Learning divergence-free and curl-free vector fields with matrix-valued kernels

Ives Macêdo [1] Rener Castro [2]

ijamj@impa.br rener@mat.puc-rio.br

[1] Instituto Nacional de Matemática Pura e Aplicada (IMPA)
[2] Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)

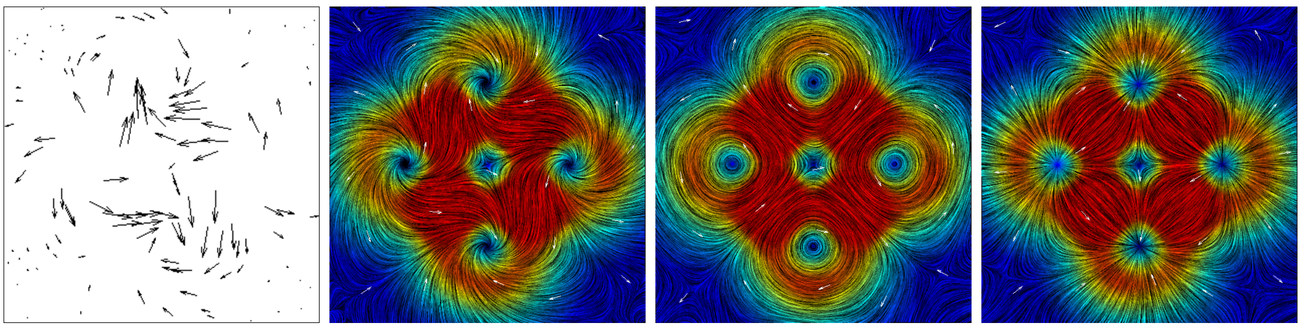

Figure 1. Learning a vector field decomposition: samples, learned field, divergence- and curl-free parts.

Abstract

We propose a novel approach for reconstructing vector fields in arbitrary dimension from an unstructured, sparse and, possibly, noisy sampling. Moreover, we are able to guarantee certain invariant properties on the reconstructed vector fields which are of great importance in applications (e.g. divergence-free vector fields corresponding to incompressible fluid flows), a difficult task which we know of no other method that can accomplish it in the same general setting we work on. Our framework builds upon recent developments in the statistical learning theory of vector-valued functions and results from the approximation theory of matrix-valued radial basis functions. As a direct byproduct of our framework, we present a very encouraging result on learning a vector field decomposition almost "for free".

Keywords: *vector field reconstruction, statistical learning, kernel methods, support vector regression, scattered data approximation, radial basis functions, geometric modeling.*

1. Introduction

Interpreted as velocities of fluid particles [1], optical flow fields [5] or directions of strokes in a painting, vector fields are ubiquitous objects in computer graphics, vision and engineering. The classical theory of physics is built upon the characterization of vector fields induced by the motion of an object. This fact is the main responsible for the pervasiveness of vector fields in graphics applications, where often they appear as measurements of real phenomena, data from physical simulations or even painting sessions.

Many physical phenomena can be characterized by vector fields with certain differential invariants, where derivatives of the component scalar fields are coupled by some relation. Important examples are the null divergence condition in incompressible fluid flows and the curl-free condition for the magnetic field in the equations of classical electrodynamics. Ensuring these physically-based constraints when representing vector fields for scientific computing or visualization has been a hard goal in designing methods for storing and manipulating them. These difficulties are even more pronounced these days with the development of techniques for measuring and simulating flow fields for aerodynamics and hydrodynamics research. They generate huge amounts of unstructured and sparse points-vectors data, and very noisy measurements [10]. This poses a hard task in designing robust methods to reconstruct globally defined vector fields, which is made even more complicated when it is required that the reconstructed vector field obeys the invariants of the field from which the samples were drawn.

Previous and related works. Since our approach for vector field reconstruction was built upon recent developments on *statistical learning* and *function approximation*, we will provide some comments on works which took one these avenues and on those theoretical efforts which influenced us.

Kuroe *et al.* [9] present a learning method for 2D vector field approximation based on artificial neural networks (ANN). They modified the classical *backpropagation algorithm* for learning a vector-valued ANN minimizing the squared loss and are able to represent the learned field as a sum of divergence- and curl-free parts (result also achieved by Mussa-Ivaldi [13], which used a related regularized least squares approach). However, their method is hard to generalize for arbitrary dimension or loss functions (drawbacks shared by [13]) and is inherently prone to the already difficult task of designing suitable ANN architectures.

Colliez *et al.* [5] adopt the classical ϵ-SVR algorithm for estimating optical flow fields in object tracking. They observe the robustness of support vector regression against outliers and its generalization abilities. Nevertheless, they apply the *linear* ϵ-SVR component-wise in the flow field not guaranteeing preservation of differential invariants.

Castro *et al.* [4] employ the *nonlinear* (kernel-based) extension of the scalar ϵ-SVR using two different representations of the vector field: *cartesian* and *polar/spherical*. The main drawbacks are that they cannot enforce invariants easily, taking spatial derivatives of a field learned in polar/spherical coordinates is cumbersome and their approach is hard to generalize for arbitrary dimensions (besides the trivial component-wise learning in cartesian coordinates).

Methods based on approximation strategies include the work of Zhong *et al.* [20], which try to both interpolate the samples and enforce the divergence-free constraint. However, this constraint enforcement only holds *at the grid points* and no such guarantee can be given elsewhere. Lage *et al.* [10] propose an efficient multilevel scheme for 2D vector field reconstruction based on the *partition of unity* method in which local polynomial approximants are estimated by a *least squares fitting* regularized using *ridge regression* (Tikhonov regularization). Nevertheless, their algorithm cannot easily enforce differential invariants for the global approximant and becomes cumbersome to implement for higher dimensions.

Recently, Micchelli and Pontil [12] have generalized the theory of scalar-valued statistical learning [8] to vector-valued functions. Under their theoretical framework, many classical statistical techniques (e.g. *ridge regression*) and some state-of-the-art learning algorithms (e.g. *support vector regression* [3]) could be extended to learn vector-valued functions in which correlations among the component functions can be captured by designing a suitable matrix-valued kernel function.

In the approximation community, Narcowich and Ward [14] introduced a construction of matrix-valued radial basis functions which happen to satisfy the properties required in [12] and induce *everywhere divergence-free* vector fields. Later, Lowitzsch [11] proved that such matrix-valued RBFs can be used to approximate arbitrarily well any sufficiently smooth divergence-free vector fields and introduced a family of compactly supported matrix-valued RBFs which also obey the properties suitable for learning (results generalized by Fuselier [7] for a construction of curl-free matrix-valued RBFs). These developments were used by Baudisch *et al.* [1] for interpolating divergence-free fluid simulation data for finite volume methods. However, their work was designed for face-averaged "clean" data located at the faces of computational cells from finite volume simulations.

Contributions. This work proposes a framework for vector field reconstruction in *arbitrary dimension* from an *unstructured*, *sparse* and, possibly, *noisy* sampling. Moreover, our approach allows us to *guarantee* that the reconstructed field is either free of *divergence* or of *curl*, the two most ubiquitous differential invariants of the vector fields encountered in classical physics (even though no other method proposed so far has been able to ensure them in such a hard setting). Our method is built upon recent developments on the *statistical learning theory of vector-valued functions* (reviewed at Section 2) and on the *approximation theory of matrix-valued radial basis functions* (whose constructions are presented at Section 3), none of which, to the best of our knowledge, has been exploited for computer graphics or vision applications so far. As proof of concept, we present some experiments performed with *synthetic*, *simulation* and *measurement* data sets at Section 4, where it is presented an encouraging result on learning a vector field *without prior invariants* and its decomposition into a sum of *divergence*- and *curl-free* parts, *with no additional cost*. Section 5 concludes this work providing directions for further research.

2. Statistical learning of vector fields

In this section, we provide a brief review of the foundations of statistical learning [6, 16] and the recent results in extending it to embrace vector-valued functions [12].

Our interest resides in learning a vector field $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ from unstructured, sparse and, possibly, noisy samples $\left(\mathbf{x}^i, \mathbf{y}^i\right)_{i=1}^N$ drawn from an *unknown* probability distribution $\mathbb{P}$ on the product space $\mathbb{R}^n \times \mathbb{R}^n$. It is desired that such a vector field, which is taken from a fixed hypothesis space $\mathcal{F}$, minimizes the expected error when evaluated on unseen data, i.e. given a *loss function* $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\mathbf{f}$ is chosen in order to minimize the *expected risk*, $\mathcal{R} : \mathcal{F} \to \mathbb{R}$,

$$\mathcal{R}[\mathbf{f}] := \mathbb{E}\left[L(\mathbf{y}, \mathbf{f}(\mathbf{x}))\right] = \int_{\mathbb{R}^n \times \mathbb{R}^n} L(\mathbf{y}, \mathbf{f}(\mathbf{x})) d\,\mathbb{P}. \qquad (1)$$

Unfortunately, since $\mathbb{P}$ is unknown, we need to take advantage of its available samples $\left(\mathbf{x}^i, \mathbf{y}^i\right)_{i=1}^N$ to be able to define a similar problem which is solvable and whose solution approximates a minimizer of (1). A natural approach is to approximate the above integral (which defines the expectation according to $\mathbb{P}$) by an empirical mean. This approach searches a minimizer of the *empirical risk* functional, $\mathcal{R}_e$,

$$\mathcal{R}_e[\mathbf{f}] := \frac{1}{N}\sum_{i=1}^N L(\mathbf{y}^i, \mathbf{f}(\mathbf{x}^i)). \tag{2}$$

However, even this new problem has its shortcomings, since the resulting minimization may admit infinitely many solutions which interpolate the samples having poor generalization performance, i.e. altought they minimize $\mathcal{R}_e$ they may be far from minimizing $\mathcal{R}$. To cope with these shortcomings, the empirical risk functional is augmented with a term which ensures uniqueness and regularity in the minimization problem. The usual practice takes $\mathcal{F}$ as a normed linear space of (nonlinear) vector fields and an increasing function $h : \mathbb{R} \rightarrow \mathbb{R}$ to define the *regularized risk*, $\mathcal{R}_r$,

$$\mathcal{R}_r[\mathbf{f}] := \frac{1}{N}\sum_{i=1}^N L(\mathbf{y}^i, \mathbf{f}(\mathbf{x}^i)) + h(\|\mathbf{f}\|), \tag{3}$$

which $\mathbf{f}$ is chosen to minimize. Notice that this approach subsumes many well established statistical techniques (e.g. *ridge regression*, where $L(\mathbf{y}^i, \mathbf{f}(\mathbf{x}^i)) = \|\mathbf{y}^i - \mathbf{f}(\mathbf{x}^i)\|^2$ and $h(x) = \frac{\lambda}{N}x^2$).

To properly pose our learning problem, we still need to define a hypothesis space $\mathcal{F}$ in which we seek a minimizer of the regularized risk functional (after all, *"learning does not take place in a vacuum"* [16]). The current practice (and theory) considers hypothesis spaces whose scalar-valued functions are arbitrarily well approximated by sums of the type $\sum_{j=1}^{M} \alpha_j k(\cdot, \mathbf{z}^j)$, in which $k : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ has certain properties (symmetry and (strict) positive definiteness) that allow to define an inner product in $\mathcal{F}$ [8].

Such spaces are important both in the theory and practice of scalar-valued ($n = 1$) learning because, under rather practical assumptions, it can be shown that,

Theorem 1 (Representer Theorem [8])**.** *If $f \in \mathcal{F}$ minimizes $\mathcal{R}_r$ then, for some $\alpha \in \mathbb{R}^N$, $f = \sum_{i=1}^{N} \alpha_i k(\cdot, \mathbf{x}^i)$.*

This means the minimization of $\mathcal{R}_r$ in $\mathcal{F}$ (usually infinite dimensional) can be restricted to the finite dimensional subspace generated by the functions $\left(k(\cdot, \mathbf{x}^i)\right)_{i=1}^N$. Such a remarkable result, with duality in nonlinear optimization [2], has allowed to design many state-of-the-art algorithms for learning scalar functions (e.g. *support vector regression*).

Recently, Micchelli and Pontil [12] generalized the basic theory developed in the scalar-valued case for learning vector-valued functions. Their work naturally subsumes the above construction of hypothesis spaces and shows that $\mathcal{F}$ should be made of vector-valued functions approximated by sums of the form $\sum_{j=1}^{M} K(\cdot, \mathbf{z}^j)\boldsymbol{\alpha}^j$, where each $\boldsymbol{\alpha}^j \in \mathbb{R}^n$ is an n-vector and the kernel $K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n\times n}$ is matrix-valued and has analogous properties to those in the scalar-valued case, to which they reduce when $n = 1$.

In [12], it is proved a version of the representer theorem for spaces of vector-valued functions induced by matrix-valued kernels which generalizes the known scalar result, in fact they prove this result in a much more general setting, altought we state a specialized version for simplicity.

Theorem 2 (Representer Theorem [12])**.**
If $\mathbf{f} \in \mathcal{F}$ minimizes $\mathcal{R}_r$ then, for some $\boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^N \in \mathbb{R}^n$,

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^N K(\mathbf{x}, \mathbf{x}^i)\boldsymbol{\alpha}^i, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

This result allowed them to design generalized support vector regression algorithms suitable for vector-valued learning. The SVR variants presented in [12] can be introduced by taking $h(x) = \lambda x^2$ and

- $L(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j|_\epsilon$, where $|\cdot|_\epsilon$ is the known ϵ-insensitive loss function $|\cdot|_\epsilon := \max(0, |\cdot| - \epsilon)$. This loss results in the following dual problem

$$\max\left\{-\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N (\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{*i})^T K(\mathbf{x}^i, \mathbf{x}^j)(\boldsymbol{\alpha}^j - \boldsymbol{\alpha}^{*j}) + \sum_{i=1}^N {\mathbf{y}^i}^T(\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{*i}) - \epsilon\sum_{i=1}^N \mathbf{e}^T(\boldsymbol{\alpha}^i + \boldsymbol{\alpha}^{*i})\right\} \tag{4}$$

  where $\boldsymbol{\alpha}^i, \boldsymbol{\alpha}^{*i} \in \mathbb{R}^n$, $\mathbf{0} \leq \boldsymbol{\alpha}^i, \boldsymbol{\alpha}^{*i} \leq \frac{1}{\lambda n N}\mathbf{e}$ and $\mathbf{e} \in \mathbb{R}^n$ is the n-vector with elements equal to 1.
- $L(\mathbf{y}, \hat{\mathbf{y}}) = \max_{j=1,\dots,n} |y_j - \hat{y}_j|_\epsilon$. This choice of loss function results in the following dual problem

$$\max\left\{-\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N (\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{*i})^T K(\mathbf{x}^i, \mathbf{x}^j)(\boldsymbol{\alpha}^j - \boldsymbol{\alpha}^{*j}) + \sum_{i=1}^N {\mathbf{y}^i}^T(\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{*i}) - \epsilon\sum_{i=1}^N \mathbf{e}^T(\boldsymbol{\alpha}^i + \boldsymbol{\alpha}^{*i})\right\} \tag{5}$$

  where $\boldsymbol{\alpha}^i, \boldsymbol{\alpha}^{*i} \geq \mathbf{0}$ and $\mathbf{e}^T\boldsymbol{\alpha}^i, \mathbf{e}^T\boldsymbol{\alpha}^{*i} \leq \frac{1}{\lambda N}$.

Both choices of loss functions (and their corresponding duals) subsume the classical support vector regression algorithm when $n = 1$, whose $L(y, \hat{y}) = |y - \hat{y}|_\epsilon$ and dual

$$\max\left\{-\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(\mathbf{x}^i, \mathbf{x}^j) + \sum_{i=1}^N y_i(\alpha_i - \alpha_i^*) - \epsilon\sum_{i=1}^N (\alpha_i + \alpha_i^*)\right\} \tag{6}$$

where $\alpha_i, \alpha_i^* \in \left[0, \frac{1}{\lambda N}\right]$. Notice that the first variant (4) reduces to the independent SVR learning of each of the n component scalar fields of $\mathbf{f}$ when the matrix-valued kernel evaluates to a diagonal matrix (its dual turns separable).

After solving the dual problem, the primal solution $\mathbf{f}$ can be recovered from the dual maximizers $(\boldsymbol{\alpha}^i, \boldsymbol{\alpha}^{*i})_{i=1}^N$ by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{N} K(\mathbf{x}, \mathbf{x}^i)(\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{*i}), \tag{7}$$

giving an analytic expression for the learned vector-field.

Although the first SVR variant (4) can also be deduced by grouping the dual problems derived for learning each component scalar field separately and employing the so called *"kernel trick"* [8] to couple the dual objectives with general matrix-valued kernels (which was how we first deduced it before knowing the work of Micchelli and Pontil [12]). We have chosen this approach in deducing the problem because it subsumes naturally a great family of important algorithms whose vector-valued variants still remain to be tested. We hope this brief discussion may encourage other graphics researchers to experiment with this framework for learning vector-valued functions in their own applications, because to the best of our knowledge this has not been done in computer graphics so far.

After developing the basic theory and the algorithms to be employed in learning vector fields with invariants, we just need a construction of matrix-valued kernel functions suitable for our purposes and which fit in our theory. This task is accomplished in the next section borrowing results from the approximation theory of radial basis functions.

3. Matrix-valued radial basis functions

In order to construct hypothesis spaces for learning divergence- and curl-free vector fields, it is sufficient to design matrix-valued kernels with the properties required by our framework and whose columns define themselves divergence- or curl-free vector fields, since the learned field (7) is a finite linear combination of them.

In a nutshell, the framework in [12] requires that,

(i) $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})^T$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;

(ii) For all $N \in \mathbb{N}$, $\mathbf{x}^1, \ldots, \mathbf{x}^N, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^N \in \mathbb{R}^n$ such that $\mathbf{x}^i \neq \mathbf{x}^j$ for $i \neq j$ and $\boldsymbol{\alpha}^k \neq \mathbf{0}$ for some k, it holds

$$\sum_{i=1}^{N} \sum_{j=1}^{N} {\boldsymbol{\alpha}^i}^T K(\mathbf{x}^i, \mathbf{x}^j) \boldsymbol{\alpha}^j > 0 \tag{8}$$

Notice that these properties subsume the scalar case in [8].

To design matrix-valued kernels whose columns are divergence- or curl-free vector fields, we adopt a construction studied in the theory of radial basis functions (RBFs) [14, 11, 7]. A *matrix-valued RBF*, $\Phi : \mathbb{R}^n \to \mathbb{R}^{n \times n}$, can be constructed from a scalar RBF, $\phi(\mathbf{x}) = \varphi(\|\mathbf{x}\|)$, by applying to ϕ a linear differential operator $\mathcal{L}$, i.e. $\Phi(\mathbf{x}) := (\mathcal{L}\phi)(\mathbf{x})$. The basic example of such an operator is the Hessian H, defined by $(\mathrm{H}\phi)_{ij} := \frac{\partial^2 \phi}{\partial x_i \partial x_j}$. This simple example will be essential in designing matrix-valued kernels for our approach to learn vector fields with those specific differential invariants.

3.1. Kernels for divergence-free vector fields

In [14], it was introduced a construction of matrix-valued RBFs in which the vector fields defined by the columns are divergence-free. These matrix-valued RBFs can be constructed by applying the operator $\mathcal{L} = \mathrm{H} - \mathrm{tr}(\mathrm{H}) \cdot \mathrm{I}$ to a scalar-valued RBF ϕ (it can be verified that the resulting matrix-valued function has divergence-free columns). Then,

$$\Phi_{df}(\mathbf{x}) = (\mathrm{H}\phi)(\mathbf{x}) - \mathrm{tr}\big\{(\mathrm{H}\phi)(\mathbf{x})\big\} \cdot \mathrm{I}, \tag{9}$$

and, as proved in [14], the (translation-invariant) matrix-valued kernel given by $K_{df}(\mathbf{x}, \mathbf{y}) = \Phi_{df}(\mathbf{x} - \mathbf{y})$ has the properties (i) and (ii) above for a popular class of scalar RBFs which includes the gaussian $e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}}$ and the inverse multiquadrics $\left(\sqrt{\|\mathbf{x}\|^2 + c^2}\right)^{-1}$.

For our results, we implemented the kernel derived from the gaussian radial basis function $\phi(\mathbf{x}) = \exp\left\{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}\right\}$,

$$K_{df}(\mathbf{x}, \mathbf{y}) = \frac{1}{\sigma^2} e^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}} \left[\left(\frac{\mathbf{x}-\mathbf{y}}{\sigma}\right)\left(\frac{\mathbf{x}-\mathbf{y}}{\sigma}\right)^T + \left((n-1) - \frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma^2}\right) \cdot \mathrm{I} \right]$$

3.2. Kernels for curl-free vector fields

Fuselier [7] introduced a construction of matrix-valued RBFs whose columns are curl-free vector fields (gradients of some scalar function) and proved results which ensure properties (i) and (ii) above for the (translation-invariant) matrix-valued kernel given by $K_{cf}(\mathbf{x}, \mathbf{y}) = \Phi_{cf}(\mathbf{x} - \mathbf{y})$, where $\mathcal{L} = -\mathrm{H}$, $\Phi_{cf}(\mathbf{x}) = -(\mathrm{H}\phi)(\mathbf{x})$ and ϕ belongs to the same class as in the divergence-free construction above.

In this case, the matrix-valued curl-free kernel induced by the gaussian function, used for our experiments, is

$$K_{cf}(\mathbf{x}, \mathbf{y}) = \frac{1}{\sigma^2} e^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}} \left[\mathrm{I} - \left(\frac{\mathbf{x}-\mathbf{y}}{\sigma}\right)\left(\frac{\mathbf{x}-\mathbf{y}}{\sigma}\right)^T \right]$$

3.3. A class of compactly-supported kernels

The results on globally supported divergence-free matrix-valued RBFs proved in [14] were generalized in [11] to an important (and widely adopted) class of compactly supported RBFs, the *Wendland's functions* [19]. In possession of such results, we are able to design compactly supported matrix-valued kernels with which we can learn vector fields with invariants (since, in [7], it was proved that the above construction of curl-free matrix-valued RBFs also induces kernels with properties (i) and (ii) for Wendland's functions).

These results have important practical consequences, because, using compactly supported kernels, the dual optimization problems may have a very sparse quadratic term and the evaluation of the sum (7) may become extremely cheaper when N is large. Although we didn't perform any experiments with those kernels in this work, we believe they provide an interesting avenue for further investigation.

4. Results

As proof of concept, we implemented the dual optimization problem (4) and the divergence-free and curl-free matrix-valued kernels induced by the gaussian RBF. We chose that extension of ϵ-SVR because it specializes to traditional component-wise kernel-based ϵ-SVR when the matrix-valued kernel is diagonal, in our case, $e^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}} \cdot \mathrm{I}$.

We designed a few representative experiments in which we reconstruct vector fields from a variety of sources: *synthetic*, *simulation* and *measurement* datasets. These inputs were preprocessed to fit in $[-1,1] \times [-1,1]$, in which case the learning parameters were hand-tuned and chosen as $\epsilon = 10^{-3}$, $\lambda = \frac{1}{nN}$ and $\sigma = 0.3$. The synthetic datasets were constructed by taking a scalar function (a mixture of four gaussians with coefficients alternating in sign) and taking its gradient ∇ as the reference curl-free vector field, its grad perp $\nabla^{\perp}$ as the reference divergence-free field and their sum as the reference (general) vector field (Figure 2).

Implementation note. Our results and visualizations were implemented with Mathworks' MATLAB™ `quadprog` routine encoutered in the *Optimization Toobox* and the *Matlab Vector Field Visualization toolkit* [18]. The simplicity of the construction of matrix-valued RBFs presented at Section 3 allowed an implementation in which the input dimension was abstracted and the learning code works for arbitrary dimensions. In fact, the only module of our code which is dimension-dependent is the visualization code. This fact highlights the elegance of the framework we propose.

In the following, we provide **relative** error *percentuals* for test sets about 16 times larger then the training datasets.

4.1. Learning divergence-free vector fields

The following experiments use the matrix-valued kernel from subsection 3.1 for divergence-free vector fields and the gaussian diagonal kernel for the component-wise ϵ-SVR.

Synthetic data. Our first experiment involves learning the synthetic divergence-free field described before from a sampling taken sparsely and uniformly distributed in $[-1,1]^2$. The training set consists of 100 points-vectors pairs depicted in Figure 3 along with our everywhere divergence-free reconstruction, a plot of the divergence of a field learned by traditional ϵ-SVR applied component-wise and that latter vector field.

Comparing these results with the reference div-free field in Figure 2, we observe that the essential features of the flow are very well retained and the main differences occur on very low sampled areas on the corners of that image (in which the magnitude of these fields is small). Quantitatively, both methods obtained small relative errors, but our method ensures the divergence-free invariant of the learned vector field besides achieving more than a fifth (0.92 percent) of the relative error obtained by component-wise learning using a scalar ϵ-SVR (5.18 percent).

Fluid simulation data. The second experiment consists of data drawn from a computer simulation of a confined incompressible fluid (whose dynamics is characterized by a divergence-free constraint on the velocity field). The task is to approximate the velocity profile from a few unstructured samples respecting the incompressibility invariant of the underlying physics.

Although the original dataset has more than four thousand points-vectors pairs, we took only 100 training samples uniformly *on the available data* illustrated in Figure 4. Our method reconstructed a divergence-free field which retained the features observable in the training samples with a relative error of 25.8 percent. The component-wise learning using scalar ϵ-SVR obtained a just slightly smaller relative error (23.0 percent) with the price of large oscillations in the divergence field of its attained reconstruction.

Measurement data. Our last result in learning a divergence-free field is the hardest one. This dataset consists of almost sixteen thousand points-vectors pairs measured from a real air flow at low speeds in a regime which is essentially incompressible. These measurements were performed using the method of *particle image velocimetry* (PIV) [17], which generates large and noisy datasets highly contaminated by outliers, posing a hard estimation task even without the added requirement of guaranteeing the mass conservation law characterized by the null divergence condition.

Figure 5 depicts the 100 training samples uniformly distributed *on the available data* and the reconstruction results. Our method captured the percetible trends in the flow field samples as well as the localized vortices tipical of slightly turbulent air motion, differently of that reconstruction obtained by component-wise ϵ-SVR learning, which introduced many sources and sinks at low sampled regions neighboring what should be localized vortices. This resulted in large variations on the divergence of that field learned by classical kernel-based support vector regression.

The relative error attained by both methods was rather large. Our divergence-free reconstruction had a 69.34 percent relative error while the component-wise ϵ-SVR method obtained 83.23. We believe this large error may be due to both a small percentage of samples used for training (less than 0.7 percent of the total data was used for training) and very noisy measurements contaminated by outliers. Even with these quantitative results, we believe that the divergence-free reconstruction retained very well both the bulk flow and the small scale motion percetible in the training dataset.

4.2. Learning a curl-free vector field

Figure 6 depicts the results we obtained on learning a curl-free vector field from 100 sparse samples located at random sites uniformly distributed on $[-1,1]^2$ whose values were taken from the synthetic curl-free field depicted in Figure 2. In Figure 6, the result obtained by training the classical ϵ-SVR on each component independently is also illustrated along a plot of its curl field (which is identically zero in the reference vector field). The relative errors of both methods are small, but the relative error achieved by our method (0.69 percent) is almost a tenth of that obtained by component-wise learning (5.19 percent), besides the guarantee we provide in obeying the curl-free constraint.

4.3. Learning a vector field decomposition

Using the fact that non-negative linear combinations of matrix-valued kernels which obey properties (i) and (ii) from Section 3 also have those properties, we can design a family of kernels for general vector fields, not just those with differential invariants. One such family just takes a convex combination of the two matrix-valued kernels constructed in Section 3 and is parameterized by a $\gamma \in [0,1]$,

$$K_\gamma(\mathbf{x},\mathbf{y}) = (1-\gamma)K_{df}(\mathbf{x},\mathbf{y}) + \gamma K_{cf}(\mathbf{x},\mathbf{y}). \qquad (10)$$

We have made some simple experiments with K_γ $\left(\text{for } \gamma = \frac{1}{2}\right)$ in which a training set with 128 samples was drawn from the reference vector field in Figure 2 located at random places distributed uniformly at $[-1,1]^2$. The coefficients were them used to reconstruct two fields using as kernels $(1-\gamma)K_{df}$ and γK_{cf}, the results of this experiment are depicted in Figure 1, the relative reconstruction error achieved was less than 1.8 percent even for a field without prior invariants.

As can be observed by comparing figures 1 and 2, the features of the flows (reference and its parts) were correctly learned without complicating the optimization problem or iterating, it was just required to use the kernel K_γ when building the quadratic term in the dual problem (4) and separating this kernel's components when evaluating (7).

Since this kind of vector field decomposition has many applications in fluid simulation and analysis of vector fields for scientific computing and visualization, learning divergence-/curl-free decompositions of vector fields from sparse unstructured noisy samples provides a very promising avenue for further research.

5. Conclusions and future works

In this work, we introduce a novel framework to design methods for reconstructing vector fields with differential invariants commonly required in applications. Our approach builds upon recent developments in the statistical learning theory of vector-valued functions and in the approximation theory of matrix-valued radial basis functions to allow vector field reconstruction in *arbitrary dimension* from *unstructured*, *sparse* and *noisy* samplings in an elegant and theoretically sound way *guaranteeing common constraints imposed by the underlying physics*.

The theoretical frameworks we use subsume many traditional statistical techniques (e.g. *ridge regression*) and some of those considered state-of-the-art (e.g. kernel-based *support vector regression*) by a separation of measures for approximation quality (loss function) and approximant complexity (induced by the regularization term). As proof of concept, we implemented a generalization of the kernel-based ϵ-SVR algorithm of which the component-wise learning of the vector field is a natural special case. We also presented an encouraging result for reconstructing vector fields without prior invariants in a manner which, with practically no added cost, it is possible to decompose the learned vector field in a sum of divergence- and curl-free parts, still retaining the elegance of the method in arbitrary dimensions.

Future works. Currently, we are working on a modification of the *sequential minimal optimization* (SMO) [15] designed for efficiently training scalar support vector machines to our problem (in which the dual variables are vector-valued and the kernels are matrix-valued). Another aspect of training we are interested in regards parameter choice (model selection) by cross-validation methods, since it is a hard (and annoying) task to hand-tune each of them for each different dataset (e.g. ϵ, λ and σ).

This work is one the first in computer graphics (if not the very first) to exploit those recent results in learning vector-valued functions and/or representing divergence-/curl-free vector fields using suitably designed matrix-valued radial basis functions. We believe that these developments have many potential applications both in graphics and vision. To mention a few, animation of incompressible fluid flows (and compression of the resulting datasets), design of vector fields for animation or rendering purposes, flow analysis in scientific visualization, estimation of mass conserving deformation fields in 2D or 3D medical images and related optical flow computations.

Acknowledgements

We would like to thank the Matmídia Laboratory of the Department of Mathematics at PUC-Rio and the Vision and Graphics Laboratory (Visgraf) at IMPA for kindly providing computational resources and datasets. This research was partially supported by Brazil's National Counsel of Technological and Scientific Development (CNPq).

References

[1] J. Baudisch, L. Bonaventura, A. Iske, and E. Miglio. Matrix-valued radial basis functions for local vector field reconstruction in computational fluid dynamic models. Technical report, Politecnico di Milano, 2006. MOX Report 75.

[2] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999. Second Edition.

[3] M. Brudnak. Vector-valued support vector regression. *International Joint Conference on Neural Networks, 2006. IJCNN'06.*, pages 1562–1569, 2006.

[4] R. Castro, M. Lage, F. Petronetto, G. Tavares, T. Lewiner, and H. Lopes. Learning vector fields by kernels. Technical report, Pontifical Catholic University of Rio de Janeiro, 2007.

[5] J. Colliez, F. Dufrenois, and D. Hamad. Optic flow estimation by support vector regression. *Engineering Applications of Artificial Intelligence*, 19(7):761–768, 2006.

[6] T. Evgeniou, M. Pontil, and T. Poggio. Statistical learning theory: A primer. *International Journal of Computer Vision*, 38(1):9–13, 2000.

[7] E. Fuselier. *Refined error estimates for matrix-valued radial basis functions*. PhD thesis, Texas A&M University, 2006.

[8] T. Hofmann, B. Schölkopf, and A. J. Smola. Kernel methods in machine learning. *Annals of Statistics*, (in press).

[9] Y. Kuroe, M. Mitsui, H. Kawakami, and T. Mori. A learning method for vector field approximation by neural networks. *IEEE International Joint Conference on Neural Networks, 2006. IJCNN'98.*, 3:2300–2305, 1998.

[10] M. Lage, F. Petronetto, A. Paiva, H. Lopes, T. Lewiner, and G. Tavares. Vector field reconstruction from sparse samples with applications. *19th Brazilian Symposium on Computer Graphics and Image Processing, 2006. SIBGRAPI'06.*, pages 297–306, 2006.

[11] S. Lowitzsch. *Approximation and interpolation employing divergence-free radial basis functions with applications*. PhD thesis, Texas A&M University, 2002.

[12] C. A. Micchelli and M. Pontil. On learning vector-valued functions. *Neural Computation*, 17(1):177–204, 2005.

[13] F. A. Mussa-Ivaldi. From basis functions to basis fields: vector field approximation from sparse data. *Biological Cybernetics*, 67(6):479–489, 1992.

[14] F. J. Narcowich and J. D. Ward. Generalized Hermite interpolation via matrix-valued conditionally positive definite functions. *Math. Comp.*, 63(208):661–687, 1994.

[15] J. C. Platt. Using analytic qp and sparseness to speed training of support vector machines. In *Advances in Neural Information Processing Systems 11*, pages 557–563. 1999.

[16] T. Poggio and S. Smale. The mathematics of learning: dealing with data. *Notices of the American Mathematical Society*, 50(5):537–544, 2003.

[17] M. Raffel, C. Willert, S. Wereley, and J. Kompenhans. *Particle Image Velocimetry: A Practical Guide*. Springer, 2007.

[18] N. B. Shamlo. Matlab toolbox for high resolution vector field visualization with application in improving the understanding of crack propagation mechanisms. Master's thesis, San Diego State University, 2005.

[19] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(4):389–396, 1995.

[20] J. Zhong, J. Weng, and T. Huang. Robust and physically-constrained interpolation of fluid flow fields. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP'92.*, 3:185–188, 1992.

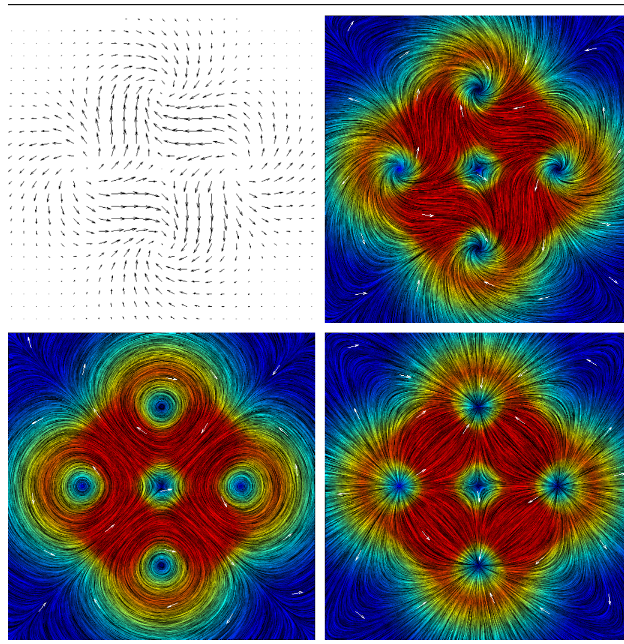

Figure 2. Reference synthetic data. Top: two visualizations of the reference vector field; Bottom: its divergence- and curl-free parts.

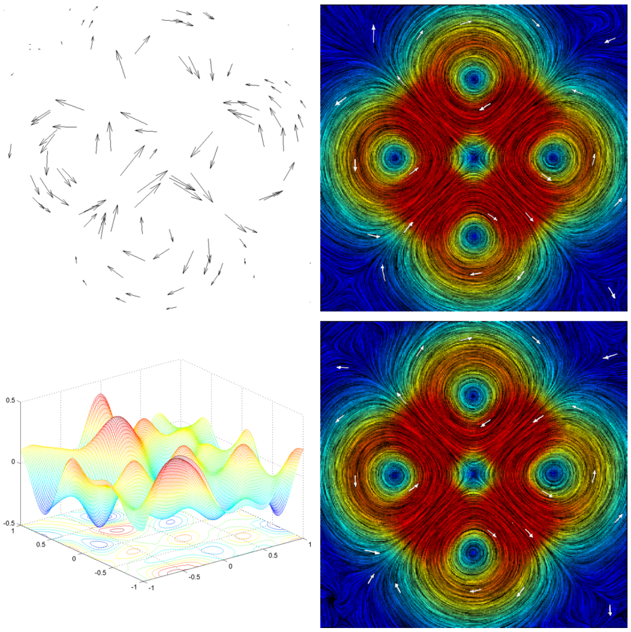

Figure 3. Divergence-free synthetic data. Top: samples and our div-free solution; Bottom: div of scalar SVR and associated field.

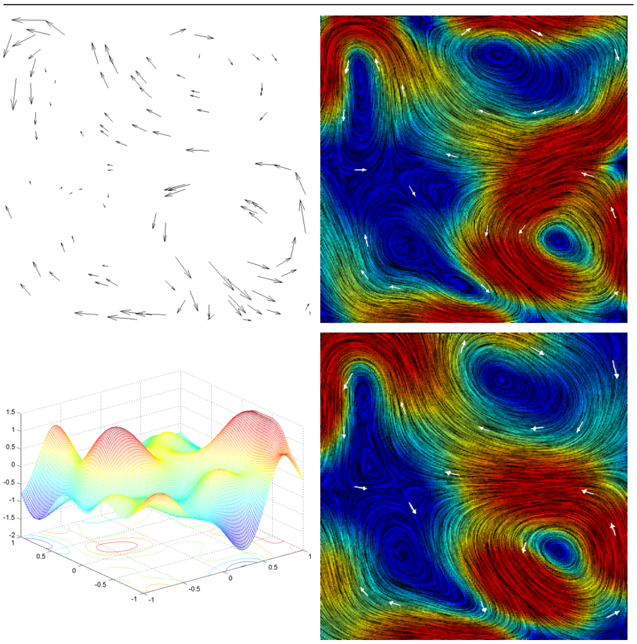

Figure 4. Fluid simulation data. Top: samples and our div-free solution; Bottom: div of scalar SVR and associated vector field.

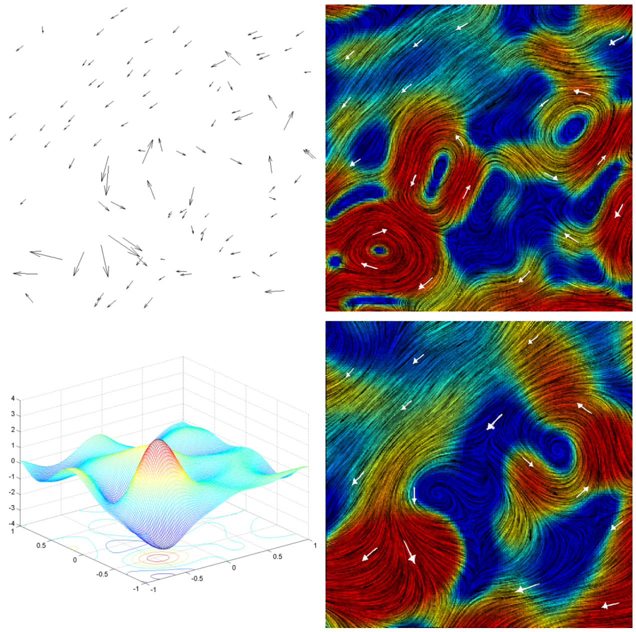

Figure 5. Measurement PIV data. Top: samples and our div-free solution; Bottom: div of scalar SVR and associated vector field.

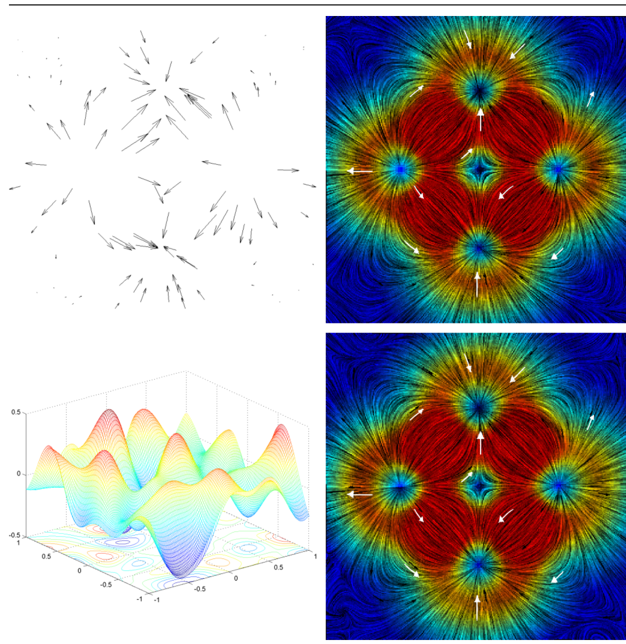

Figure 6. Curl-free synthetic data. Top: samples and our curl-free solution; Bottom: curl of scalar SVR and associated vector field.