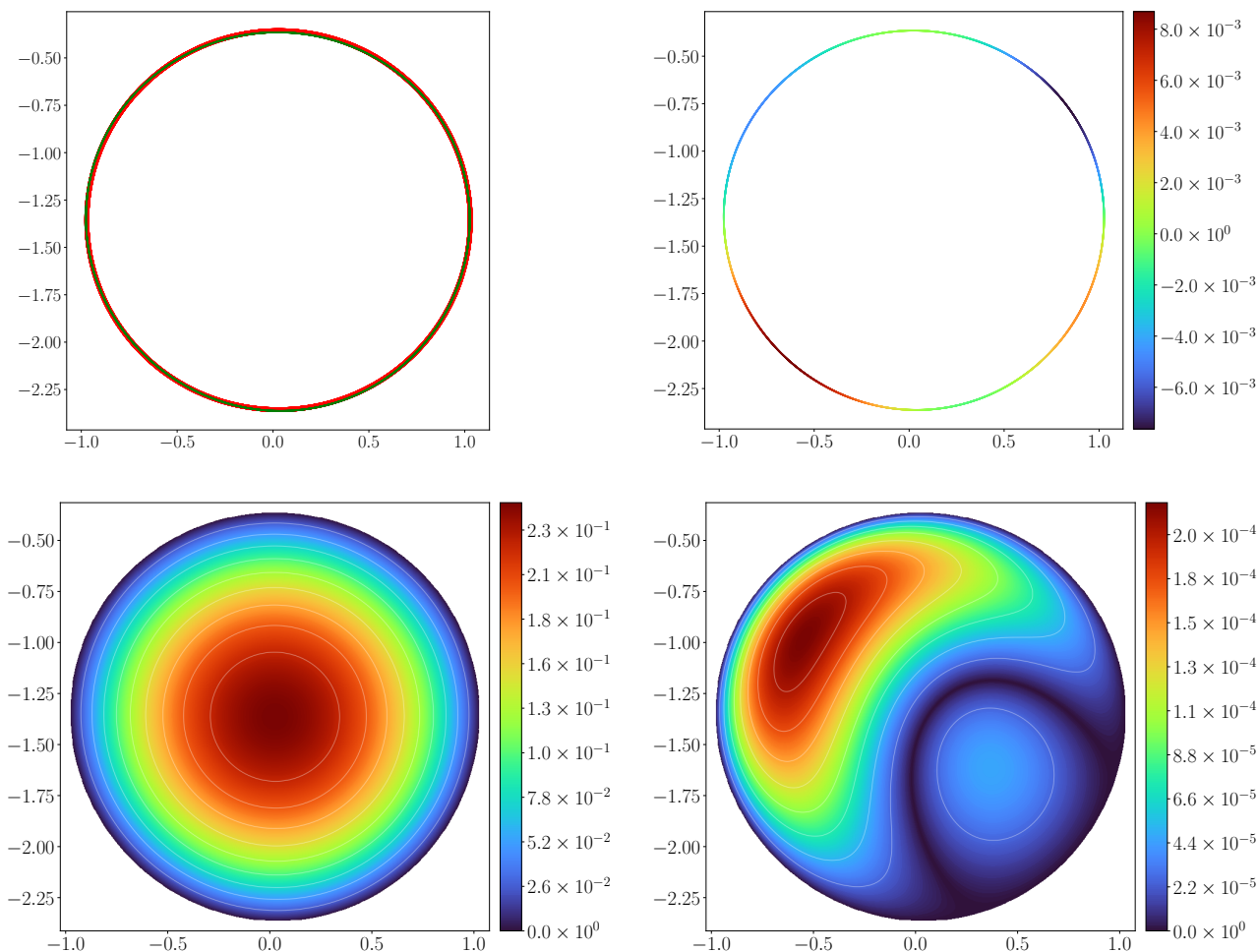


Figure 6: Shape optimization for the Dirichlet energy, for the Poisson equation with source term  $f = 1$ . Top left panel: learned shape (green line) and reference shape (red line). Top right panel: deviation from the average of the optimality condition in the learned shape. Bottom left panel: approximate PDE solution. Bottom right panel: pointwise error between the true and approximate PDE solutions.

Table 3: Relevant metrics related to the shape optimization of the Poisson equation with  $f = 1$ .

Hausdorff distance	Optimality error	$L^2$ error
$3.09 \times 10^{-2}$	$4.34 \times 10^{-3}$	$9.15 \times 10^{-5}$

### 3.3.2 Non-parametric problem: results with $f$ given by (25) with $\mu = 0.5$

Next, we turn to a more complex, non-constant source term. Namely, we take  $f$  given by (25) with  $\mu = 0.5$ , and display the results on Figure 7. This time, the reference shape has been obtained by using FreeFem++, with a fixed point algorithm based on [9]. As a consequence, we only consider it as a point of reference rather than a ground truth. We observe a good agreement between the learned shape and the reference one, much like in Section 3.3.1. Moreover, to quantify the approximation quality, we report in Table 4 the same



metrics as in [Section 3.3.1](#).

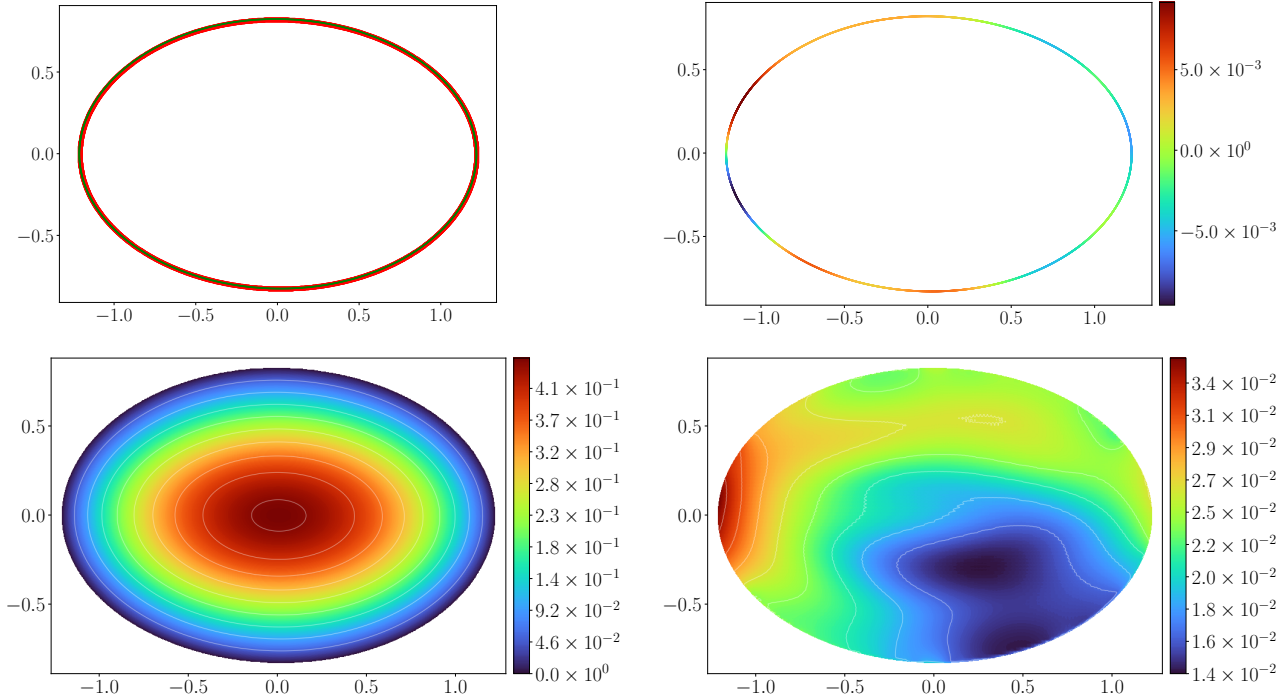


Figure 7: Shape optimization for the Dirichlet energy, for the Poisson equation with source term  $f$  given by (25) with  $\mu = 0.5$ . The results of a fixed-point algorithm [9] is our reference for this numerical experiment. Top left panel: learned shape (green line) and reference shape (red line). Top right panel: deviation from the average of the optimality condition in the learned shape. Bottom left panel: approximate PDE solution. Bottom right panel: pointwise error between the reference and approximate PDE solutions.

Table 4: Relevant metrics related to the shape optimization of the Poisson equation with source term  $f$  given by (25) with  $\mu = 0.5$ .

Hausdorff distance	Optimality error	$L^2$ error
$8.29 \times 10^{-3}$	$4.43 \times 10^{-3}$	$2.20 \times 10^{-2}$

### 3.3.3 Parametric problem: results with constant $f$

We now turn to a parametric shape optimization problem, where the source term is a constant  $\mu \in \mathbb{M} = (0.5, 1.5)$ . The results and errors are displayed on [Figure 8](#), where there is a good agreement between approximate and exact solution. In addition, we display the optimal shapes and the optimality error for 10 random values of  $\mu$  on [Figure 9](#). We note that the first order optimality condition is indeed satisfied, even though the shapes themselves are not necessarily centered at the origin. Indeed, like in [Section 3.3.1](#), the



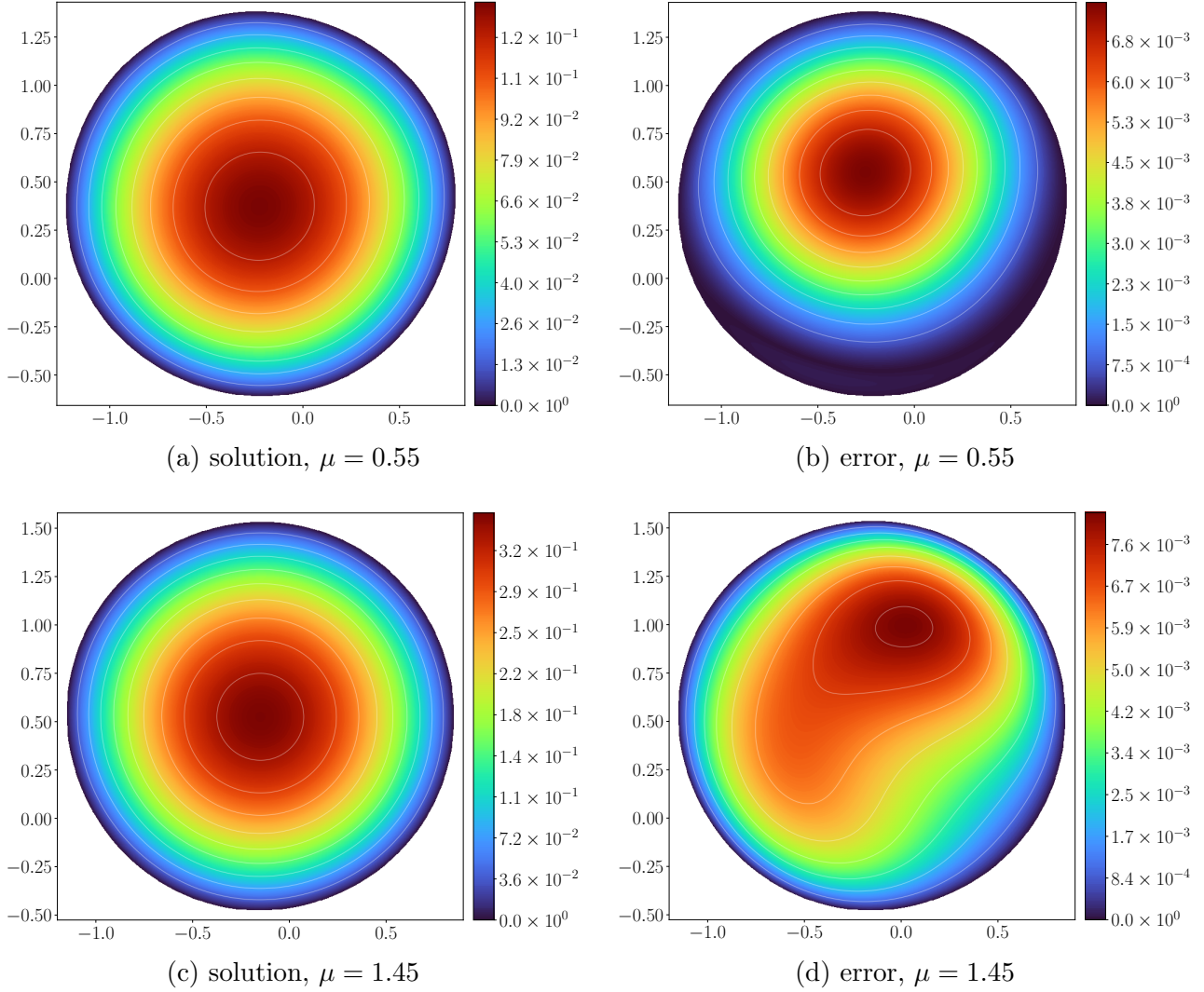


Figure 8: Shape optimization for the Dirichlet energy, for the Poisson equation with source term  $f = \mu$ . For two values of  $\mu$ , we display the approximate solution (left column) and the pointwise error between the approximate and exact solutions (right column).

optimal shape is the unit sphere, but the center of the sphere is not fixed. Finally, we report some statistics on our three main metrics in [Table 5](#), which confirm the relevance of our approach for a parametric problem.

### 3.3.4 Parametric problem: results with $f$ given by (25)

We now go back to the Poisson problem with source term given by (25). We learn the optimal shape for  $\mu \in \mathbb{M} = (0.5, 2)$ , and display the results on [Figure 10](#). Namely, we observe that the optimality condition is almost satisfied for all values of  $\mu$ . Then, we also report statistics on the optimality condition and, since no exact solution is available, the value of the integral in (26). We also observe a good quantitative behavior

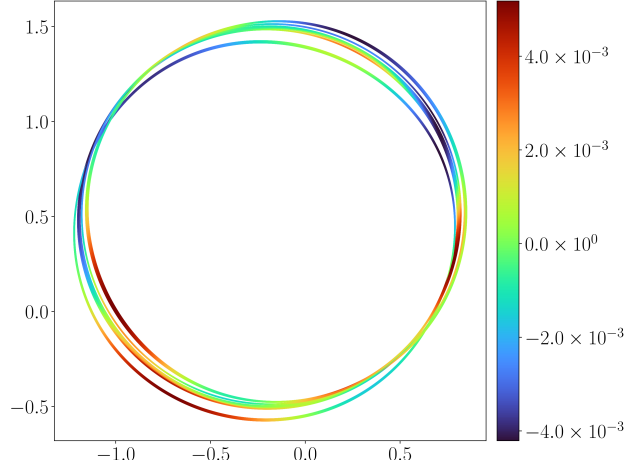


Figure 9: Shape optimization for the Dirichlet energy, for the Poisson equation with source term  $f = \mu$ : optimal shapes and deviation from the average of the optimality condition for 10 random values of  $\mu$ .

Table 5: Statistics of relevant metrics in the case of a parametric constant source term  $f = \mu$ , obtained by computing each metric for  $10^3$  values of  $\mu$ .

Metric	Mean	Max	Min	Standard deviation
Hausdorff distance	$3.16 \times 10^{-2}$	$3.56 \times 10^{-2}$	$2.89 \times 10^{-2}$	$1.86 \times 10^{-3}$
optimality error	$4.23 \times 10^{-3}$	$7.55 \times 10^{-3}$	$2.11 \times 10^{-3}$	$3.67 \times 10^{-3}$
$L^2$ error	$4.59 \times 10^{-3}$	$7.38 \times 10^{-3}$	$1.74 \times 10^{-3}$	$1.81 \times 10^{-3}$

of our approach.

Table 6: Statistics of relevant metrics in the case of a parametric source term given by (25), obtained by computing each metric for  $10^3$  values of  $\mu$ .

Metric	Mean	Max	Min	Standard deviation
optimality error	$1.37 \times 10^{-2}$	$2.23 \times 10^{-2}$	$8.31 \times 10^{-3}$	$9.19 \times 10^{-3}$
variational formulation	$5.89 \times 10^{-2}$	$1.58 \times 10^{-1}$	$3.21 \times 10^{-3}$	$2.89 \times 10^{-2}$

### 3.3.5 Parametric problem: results with a non-radial source term

For this last test case concerning the Poisson problem with Dirichlet boundary conditions, we take a source term given by  $f(x, y; \lambda) = \exp(1 - \|T_\lambda(x, y)\|^2)$ , with  $T_\lambda$  the symplectic map defined in (24) and  $\lambda \in (0.5, 2)$ . The results are depicted on Figure 11, with a good observed behavior of our approach. The same statistics as in Section 3.3.4 are reported in Table 7, which further confirm the relevance of our method.

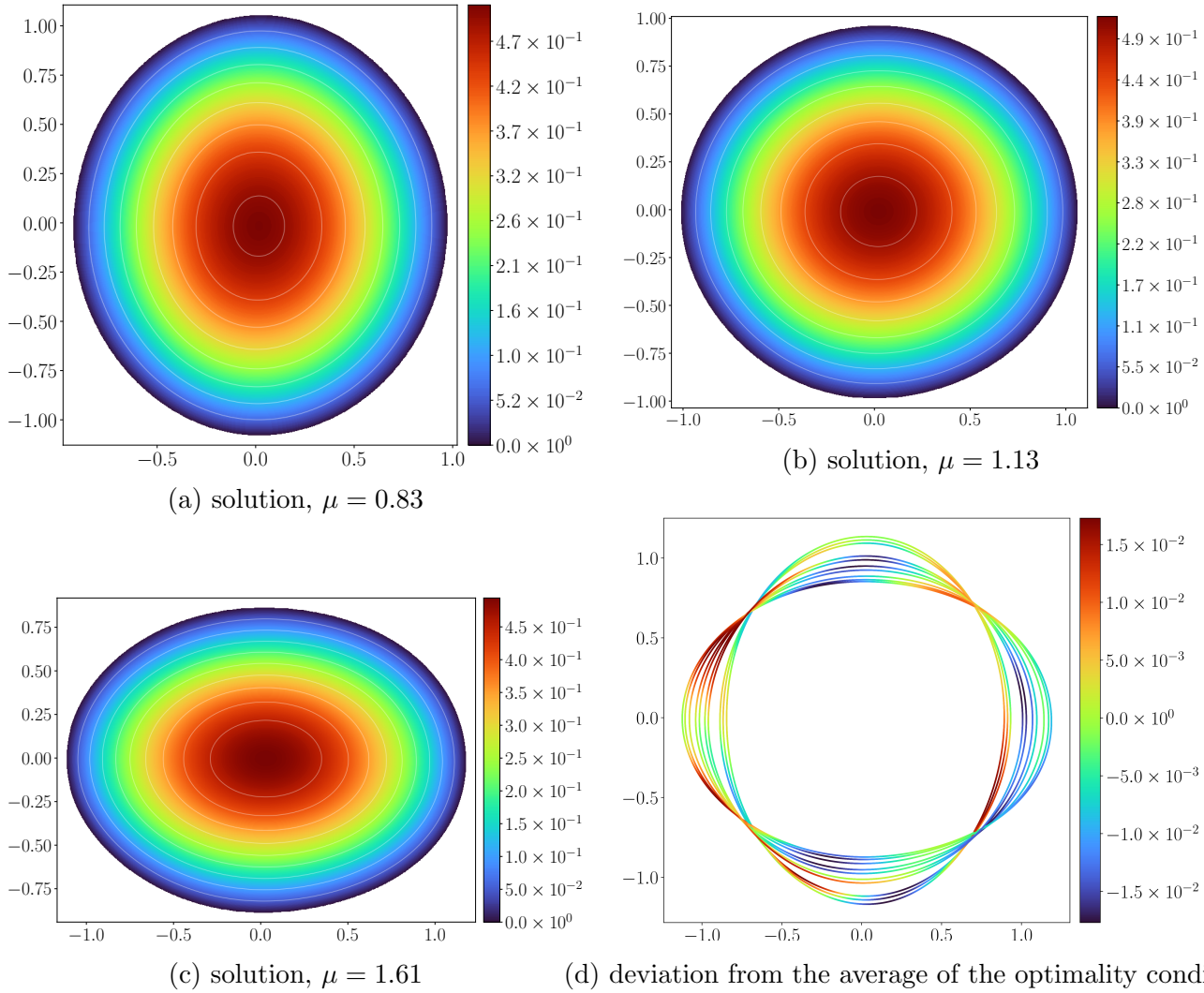


Figure 10: Shape optimization for the Dirichlet energy, for the Poisson equation with source term given by (25), in the context of Section 3.3.4. For three values of  $\mu$ , we display the approximate solution (top left, top right, and bottom left panels) and the deviation from the average of the optimality condition for 10 random values of  $\mu$  (bottom right panel).

Table 7: Statistics of relevant metrics for the parametric source term given in Section 3.3.5, obtained by computing each metric for  $10^3$  values of  $\lambda$ .

Metric	Mean	Max	Min	Standard deviation
optimality error	$2.04 \times 10^{-2}$	$3.36 \times 10^{-2}$	$1.24 \times 10^{-2}$	$1.57 \times 10^{-2}$
variational formulation	$3.36 \times 10^{-2}$	$1.12 \times 10^{-1}$	$7.27 \times 10^{-5}$	$2.15 \times 10^{-2}$

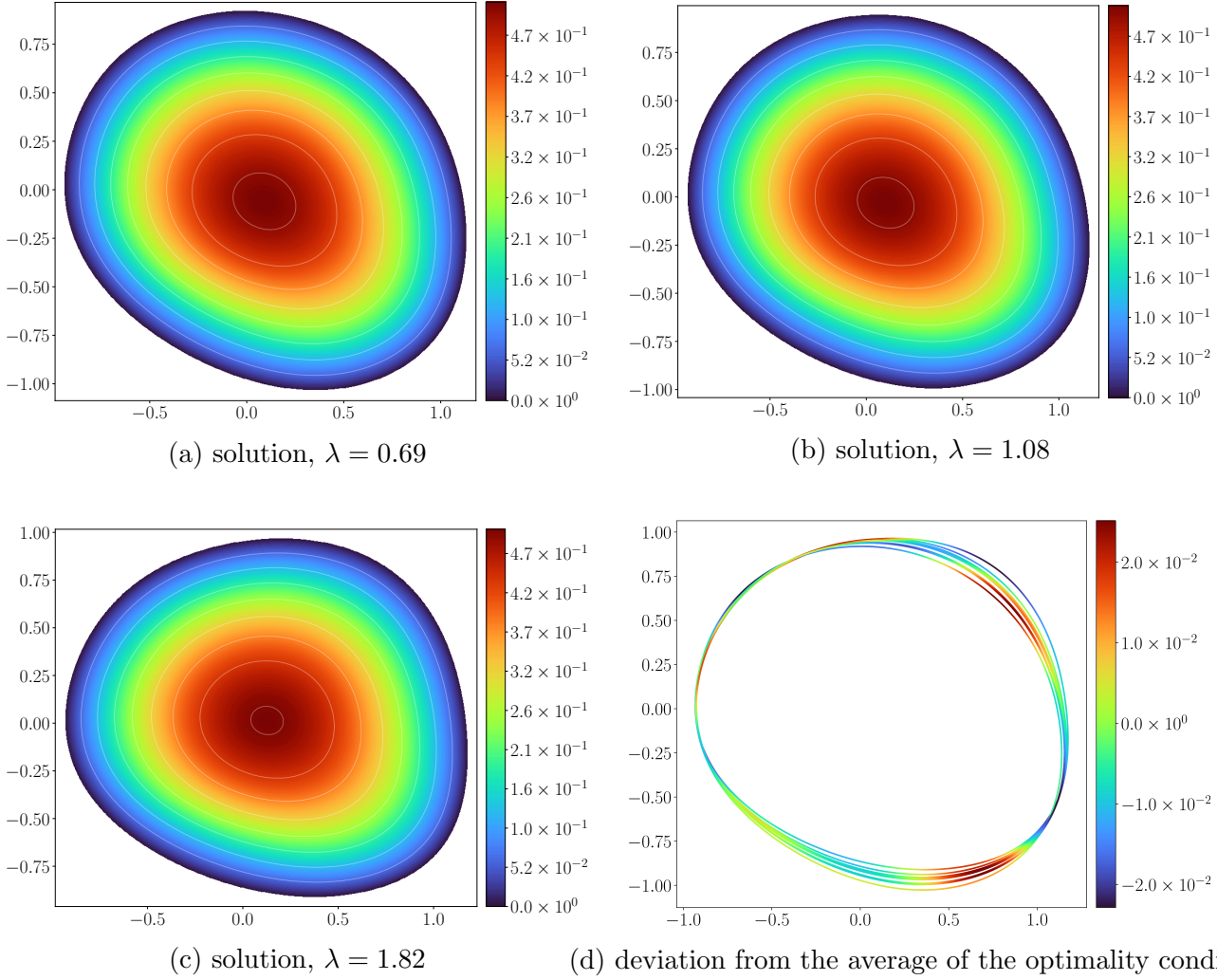
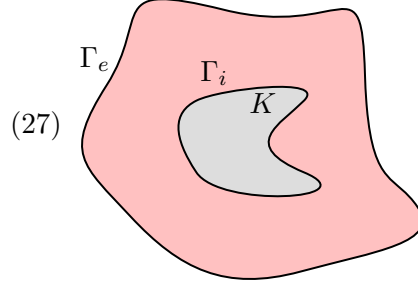


Figure 11: Parametric shape optimization for the Dirichlet energy in the context of [Section 3.3.5](#). For three values of  $\lambda$ , we display the approximate solution (top left, top right, and bottom left panels) and the deviation from the average of the optimality condition for 10 random values of  $\lambda$  (bottom right panel).

### 3.4 The exterior Bernoulli free-boundary problem

To check the relevance of our methodology in a more complex case, we consider the exterior Bernoulli free-boundary problem, described for instance in [\[19\]](#). Let  $\Omega$  be an open bounded connected set in  $\mathbb{R}^n$ , such that  $\partial\Omega = \Gamma_e \cup \Gamma_i$ , and  $K \subset \Omega$  a compact set such that  $\partial K = \Gamma_i$ . These sets are represented to the right of equation [\(27\)](#). The Bernoulli problem is an overdetermined PDE, whose unique solution  $u_\Omega^K \in H_0^1(\Omega)$  satisfies:

$$(27) \quad \begin{cases} -\Delta u_{\Omega}^K = 0 & \text{in } \Omega, \\ u_{\Omega}^K = 1 & \text{on } \Gamma_i, \\ u_{\Omega}^K = 0 & \text{on } \Gamma_e, \\ |\nabla u_{\Omega}^K| = c & \text{on } \Gamma_e, \end{cases}$$



for a given constant  $c > 0$ . This problem has been extensively studied. For instance, it can be demonstrated that if  $K$  is star-shaped (resp. convex), there exists a unique regular solution  $\Omega$  that is also star-shaped (resp. convex) [18, Chapter 6].

The solution  $u_{\Omega}^K$  of (27) remains a minimum of the Dirichlet energy  $\mathcal{E}_K(\Omega) = \frac{1}{2} \int_{\Omega} |\nabla u_{\Omega}^K|^2$ . Therefore, our method based on DeepRitz and SympNets remains relevant to solve this problem. We require adapting this strategy to the problem at hand; this is described in Section 3.4.1. Numerical results are then presented in Section 3.4.2.

### 3.4.1 Numerical strategy for solving Bernoulli problem with NNs

To implement Bernoulli's overdetermined problem, we need to know  $\varphi_i$  and  $\varphi_e$ , the respective level-set functions of  $\partial K = \Gamma_i$  and  $\Gamma_e$ . Here, we assume  $\varphi_i$  to be known, since  $K$  is a given set. As an example and for simplicity,  $K$  is taken as an ellipse centered at the origin, with width  $2a$  and height  $2b$ .

Like in Section 2.4, we want to transform an initial shape  $\Omega$  by a symplectic map  $\mathcal{T}$  into an optimal shape  $\Omega^*$ . Following (27), we denote by  $\Gamma_i$  the inner boundary of  $\Omega$  and  $\Gamma_e$  the outer boundary of  $\Omega$ . The main numerical challenge lies in mapping  $\Gamma_i$  onto  $\partial K$  by the symplectic map  $\mathcal{T}$ . A solution could be using penalization in the loss function, but it leads to a numerically ill-posed problem. Therefore, we look for a more intrinsic method.

On Figure 12, the numerical implementation is explained. We start from an initial shape  $\Omega$ , which is the unit ball  $\mathbb{B}^2$ . After applying the symplectic map  $\mathcal{T}$  on  $\mathbb{B}^2$  to obtain a computational domain  $\mathcal{T}\mathbb{B}^2$ , a mask is applied in order to delete the collocation points mapped within  $K$ .

To implement the boundary conditions, we follow the strategy presented in Section 2.4. Equipped with a PINN  $u_{\theta} : \mathbb{B}^2 \rightarrow \mathbb{R}$  and a SympNet  $T_{\omega} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , the approximate solution  $v_{\theta, \omega}$  is represented as follows:

$$v_{\theta, \omega} : \quad \mathbb{B}^2 \quad \rightarrow \quad \mathbb{R} \\ (x_1, x_2) \quad \mapsto \quad \alpha_{\omega}(x_1, x_2)(u_{\theta} \circ T_{\omega})(x_1, x_2) + \beta_{\omega}(x_1, x_2), \quad (28)$$

with  $\alpha_{\omega} = 0$  on  $\Gamma_i \cup \Gamma_e$ ,  $\beta_{\omega} = 0$  on  $\Gamma_e$  and  $\beta_{\omega} = 1$  on  $\Gamma_i$ . It is enough to choose  $\alpha_{\omega}, \beta_{\omega} : \mathbb{B}^2 \setminus T_{\omega}^{-1}K \rightarrow T_{\omega}\mathbb{B}^2 \setminus K$  such that

$$\alpha_{\omega} = \varphi_{\omega} \varphi_{\mathbb{B}^2} \quad \text{and} \quad \beta_{\omega} = \frac{\varphi_{\mathbb{B}^2}}{\varphi_{\omega} - \varphi_{\mathbb{B}^2}}, \quad (29)$$

where, for  $(x_1, x_2) \in \mathbb{B}^2 \setminus T_{\omega}^{-1}K$ ,

$$\varphi_{\omega}(x_1, x_2) = \left( \frac{x_1^{T_{\omega}}}{a} \right)^2 + \left( \frac{x_2^{T_{\omega}}}{b} \right)^2 - 1, \quad \text{with } (x_1^{T_{\omega}}, x_2^{T_{\omega}}) = T_{\omega}(x_1, x_2); \\ \varphi_{\mathbb{B}^2}(x_1, x_2) = x^2 + y^2 - \rho_{\max}^2, \quad \text{with } \rho_{\max} \text{ the radius of the ball } \mathbb{B}^2.$$

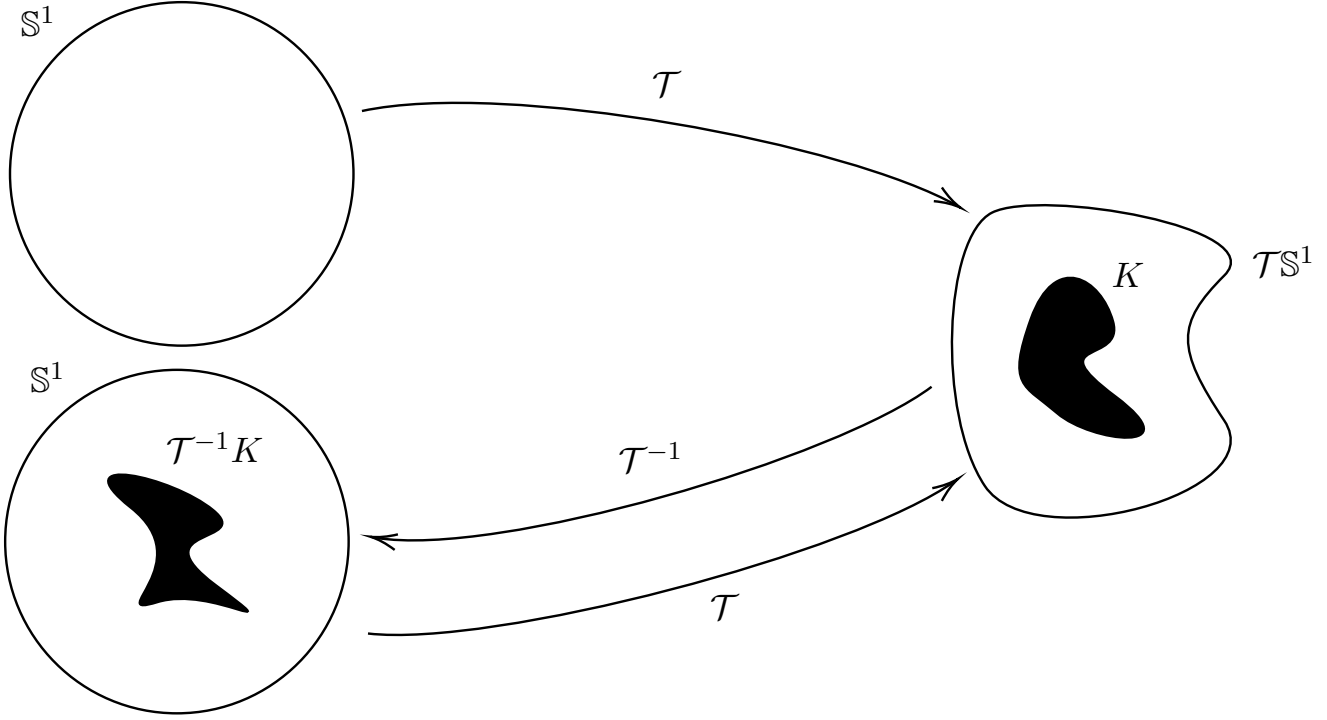


Figure 12: Sampling method for the numerical resolution of the Bernoulli overdetermined PDE.

### 3.4.2 Numerical results

We start with a case where the exact solution is known: with a centered disk-shaped obstacle (i.e., with  $a = 0.5$ ), the optimal shape will also be a disk. Indeed, assume that  $\Omega$  is a centered disk with radius  $R > 0$ . Then, one easily check that a particular solution of the Bernoulli exterior problem (27) is  $u(x, y) = (\log R - \log \sqrt{x^2 + y^2}) / \log(2R)$ , where  $R$  is the unique positive number such that  $R |\log(2R)| = 1/c$ . We conclude by uniqueness of the solution to (27) (see e.g. [19, Theorem 1.1]). The results are displayed in the top panels of Figure 13, where we observe that the optimality condition is satisfied (the optimality error is  $6.83 \times 10^{-3}$ ). Moreover, the Hausdorff distance between the learned shape and  $\mathbb{B}^2$  is  $6.23 \times 10^{-3}$ , which further confirms the relevance of our approximation.

Then, we tackle the case of an elongated ellipse ( $a = 0.65$ ). This case is harder to solve, as the optimal shape is not  $\mathbb{B}^2$ , and the SympNet has to learn a complex transformation, involving constraints on both  $\Gamma_i$  and  $\Gamma_e$ . In this case, to help train the model, we have added a term in the loss function to penalize the optimality condition on  $\Gamma_e$ . This term is activated after 22 000 epochs, and is accompanied by a reduction in the learning rate, from  $1 \times 10^{-2}$  to  $1 \times 10^{-4}$ . The results are displayed in the bottom panels of Figure 13. As we do not know the optimal shape in this case, we cannot report the Hausdorff distance; however, the optimality error is  $7.04 \times 10^{-3}$ , which is a good indication that the final shape is close to being optimal.

In both cases, we also report in Table 8 the values of the variational formulation (22), like in the previous sections, to further confirm the relevance of our approach. We observe that the values of the variational formulation are slightly larger than for Poisson's equation, but that is attributable to the difficulty of the problem as well as the larger values taken by  $u$  in this case.



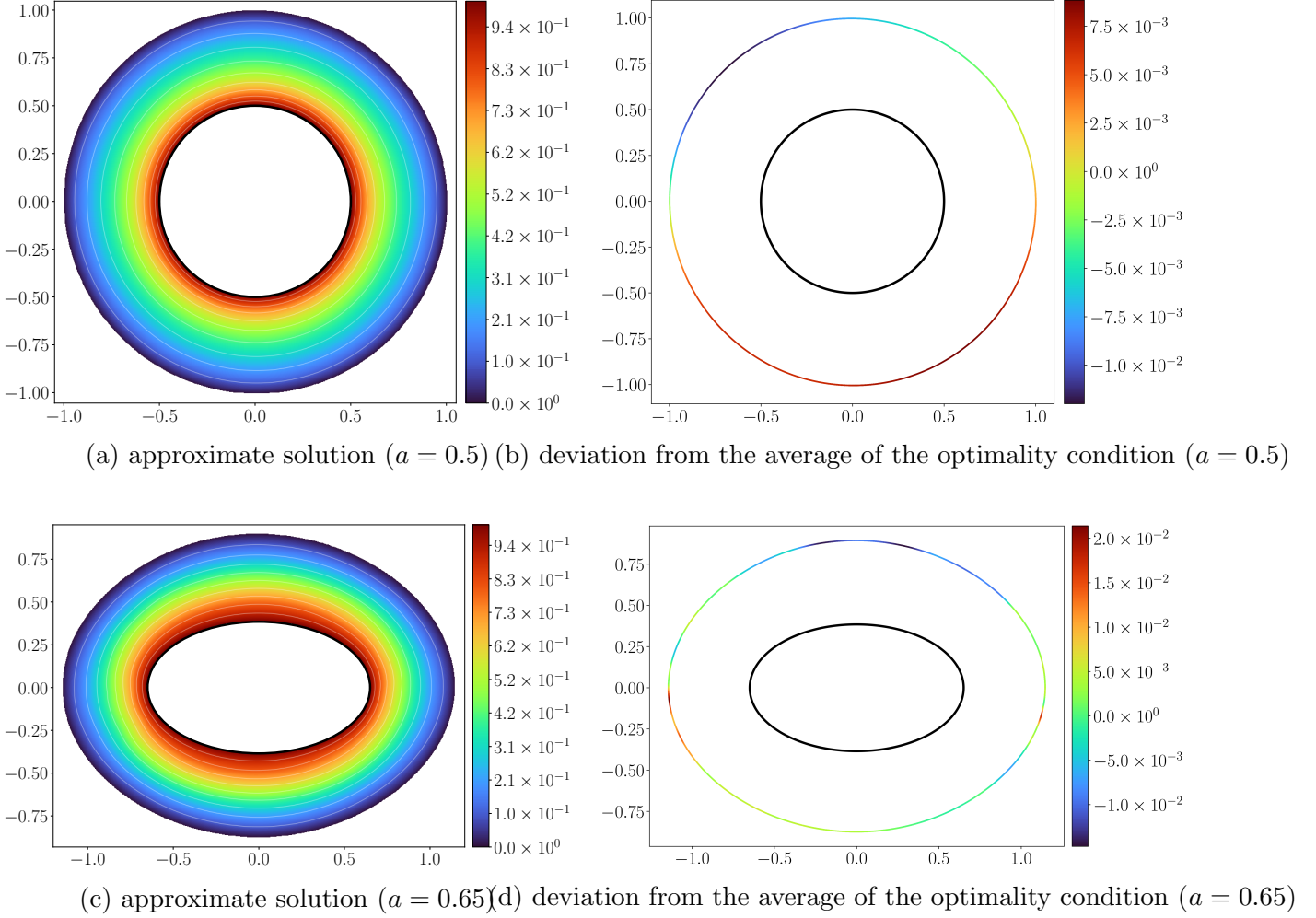


Figure 13: Approximate solutions (left panels) and optimality conditions (right panels) for the Bernoulli problem with an ellipsoid-shaped obstacle (top panels:  $a = 0.5$ , i.e., the case of a disk; bottom panels:  $a = 0.65$ ). On all panels, the black line corresponds to  $\partial K = \Gamma_i$ .

Table 8: Statistics of the variational formulation for the overdetermined Bernoulli problem, described in Section 3.4, for different values of  $a$ .

Value of $a$	Mean	Max	Min	Standard deviation
$a = 0.5$	$1.04 \times 10^{-1}$	$2.45 \times 10^{-1}$	$1.50 \times 10^{-4}$	$5.18 \times 10^{-2}$
$a = 0.65$	$8.95 \times 10^{-2}$	$2.29 \times 10^{-1}$	$5.61 \times 10^{-4}$	$4.60 \times 10^{-2}$

## 4 Minimizing the Dirichlet Energy with prescribed volume for Robin boundary conditions

To show how this methodology is generalizable to other boundary conditions, we introduce again  $D$  a compact set of  $\mathbb{R}^n$ , as well as  $\Omega$  an open subset of  $D$ ,  $f \in H^{-1}(D)$ , and a Robin coefficient  $\kappa \in \mathbb{R}^+$ . We

denote by  $u_{\Omega}^{f,\kappa} \in H^1(\Omega)$  the unique solution of the Poisson equation in  $\Omega$  with Robin boundary conditions

$$\begin{cases} -\Delta u_{\Omega}^{f,\kappa} = f & \text{in } \Omega, \\ \partial_n u_{\Omega}^{f,\kappa} + \kappa u_{\Omega}^{f,\kappa} = 0 & \text{on } \partial\Omega. \end{cases} \quad (30)$$

Moreover,  $u_{\Omega}^{f,\kappa}$  is the minimizer to an energy  $\mathcal{J}$  defined for  $v \in H^1(\Omega)$  by

$$\mathcal{J}(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 + \frac{\kappa}{2} \int_{\partial\Omega} v^2 - \int_{\Omega} f v. \quad (31)$$

This leads to the definition of a new Dirichlet energy, defined for  $\Omega \subset \mathbb{R}^n$  by

$$\mathcal{E}(\Omega) = \frac{1}{2} \int_{\Omega} |\nabla u_{\Omega}^{f,\kappa}|^2 + \frac{\kappa}{2} \int_{\partial\Omega} u_{\Omega}^{f,\kappa^2} - \int_{\Omega} f u_{\Omega}^{f,\kappa}$$

For a prescribed volume  $V_0 > 0$ , the new shape optimization problem for the Dirichlet energy of the Robin-Poisson [Equation \(30\)](#) reads

$$\inf\{\mathcal{E}(\Omega), \Omega \text{ bounded open set of } D, \text{ such that } |\Omega| = V_0\}. \quad (32)$$

**Remark 8.** Shape optimization problems involving elliptic equations with Robin boundary conditions are generally more difficult to study than those with Dirichlet conditions. In a recent series of works, the authors introduced a suitable relaxation of this type of problem in the space of special functions with bounded variation, as developed by De Giorgi and Ambrosio. This approach yields the existence of a generalized minimizer [\[5, 7\]](#). In certain problems, the authors conducted further analysis to prove the existence of a "regular" optimal domain (such as an open set). Let us highlight that the case where  $f(\cdot) = 1$  has been fully studied: Saint-Venant's inequality indicates that the optimal domain is a ball [\[6\]](#). Generally, determining the regularity of the minimizers for this problem is not straightforward. In the scope of this article, we do not dwell on these questions, especially since we are seeking a local minimizer for this shape optimization problem.

**Theorem 2.** *Assume that there exists a  $C^2$  optimal shape  $\Omega^*$ . If  $f \in H_{loc}^1(\mathbb{R}^n)$ , then the solution of [\(30\)](#)  $u^*$  belongs to  $H^3(\Omega^*)$ . Moreover, the first order optimality condition on  $\partial\Omega^*$  reads*

$$\exists c \in \mathbb{R} \quad | \quad \frac{1}{2} |\nabla u^*|^2 + \frac{u^*}{2} (\kappa H^* - 2\kappa^2) - f = c, \quad (33)$$

where  $H^*$  is the mean curvature of  $\partial\Omega^*$ .

For the sake of completeness, the proof of this result, although standard, is provided in [Appendix A](#).

Equipped with these results, we now explain how to solve the shape optimization problem in [Section 4.1](#). Then, we validate our methodology on several numerical experiments in [Section 4.2](#).

## 4.1 Solving the shape optimization problem with Robin boundary conditions

To solve the shape optimization problem [\(32\)](#) associated to the Robin-Poisson problem [\(30\)](#), the same change of variables as in [Lemma 2](#) is required. We introduce  $w = u \circ \mathcal{T}$  with  $\mathcal{T}$  a symplectic map and  $u$  the solution of the Poisson-Robin problem in  $\mathcal{T}\mathbb{B}^2$ . The following result states the equation whose solution is  $w$  in  $\mathbb{B}^2$ .



**Lemma 3.** Let  $\mathcal{T}$  be a symplectic map on  $\mathbb{R}^2$  and  $u_{\mathcal{T}} \in H^1(\mathcal{T}\mathbb{B}^2)$ , such that  $w = u_{\mathcal{T}} \circ \mathcal{T} \in H^1(\mathbb{B}^2)$ . If  $u_{\mathcal{T}} \in H_0^1(\mathcal{T}\mathbb{B}^2)$  is the solution of the Robin-Poisson problem (30), then  $w \in H^1(\mathbb{B}^2)$  is the weak solution of

$$\inf_{v \in H^1(\mathbb{B}^2)} \left\{ \frac{1}{2} \int_{\mathbb{B}^2} (A \nabla v \cdot \nabla v) + \frac{\kappa}{2} \int_{\partial \mathbb{B}^2} |J_{\mathcal{T}}^{-\top} n| v^2 - \int_{\mathbb{B}^2} \tilde{f} v \right\}, \quad (34)$$

where  $A : \mathbb{B}^2 \rightarrow \mathcal{M}_2(\mathbb{R})$  is defined by  $A = J_{\mathcal{T}}^{-1} J_{\mathcal{T}}^{-\top}$  (with  $J_{\mathcal{T}}$  the Jacobian matrix of  $\mathcal{T}$ ),  $\tilde{f} = f \circ \mathcal{T}$ , and  $n$  is the outwards unit normal of  $\mathbb{S}^1$ .

*Proof.* The proof of Lemma 3 follows the same reasoning as the proof of Lemma 2. Therefore, we omit it here. Nevertheless, we point out that the variational formula of (30) has a boundary term that involves the tangential Jacobian determinant  $|J_{\mathcal{T}}^{-\top} n|$  during the change of variables.  $\square$

**Remark 9.** The boundary conditions of (30) can be derived from the variational formulation associated with the energetic formulation (34). Notably, they do not affect the space of test functions. Consequently, for numerical methods, minimizing (31) is sufficient to satisfy the Robin boundary conditions. Thus, the methodology for the numerical resolution of the Dirichlet energy minimization presented in the previous section can be directly adapted for (30) without the need for any additional function to enforce the boundary condition.

The loss function to be minimized with respect to the PINN and the SympNet thus becomes

$$\begin{aligned} \mathcal{J}_{\text{Robin}}(\theta, \omega; \{x_{\Omega}^i, x_{\Gamma}^i, \mu^i\}_{i=1}^N) &= \frac{V_0}{N} \sum_{i=1}^N \left\{ \frac{1}{2} A_{\omega} \nabla v_{\theta, \omega} \cdot \nabla v_{\theta, \omega} \right\} (x_{\Omega}^i; \mu^i) \\ &+ \kappa \frac{2\sqrt{\pi} V_0}{N} \sum_{i=1}^N \frac{1}{2} \left\{ \mathcal{J}_{\omega}^{\Gamma} v_{\theta, \omega}^2 \right\} (x_{\Gamma}^i; \mu^i) \\ &- \frac{V_0}{N} \sum_{i=1}^N \left\{ \tilde{f}_{\omega} v_{\theta, \omega} \right\} (x_{\Omega}^i; \mu^i), \end{aligned} \quad (35)$$

with

- $\theta$  the trainable weights of the PINN,  $\omega$  the trainable weights of the SympNet;
- $T_{\omega} : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$  the SympNet;
- $u_{\theta} : T_{\omega} \mathbb{B}^2 \rightarrow \mathbb{R}$  the PINN;
- $v_{\theta, \omega} = u_{\theta} \circ T_{\omega}$  the solution of the Poisson problem set in  $T_{\omega} \mathbb{B}^2$ ;
- $A_{\omega}$  the diffusion matrix defined by  $A_{\omega} = J_{T_{\omega}}^{-1} J_{T_{\omega}}^{-\top}$ ;
- $J_{\omega}^{\Gamma} = |J_{T_{\omega}}^{-\top} n|$  the tangential Jacobian determinant defined on  $\mathbb{S}^1$ , with  $n$  the normal outward vector of  $\mathbb{S}^1$ , and  $|\cdot|$  the euclidean norm in  $\mathbb{R}^2$ ;
- $\tilde{f}_{\omega} : x \in \mathbb{B}^2 \mapsto (f \circ T_{\omega})(x) \in \mathbb{R}$ ;
- $N$  the number of collocation points;
- $\{x_{\Omega}^i, x_{\Gamma}^i, \mu^i\}_{i=1}^N \in (\mathbb{B}^2 \times \mathbb{S}^1 \times \mathbb{M})^N$  a set of random collocation points and values of the parameters.

## 4.2 Numerical experiments

Now, we proceed with numerical experiments. We first treat non-parametric shape optimization problems, where the source term  $f$  is either constant (in [Section 4.2.1](#)) or given by a radial function (in [Section 4.2.2](#)). Then, in [Section 4.2.3](#), we consider a parametric shape optimization problem, where the Robin coefficient is parametric.

**Remark 10.** For the remainder of the article, and with a slight abuse of notation, the wording “optimality error” will refer to the standard deviation of (33). More precisely, we will verify the local optimality of the obtained shape by examining if

$$\text{optimality error} := \sqrt{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} (\rho(x) - \bar{\rho})^2 \, d\sigma} = 0,$$

where  $\rho = \frac{1}{2}|\nabla u^*|^2 + \frac{u^*}{2}(\kappa H^* - 2\kappa^2) - f$  on  $\partial\Omega$ , and  $\bar{\rho}$  denotes the average value of  $\rho$  on  $\partial\Omega$ .

### 4.2.1 Non-parametric problem: results with $f = 1$ and $\kappa = 1$

The results with  $f = 1$  are displayed on [Figure 14](#). In this case, the optimal shape is known to be the unit sphere and the solution of the Robin-Poisson problem (30) writes  $u(x, y) = \frac{3-x^2-y^2}{4}$  [6]. Remark that, as the source term is constant, the optimal shape is not necessary centered at the origin, like in [Section 3.3.1](#). To get an idea of the quality of our results presented in [Figure 14](#), we looked at the distance between the learned and the optimal shape (top left panel), as well as the deviation from the average of the optimality condition (top right panel), and the pointwise error between the exact solution and the approximated solution (bottom right panel). Of course, the approximate solution to the Robin-Poisson problem (30) is also displayed (bottom left panel).

As with Dirichlet boundary conditions, we report some metrics in [Table 9](#). Remark that the Hausdorff distance, the standard deviation of the optimality condition and the  $L^2$  error reported in [Table 9](#) are of the same order as the one obtained in [Table 3](#). This is a good indicator of the robustness of our approach with respect to the boundary conditions

Table 9: Relevant metrics related to the shape optimization of the Robin-Poisson equation with  $f = 1$  and  $\kappa = 1$ .

Hausdorff distance	Optimality error	$L^2$ error
$9.86 \times 10^{-3}$	$4.99 \times 10^{-3}$	$4.13 \times 10^{-3}$

### 4.2.2 Non-parametric problem: results with $f$ given by (25) with $\mu = 0.8$ and with $\kappa = 1$

To graduate the difficulty of problems to be solved numerically, we turn to an exponential source term  $f$  given by (25) with  $\mu = 0.8$ , and display the results on [Figure 15](#). Because of the lack of knowledge we have on the potential optimal solution, the only relevant metric we can look at is the optimality condition (33). For this numerical experiment the optimality error is  $1.87 \times 10^{-2}$ . In our opinion, this is a very good result.

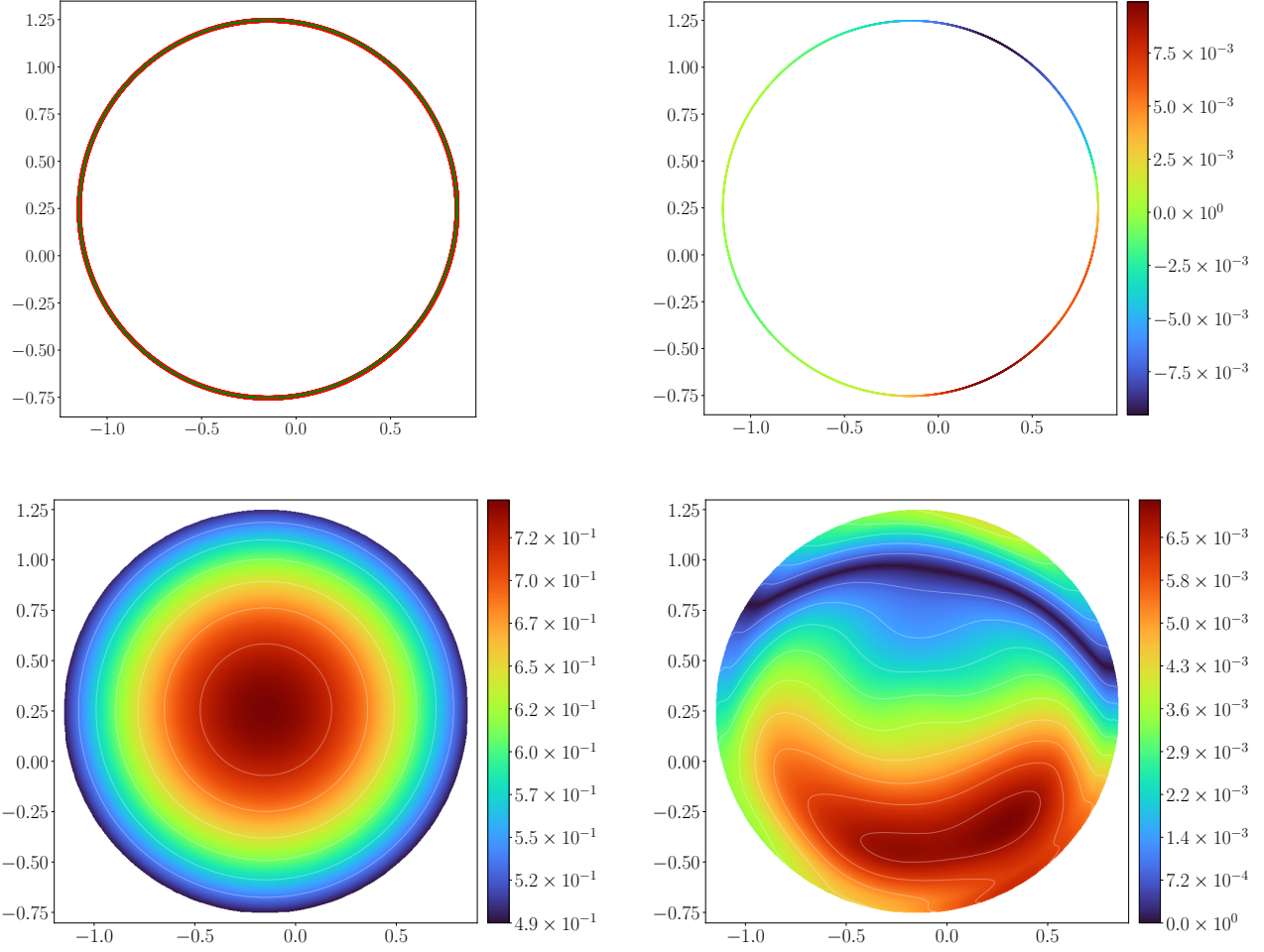


Figure 14: Shape optimization for the Dirichlet energy, for the Robin-Poisson equation with source term  $f = 1$  and  $\kappa = 1$ . Top left panel: learned shape (green line) and reference shape (red line). Top right panel: deviation from the average of the optimality condition in the learned shape. Bottom left panel: approximate PDE solution. Bottom right panel: pointwise error between the true and approximate PDE solutions.

#### 4.2.3 Parametric problem: results with $f = 1$ and $\kappa \in (0.5, 1.5)$

We now turn to a parametric shape optimization problem for  $f = 1$  and  $\kappa \in \mathbb{M} = (0.5, 1.5)$ . As described in [6], the optimal shape is a disk, and thus the solution of the Robin-Poisson problem is  $u(x, y) = \frac{2+\kappa}{4\kappa} - \frac{x^2+y^2}{4}$ . The results and errors are displayed on Figure 16, where there is a good agreement between approximate and exact solution. In addition, we display the optimal shapes and optimality error for 10 random values of  $\mu$  on Figure 17. Finally, we report some statistics on our three main metrics in Table 10, which confirm the relevance of our approach for a parametric problem.

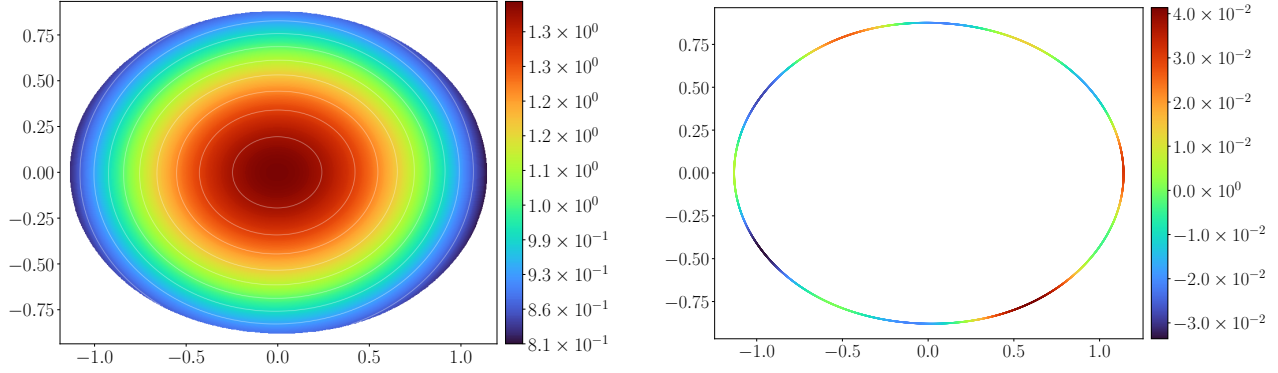


Figure 15: Shape optimization for the Dirichlet energy, for the Poisson equation with source term  $f$  given by (25) with  $\mu = 0.8$  and  $\kappa = 1$  and with Robin boundary conditions. The results of a fixed-point algorithm [9] is our reference for this numerical experiment. Left panel: approximate PDE solution in the learned shape. Right panel: deviation from the average of the optimality condition.

Table 10: Statistics of relevant metrics in the case of a constant source term  $f = 1$  and a parametric constant Robin coefficient  $\kappa$ , obtained by computing each metric for  $10^3$  values of  $\kappa$ .

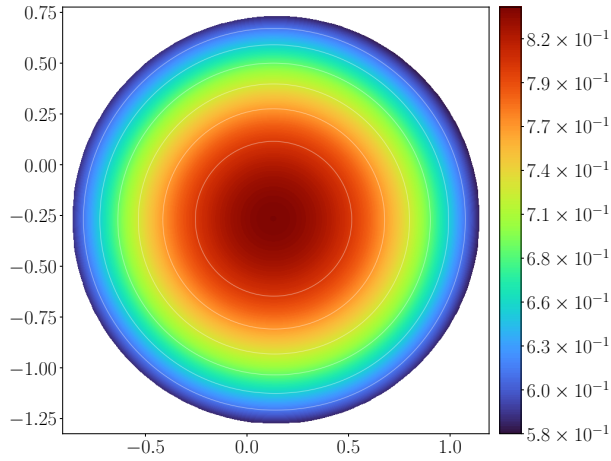
Metric	Mean	Max	Min	Standard deviation
Hausdorff distance	$5.96 \times 10^{-3}$	$1.08 \times 10^{-2}$	$2.11 \times 10^{-3}$	$2.87 \times 10^{-3}$
optimality error	$4.22 \times 10^{-3}$	$8.81 \times 10^{-3}$	$2.71 \times 10^{-3}$	$3.50 \times 10^{-3}$
$L^2$ error	$4.20 \times 10^{-3}$	$1.05 \times 10^{-2}$	$3.27 \times 10^{-3}$	$8.08 \times 10^{-4}$

## 5 Conclusion

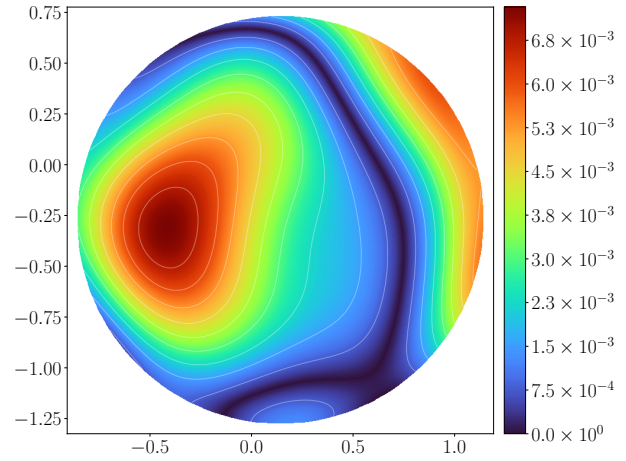
In this work, we developed a flexible Proof of Concept for geometric shape optimization using neural networks. We minimized the Dirichlet energy under volume constraints by approximating the Poisson equation’s solution and shape transformations with physics-informed neural networks (PINNs). Our combined optimization algorithm successfully handled Dirichlet and Robin boundary conditions, parametric right-hand sides or Robin coefficients, and Bernoulli-type free boundary problems. Moreover, it is very easily parallelizable.

A significant advantage of our methodology is that it does not require the calculation of the shape derivative. Additionally, our approach has the potential to address non-self-adjoint cases and could be adapted to handle non-energetic criteria. This flexibility makes our method a promising tool for various complex shape optimization problems.

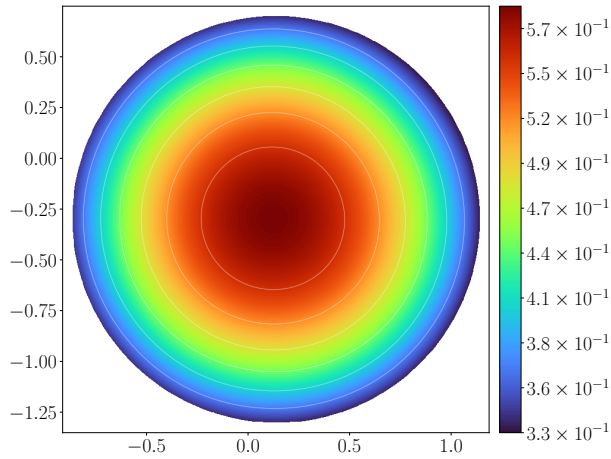
This work should be regarded as an initial step, with many questions still open for future exploration. One significant challenge is extending our approach to PDE physical models in higher dimensions. While our current work addresses a two-dimensional case, our goal is to develop methods capable of handling problems in dimensions larger than three. This challenge is intricately linked to incorporating volume constraints and, more broadly, managing general manufacturing constraints in any dimension. A simplistic approach might involve penalizing constraints, but this often conflicts with learning algorithms, as it can lead to poorly conditioned systems.



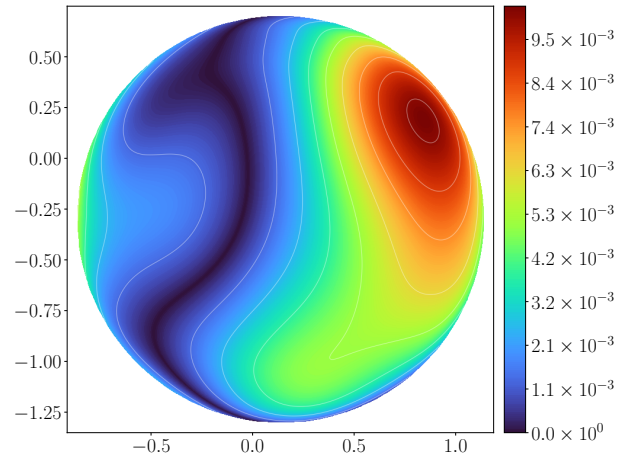
(a) solution,  $\mu = 0.86$



(b) error,  $\mu = 0.86$



(c) solution,  $\mu = 1.48$



(d) error,  $\mu = 1.48$

Figure 16: Shape optimization for the Dirichlet energy, for the Robin-Poisson equation with source term  $f = 1$  and  $\kappa \in (0.5, 1.5)$ . For two values of  $\kappa$ , we display the approximate solution (left column) and the pointwise error between the approximate and exact solutions (right column).

A particularly ambitious direction for future research is to tackle physical problems that traditional shape optimization methods find challenging. For instance, optimizing criteria related to solutions of turbulent fluid dynamics equations is a complex area that warrants further investigation. This will be a key focus of our future efforts.

## Acknowledgments

We would like to sincerely thank Caroline Vernier and Killian Lutz for helpful conversations and their comments on the problem. We would like to thank Konrad Janik as well, for enlightening discussions on parametric SympNets.

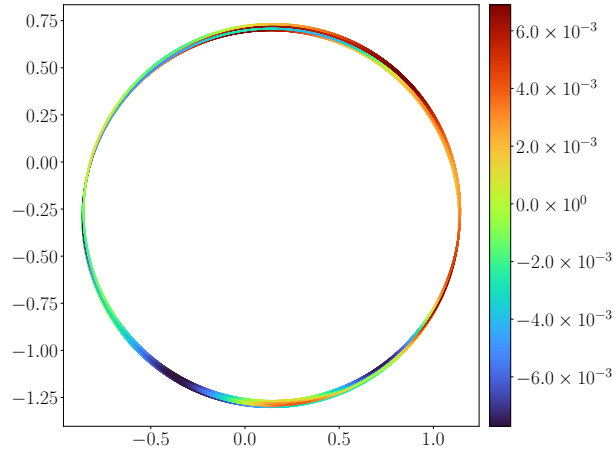


Figure 17: Shape optimization for the Dirichlet energy, for the Robin-Poisson equation with source term  $f = 1$  and  $\kappa \in [0.5, 1.5]$ : optimal shapes and deviation from the average of the optimality condition for 10 random values of  $\kappa$ .

This work was supported by PEPR PDE-AI and by PEPR Numpex Exa-MA. The last author were partially supported by the ANR Project “STOIQUES”.

## References

- [1] G. Allaire. *Conception optimale de structures*, volume 58 of *Mathématiques et Applications*. Springer Berlin Heidelberg, 2006.
- [2] G. Allaire, C. Dapogny, and F. Jouve. Shape and topology optimization. In *Handbook of numerical analysis*, volume 22, pages 1–132. Elsevier, 2021.
- [3] G. Allaire, F. Jouve, and A.-M. Toader. A level-set method for shape optimization. *C. R. Math. Acad. Sci. Paris*, 334(12):1125–1130, 2002.
- [4] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics. Springer New York, 1989.
- [5] D. Bucur and A. Giacomini. Faber–krahm inequalities for the robin-laplacian: A free discontinuity approach. *Archive for Rational Mechanics and Analysis*, 218(2):757–824, 2015.
- [6] D. Bucur and A. Giacomini. The saint-venant inequality for the laplace operator with robin boundary conditions. *Milan Journal of Mathematics*, 83(2):327–343, 2015.
- [7] D. Bucur, A. Giacomini, and P. Trebeschi. The robin–laplacian problem on varying domains. *Calculus of Variations and Partial Differential Equations*, 55:1–29, 2016.
- [8] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.*, 7:1–49, 1998.
- [9] A. Chambolle, I. Mazari-Fouquer, and Y. Privat. Stability of optimal shapes and convergence of thresholding algorithms in linear and spectral optimal control problems. preprint, June 2023.

- [10] S. Cuomo, F. Giampaolo, S. Izzo, C. Nitsch, F. Piccialli, and C. Trombetti. A physics-informed learning approach to Bernoulli-type free boundary problems. *Comput. Math. Appl.*, 128:34–43, 2022.
- [11] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries: metrics, analysis, differential calculus, and optimization*. SIAM, 2011.
- [12] W. E and B. Yu. *The Deep Ritz Method: A Deep Learning-Based Numerical Algorithm for Solving Variational Problems*, 2018.
- [13] E. Franck, V. Michel-Dansac, and L. Navoret. Approximately well-balanced Discontinuous Galerkin methods using bases enriched with Physics-Informed Neural Networks. *J. Comput. Phys.*, 2024.
- [14] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [15] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer Series in Computational Mathematics. Springer-Verlag, 2006.
- [16] F. Hecht. New Development in FreeFem++. *J. Numer. Math.*, 20(3–4), 2012.
- [17] A. Henrot. *Shape optimization and spectral theory*. De Gruyter Open, 2017.
- [18] A. Henrot and M. Pierre. *Shape Variation and Optimization*. European Mathematical Society Publishing House, Feb. 2018.
- [19] A. Henrot and H. Shahgholian. Existence of classical solutions to a free boundary problem for the p-Laplace operator: (I) the exterior convex case. *J. reine angew. Math.*, 2000(521), 2000.
- [20] H. Jeong, J. Bai, C. Batuwatta-Gamage, C. Rathnayaka, Y. Zhou, and Y. Gu. A physics-informed neural network-based topology optimization (pinnto) framework for structural optimization. *Engineering Structures*, 278:115484, 2023.
- [21] P. Jin, Z. Zhang, I. G. Kevrekidis, and G. E. Karniadakis. Learning Poisson Systems and Trajectories of Autonomous Systems via Poisson Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.*, 34(11):8271–8283, 2023.
- [22] P. Jin, Z. Zhang, A. Zhu, Y. Tang, and G. E. Karniadakis. SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems. *Neural Networks*, 132:166–179, 2020.
- [23] M. Kac. Can one hear the shape of a drum? *The american mathematical monthly*, 73(4P2):1–23, 1966.
- [24] D. Kingma and J. Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.
- [25] I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans. Neural Netw.*, 9(5):987–1000, 1998.
- [26] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. OUP Oxford, 2009.
- [27] M. Nakahara. *Geometry, Topology and Physics*. CRC Press, 2018.
- [28] A. Odot, G. Mestdagh, Y. Privat, and S. Cotin. Real-time elastic partial shape matching using a neural network-based adjoint method. In *International Conference on Optimization and Learning*, pages 137–147. Springer, 2023.



- [29] A. Paszke and S. e. a. Gross. *PyTorch: an imperative style, high-performance deep learning library*, pages 8026–8037. Curran Associates Inc., Red Hook, NY, USA, 2019.
- [30] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, 378:686–707, 2019.
- [31] A. A. Taha and A. Hanbury. An Efficient Algorithm for Calculating the Exact Hausdorff Distance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(11):2153–2163, 2015.
- [32] M. Tonnel, A. Wirbel, F. Oesterle, and J.-T. Fischer. AvaFrame com1DFA (v1.3): a thickness-integrated computational avalanche module – theory, numerics, and testing. *Geosci. Model Dev.*, 16(23):7013–7035, 2023.
- [33] Z. Zhang, C. Lin, and B. Wang. Physics-informed shape optimization using coordinate projection. *Scientific Reports*, 14(1):6537, 2024.

## A Necessary first order optimality condition for Problem (32)

Let us compute the shape derivative of the functional  $\mathcal{E}$  defined by

$$\mathcal{E}(\Omega) = \frac{1}{2} \int_{\Omega} |\nabla u|^2 + \frac{\kappa}{2} \int_{\partial\Omega} u^2 - \int_{\Omega} fu, \quad (36)$$

with  $u$  solution of the Robin-Poisson problem (30).

In what follows, the notation  $\nabla_{\Gamma}$  stands for the tangential gradient.

By multiplying the main equation of (30) by  $u$  and integrating then by parts drives us to

$$\int_{\Omega} |\nabla u|^2 + \kappa \int_{\partial\Omega} u^2 = \int_{\Omega} fu, \quad (37)$$

so that

$$\mathcal{E}(\Omega) = -\frac{1}{2} \int_{\Omega} fu. \quad (38)$$

Let  $V \in W^{2,\infty}(\mathbb{R}^2, \mathbb{R}^2)$  denote a compactly supported a vector field. In what follows, we extend the normal vector  $n$  to a neighborhood of  $\partial\Omega$ , in such a way that  $n$  remains unitary in a neighborhood of  $\partial\Omega$ . This way, all the upcoming calculations will make sense and, in particular, will allow us to define the quantity  $\nabla_{\Gamma}(V \cdot n)$ . Therefore,  $\|n\|^2 = 1$  in this neighborhood which implies in particular that  $(Dn)n = 0$ .

Then, according to [18, Theorem Chapter 5],

$$\langle d\mathcal{E}(\Omega), V \rangle = -\frac{1}{2} \left( \int_{\Omega} fu' + \int_{\partial\Omega} fu(V \cdot n) \right). \quad (39)$$

where the Eulerian derivative  $u'$  of  $u$  solves

$$\begin{cases} -\Delta u' = 0 & \text{in } \Omega; \\ \frac{\partial u'}{\partial n} + \kappa u' = +\nabla u \cdot \nabla_{\Gamma}(V \cdot n) - \left( \frac{\partial^2 u}{\partial n^2} + \kappa \frac{\partial u}{\partial n} \right) (V \cdot n) & \text{on } \partial\Omega. \end{cases} \quad (40)$$

We now multiply the equation (30) solved by  $u$  by  $u'$  and we integrate by parts to obtain with Green formula

$$\int_{\Omega} \nabla u \cdot \nabla u' + \kappa \int_{\partial\Omega} uu' = \int_{\Omega} fu', \quad (41)$$



symmetrically, by reversing the role of  $u$  and  $u'$ , and by using (40), we get

$$\int_{\Omega} \nabla u' \cdot \nabla u + \kappa \int_{\partial\Omega} u' u - \int_{\partial\Omega} u \nabla u \cdot \nabla_{\Gamma}(V \cdot n) + \int_{\partial\Omega} u \left( \frac{\partial^2 u}{\partial n^2} + \kappa \frac{\partial u}{\partial n} \right) (V \cdot n) = 0. \quad (42)$$

By combining (41) and (42), we obtain

$$\int_{\Omega} f u' = + \int_{\partial\Omega} u \nabla u \cdot \nabla_{\Gamma}(V \cdot n) - \int_{\partial\Omega} u \left( \frac{\partial^2 u}{\partial n^2} + \kappa \frac{\partial u}{\partial n} \right) (V \cdot n). \quad (43)$$

By simplifying  $\int_{\partial\Omega} u \nabla u \cdot \nabla_{\Gamma}(V \cdot n)$  with an integration by parts on  $\partial\Omega$  and using an orthogonal decomposition of the gradient, we obtain

$$\int_{\partial\Omega} u \nabla u \cdot \nabla_{\Gamma}(V \cdot n) = \int_{\partial\Omega} \frac{1}{2} \nabla_{\Gamma}(u^2) \cdot \nabla_{\Gamma}(V \cdot n). \quad (44)$$

Since  $\partial\Omega$  is assumed to be  $C^2$ ,  $V \cdot n \in H^2(\Omega)$  and  $u \in H^3(\Omega)$ , we have [18, Section 5.4, eq. 5.59]

$$\int_{\partial\Omega} u \nabla u \cdot \nabla_{\Gamma}(V \cdot n) = \int_{\partial\Omega} (V \cdot n) \left( -\frac{1}{2} \Delta_{\Gamma}(u^2) \right), \quad (45)$$

where  $\Delta_{\Gamma}$  denotes the Laplace-Beltrami operator. Using

$$\frac{1}{2} \Delta_{\Gamma}(u^2) = u \Delta_{\Gamma} u + |\nabla_{\Gamma} u|^2, \quad (46)$$

and the orthogonal decomposition of the Laplacian

$$\Delta_{\Gamma} u = \Delta u - H \frac{\partial u}{\partial n} - \frac{\partial^2 u}{\partial n^2} = -f - H \frac{\partial u}{\partial n} - \frac{\partial^2 u}{\partial n^2}, \quad (47)$$

where  $H = \operatorname{div}_{\Gamma}(n)$  denotes the mean curvature of  $\partial\Omega$ , and  $\operatorname{div}_{\Gamma}$  the tangential divergence, we come back to (43) and infer that

$$\int_{\Omega} f u' = \int_{\partial\Omega} (V \cdot n) \left( u \left( f + H \frac{\partial u}{\partial n} + \frac{\partial^2 u}{\partial n^2} \right) - |\nabla_{\Gamma} u|^2 - u \left( \frac{\partial^2 u}{\partial n^2} + \kappa \frac{\partial u}{\partial n} \right) \right). \quad (48)$$

Note that

$$\frac{\partial u}{\partial n} = -\kappa u \quad \text{a.e. on } \partial\Omega, \quad (49)$$

so that

$$|\nabla_{\Gamma} u|^2 = |\nabla u|^2 - \left| \frac{\partial u}{\partial n} \right|^2 = |\nabla u|^2 - \kappa^2 u^2. \quad (50)$$

Now, (48) rewrites

$$\int_{\Omega} f u' = \int_{\partial\Omega} (V \cdot n) (f u - u^2(\kappa H - 2\kappa^2) - |\nabla u|^2). \quad (51)$$

Coming back to (39), the computations above yield to

$$\langle d\mathcal{E}(\Omega), V \rangle = \int_{\partial\Omega} (V \cdot n) \left( \frac{1}{2} |\nabla u|^2 + \frac{u^2}{2} (\kappa H - 2\kappa^2) - f u \right). \quad (52)$$

Thus, because of the volume constraint, there exists a Lagrange multiplier  $c \in \mathbb{R}$  such that

$$\frac{1}{2} |\nabla u|^2 + \frac{u^2}{2} (\kappa H - 2\kappa^2) - f u = c, \quad \text{on } \partial\Omega, \quad (53)$$

we refer for instance to [18, Section 6.1.3].

## B Network hyperparameters

This appendix provides the hyperparameters used for the networks in the numerical experiments. In [Appendix B](#),  $\eta_P$  and  $\eta_S$  are the learning rates of the PINN and SympNet, respectively,  $\sigma_P$  and  $\sigma_S$  are their respective activation functions,  $N$  is the number of collocation points, and  $q$  and  $\delta$  are respectively the number of layers and width of the SympNet.

Table 11: Hyperparameters of the networks used in [Section 3](#) and in [Section 4](#).

Figure	layers	$\eta_P$	$\sigma_P$	$q$	$\delta$	$\eta_S$	$\sigma_S$	$N$	epochs
<a href="#">Figure 4</a>	[20, 40, 40, 20]	$1 \times 10^{-3}$	tanh	—	—	—	—	20 000	1500
<a href="#">Figure 5</a>	—	—	—	8	10	$1 \times 10^{-2}$	sigmoid	10 000	1000
<a href="#">Figure 6</a>	[10, 20, 20, 10]	$5 \times 10^{-3}$	tanh	4	5	$5 \times 10^{-3}$	sigmoid	5000	750
<a href="#">Figure 7</a>	[10, 20, 40, 40, 20, 10]	$5 \times 10^{-3}$	tanh	8	4	$5 \times 10^{-3}$	sigmoid	5000	9000
<a href="#">Figure 8</a>	[10, 20, 40, 40, 20, 10]	$5 \times 10^{-3}$	tanh	8	4	$5 \times 10^{-3}$	tanh	10 000	1000
<a href="#">Figure 10</a>	[10, 20, 40, 40, 20, 10]	$5 \times 10^{-3}$	tanh	8	6	$5 \times 10^{-3}$	tanh	10 000	2000
<a href="#">Figure 11</a>	[10, 20, 40, 40, 20, 10]	$5 \times 10^{-3}$	tanh	8	6	$5 \times 10^{-3}$	tanh	10 000	2000
<a href="#">Figure 13</a>	[10, 20, 40, 40, 20, 10]	$1 \times 10^{-2}$	tanh	8	4	$1 \times 10^{-2}$	sigmoid	2500	15 000
<a href="#">Figure 14</a>	[10, 20, 40, 40, 20, 10]	$1 \times 10^{-2}$	tanh	8	5	$1 \times 10^{-2}$	tanh	10 000	10 000
<a href="#">Figure 15</a>	[10, 20, 40, 40, 20, 10]	$1 \times 10^{-2}$	tanh	8	5	$1 \times 10^{-2}$	tanh	10 000	10 000
<a href="#">Figure 16</a>	[10, 20, 40, 40, 20, 10]	$1 \times 10^{-2}$	tanh	8	5	$1 \times 10^{-2}$	tanh	10 000	10 000