

Translation-invariant interpolation of parametric dictionaries

Frédéric Champagnat, Cédric Herzet

▶ **To cite this version:**

Frédéric Champagnat, Cédric Herzet. Translation-invariant interpolation of parametric dictionaries. iTwist 2020 - International Traveling Workshop on Interactions between low-complexity data models and Sensing Techniques, Dec 2020, Nantes, France. pp.1-3. hal-03070414

HAL Id: hal-03070414
https://inria.hal.science/hal-03070414

Submitted on 15 Dec 2020

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.

Translation-invariant interpolation of parametric dictionaries

Frédéric Champagnat
ONERA, The French Aerospace Lab
Palaiseau, France
frederic.champagnat@onera.fr

Cédric Herzet
INRIA Rennes-Bretagne Atlantique
Rennes, France
cedric.herzet@inria.fr

***Abstract*—In this communication, we address the problem of approximating the atoms of a parametric dictionary $\mathcal{A} = \{\boldsymbol{a}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$, commonly encountered in the context of sparse representations in "continuous" dictionaries. We focus on the case of translation-invariant dictionaries, where the inner product between $\boldsymbol{a}(\boldsymbol{\theta})$ and $\boldsymbol{a}(\boldsymbol{\theta}')$ only depends on the difference $\boldsymbol{\theta} - \boldsymbol{\theta}'$. We investigate the following general question: is there some low-rank approximation of $\mathcal{A}$ which interpolates a subset of atoms $\{\boldsymbol{a}(\boldsymbol{\theta}_j)\}_{j=1}^J$ while preserving the translation-invariant nature of the original dictionary? We derive necessary and sufficient conditions characterizing the existence of such an "interpolating" and "translation-invariant" low-rank approximation. Moreover, we provide closed-form expressions of such a dictionary when it exists. We illustrate the applicability of our results in the case of a two-dimensional isotropic Gaussian dictionary. We show that, in this particular setup, the proposed approximation framework outperforms standard Taylor approximation.**

***Index Terms*—Sparse representations, continuous dictionaries, translation invariance, interpolating approximations.**

I. Introduction

Sparse representations aim at representing a signal of interest as the linear combination of a few elements of a dictionary $\mathcal{A}$. Recently, this problem has been reformulated in a "continuous" setting, where the elements of $\mathcal{A}$ are continuously indexed by some parameter $\boldsymbol{\theta}$:

$$\mathcal{A} = \{\boldsymbol{a}(\boldsymbol{\theta}) \in \mathcal{H} : \boldsymbol{\theta} \in \Theta\} \tag{1}$$

where Θ is a square interval of $\mathbb{R}^d$, $\mathcal{H}$ is a Hilbert space over the real field $\mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$, and $\boldsymbol{a}(\cdot)$ some continuous function from Θ to $\mathcal{H}$, see *e.g.*, [1]–[3]. We adopt the common hypothesis that $\|\boldsymbol{a}(\boldsymbol{\theta})\| = 1 \ \forall \boldsymbol{\theta} \in \Theta$. We focus on families of translation invariant parametric dictionaries: we consider the case where

$$\big\langle \boldsymbol{a}(\boldsymbol{\theta}), \boldsymbol{a}(\boldsymbol{\theta}') \big\rangle = \big\langle \boldsymbol{a}(\boldsymbol{\theta} - \boldsymbol{\tau}), \boldsymbol{a}(\boldsymbol{\theta}' - \boldsymbol{\tau}) \big\rangle, \tag{2}$$

that is the inner product between two atoms of the dictionary only depends on $\boldsymbol{\theta} - \boldsymbol{\theta}'$. This setup is ubiquitous in many physical, chemical or biological problems where the observed signal is the linear combination of shifted copies of the system's impulse response, see *e.g.*, [3], [4].

The authors thank the project "ANR BECOSE" (ANR-15-CE23-0021) for its support.

Continuous dictionaries contain an infinite uncountable number of elements and induce therefore new difficulties. Trivial operations in the discrete setting may become challenging in the continuous framework. For instance, consider the well-known "atom selection" problem [5], [6]:

$$\text{Find } \boldsymbol{\theta}^\star = \arg\max_{\boldsymbol{\theta} \in \Theta} \langle \boldsymbol{a}(\boldsymbol{\theta}), \boldsymbol{r} \rangle \quad \text{for some } \boldsymbol{r} \in \mathcal{H}, \tag{3}$$

which only involves the evaluation of a finite number of inner products in the discrete setting but can turn out to be a difficult optimization task in the continuous framework.

In order to circumvent this problem, several contributions of the literature have proposed to tackle the infinite-dimensional nature of continuous dictionaries by resorting to "low-rank" approximations of $\boldsymbol{a}(\boldsymbol{\theta})$. More formally, the idea consists in approximating the elements of $\mathcal{A}$ as linear combinations of a few vectors $\{\boldsymbol{v}_k\}_{k=1}^K$:

$$\hat{\boldsymbol{a}}(\boldsymbol{\theta}) \triangleq \sum_{k=1}^K \boldsymbol{v}_k c_k(\boldsymbol{\theta}) \tag{4}$$

for some functions $\{c_k : \Theta \to \mathbb{R}\}_{k=1}^K$. This approach was for example used in [7] and [8] to transform "continuous" sparse-representation problems into approximate (but tractable) finite-dimensional ones. The quality of the approximation obtained with this strategy obviously depends on the choice of the vectors $\{\boldsymbol{v}_k\}_{k=1}^K$ and functions $\{c_k\}_{k=1}^K$. Several options were considered in [7]–[9]. In [7], the authors introduced the "Taylor" and "polar" approximations: the former is based on a Taylor decomposition of $\boldsymbol{a}(\boldsymbol{\theta})$; the latter is constructed so that $\hat{\boldsymbol{a}}(\boldsymbol{\theta})$ has a unit norm $\forall \boldsymbol{\theta} \in \Theta$ and interpolates $\boldsymbol{a}(\boldsymbol{\theta})$ for some $\{\boldsymbol{\theta}_j\}_{j=1}^3$. In [8], the authors suggested to use a singular-value decomposition of $\boldsymbol{a}(\boldsymbol{\theta})$ to identify the approximation subspace minimizing the projection error in a $\|\cdot\|$-sense. In [9], the authors pointed out some connection between the approximations considered in [7] and [8].

The present communication is mainly of theoretical nature. We address the following general question: is there some low-rank approximation of the form (4) which: *(i)* interpolates the original dictionary over some subset of parameters $\{\boldsymbol{\theta}_j\}_{j=1}^J$; *(ii)* preserves the translation invariance of the original dictionary? The formal statement of this question is given below, all the formal derivations are available in [10].

II. Problem statement and main result

We focus on low-rank approximations of the form (4) obeying the two following properties:

- *Interpolation*:

$$\hat{\boldsymbol{a}}(\boldsymbol{\theta}) = \boldsymbol{a}(\boldsymbol{\theta}) \quad \text{for } \boldsymbol{\theta} \in \{\boldsymbol{\theta}_j\}_{j=1}^{J}. \tag{5}$$

- *Translation invariance*:

$$\left\langle \hat{\boldsymbol{a}}(\boldsymbol{\theta}), \hat{\boldsymbol{a}}(\boldsymbol{\theta}') \right\rangle = \lambda_0 + \sum_{\ell=1}^{L} \lambda_\ell \cos(\boldsymbol{\omega}_\ell^{\mathrm{T}}(\boldsymbol{\theta} - \boldsymbol{\theta}')) \tag{6}$$

for some $\{\lambda_\ell\}_{\ell=0}^{L}$ and $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$.

Property (5) imposes that the approximation $\hat{\boldsymbol{a}}(\boldsymbol{\theta})$ perfectly interpolates the original atom $\boldsymbol{a}(\boldsymbol{\theta})$ for some values of the parameters $\{\boldsymbol{\theta}_j\}_{j=1}^{J}$. Property (6) enforces the inner product of the approximated atoms to obey a "raised-cosine" kernel. In particular, we note that satisfying (6) ensures the translation-invariance of $\hat{\boldsymbol{a}}(\boldsymbol{\theta})$.

In the rest of this section, we answer the two following questions: 1) is there some low-rank approximation $\hat{\mathcal{A}} = \{\hat{\boldsymbol{a}}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ verifying properties (5) and (6)? 2) if such an approximation exists, how to build it?

Before presenting our main result, we make two important remarks, which will allow us to give a more formal statement of the two above questions.

First, in order to simplify our exposition, we assume that $\{\lambda_\ell\}_{\ell=0}^{L}$ and $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$ are such that the right-hand side of (6) defines a semidefinite kernel. This assumption makes sense since the kernel induced by any family of parametric atoms must necessarily be positive semidefinite. Second, consistency of the different decompositions imply constraints on parameters (J, K, L), more precisely

$$\begin{aligned} J &= K \\ L &= \lfloor K/2 \rfloor \end{aligned}$$

where $\lfloor \cdot \rfloor$ returns the greatest integer less than or equal to its input argument. We also assume that the interpolated vectors $\{\boldsymbol{a}(\boldsymbol{\theta}_j)\}_{j=1}^{J}$ are linearly independent. The main question addressed herefafter thus takes the following form:

Main Question. *Is there a rank-K approximation* (4) *that interpolates a family of K linearly independent atoms $\{\boldsymbol{a}(\boldsymbol{\theta}_j)\}_{j=1}^{K}$ and verifies* (6) *for some $\{\lambda_\ell\}_{\ell=0}^{\lfloor K/2 \rfloor}$ and $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{\lfloor K/2 \rfloor}$? If so, how to build such an approximation?*

The answer to this question is provided in Theorem 1 below. We present necessary and sufficient conditions that ensure the existence of the desired approximation. Moreover, when it exists, we provide the closed-form expressions of $\{\boldsymbol{v}_k\}_{k=1}^{K}$ and $\{c_k(\boldsymbol{\theta})\}_{k=1}^{K}$ defining the approximation.

Our main result relates the existence of a low-rank approximation verifying (5)-(6) to some constraints on the Gram matrix of the family of interpolated atoms $\{\boldsymbol{a}(\boldsymbol{\theta}_j)\}_{j=1}^{K}$. More specifically, letting

$$\mathbf{G} \triangleq [\langle \boldsymbol{a}(\boldsymbol{\theta}_i), \boldsymbol{a}(\boldsymbol{\theta}_j) \rangle]_{i,j} \in \mathbb{R}^{K \times K}, \tag{7}$$

the following result holds:

Theorem 1. *Let $K \in \mathbb{N}$. If there exists a family of atoms* (4) *satisfying* (5)-(6)*, then*

$$\mathbf{G}(i,j) = \lambda_0 + \sum_{\ell=1}^{L} \lambda_\ell \cos(\boldsymbol{\omega}_\ell^{\mathrm{T}}(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j)) \quad \forall i,j \tag{8}$$

for some $\{\lambda_\ell\}_{\ell=0}^{L}$, $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$.

Conversely, assume there exist $\{\lambda_\ell\}_{\ell=0}^{L}$, $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$ and $\{\boldsymbol{\theta}_j\}_{j=1}^{K}$ such that (8) *holds. Then, low-rank approximation* (4) *with*

$$c_k(\boldsymbol{\theta}) = \lambda_0 + \sum_{\ell=1}^{L} \lambda_\ell \cos(\boldsymbol{\omega}_\ell^{\mathrm{T}}(\boldsymbol{\theta} - \boldsymbol{\theta}_k)) \tag{9}$$

$$\boldsymbol{v}_k = \sum_{j=1}^{K} \boldsymbol{a}(\boldsymbol{\theta}_j)\, \mathbf{G}^{-1}(k,j), \tag{10}$$

verifies (6) *and interpolates $\{\boldsymbol{a}(\boldsymbol{\theta}_j)\}_{j=1}^{K}$.*

A proof of this result is available in [10]. Theorem 1 provides necessary and sufficient conditions for our "Main Question" to have a positive answer. Interestingly, we see that the existence of a low-rank decomposition verifying (5)-(6) is exclusively conditioned on the existence of some particular factorization of the Gram matrix $\mathbf{G}$ (see condition (8)). Hence, Theorem 1 transforms the question of the existence of an interpolating and translation-invariant low-rank approximation into an algebraic problem where one must find a set of parameters $\{\lambda_\ell\}_{\ell=0}^{L}$, $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$ verifying equation (8).

The existence of decomposition (8) will depend on the nature of the original dictionary $\mathcal{A}$ and the choice of the interpolation parameters $\{\boldsymbol{\theta}_j\}_{j=1}^{K}$. Providing a general answer to this algebraic problem is therefore a broad question which is out of the scope of this communication. Note however that in the case of interpolated atoms on a regular grid, the Gram matrix problem decomposition is linked to the well-documented Vandermonde decomposition problem [11]. The Vandermonde decomposition is not fully resolved in the multi-dimensional case but we provide a solution to this problem in the particular case of a separable isotropic dictionary $\mathcal{A}$.

To conclude this section, let us mention that the expressions of the vectors $\{\boldsymbol{v}_k\}_{k=1}^{K}$ and functions $\{c_k\}_{k=1}^{K}$ provided in Theorem 1 only involve simple operations and are thus easy to evaluate numerically. Hence, the construction a low-rank dictionary verifying (5)-(6) is straightforward once the set of paremeters $\{\lambda_\ell\}_{\ell=0}^{L}$ and $\{\boldsymbol{\omega}_\ell\}_{\ell=1}^{L}$ satisfying conditions (8) has been identified.

We will illustrate the applicability of our results in the case of a two-dimensional isotropic Gaussian dictionary. We show that, in this particular setup, the proposed approximation framework outperforms standard Taylor approximation.

REFERENCES

[1] Kristian Bredies and Hanna Katriina Pikkarainen, "Inverse problems in spaces of measures," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 19, no. 1, pp. 190–218, 2013.

[2] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Comm. Pure Appl. Math*, vol. 67, no. 6, pp. 906–956, June 2014.

[3] Quentin Denoyelle and Gabriel Peyré, *Theoretical and Numerical Analysis of Super-Resolution Without Grid*, Ph.D. thesis, Université Paris-Dauphine, 2018.

[4] Zai Yang, Jian Li, Petre Stoica, Lihua Xie, Rama Chellappa, and Sergios Theodoridis, *Chapter 11 - Sparse methods for direction-of-arrival estimation*, pp. 509–581, Academic Press, 2018.

[5] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Ann. Asilomar Conf. Signals, Systems, and Computers*, 1993.

[6] Martin Jaggi, "Revisiting frank-wolfe: Projection-free sparse convex optimization," in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*. 2013, ICML'13, pp. I–427–I–435, JMLR.org.

[7] Chaitanya Ekanadham, Daniel Tranchina, and Eero P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *Signal Processing, IEEE Transactions on*, vol. 59, no. 10, pp. 4735–4744, Oct. 2011.

[8] Karin C. Knudson, Jacob L. Yates, Alexander C. Huk, and Jonathan W. Pillow, "Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit," in *Advances in Neural Information Processing*, 2014.

[9] F. Champagnat and C. Herzet, "Atom selection in continuous dictionaries: Reconciling polar and svd approximations," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 5516–5520.

[10] F. Champagnat and C. Herzet, "Interpolating and translation-invariant approximations of parametric dictionaries," Tech. Rep., INRIA-Rennes Bretagne Atlantique, http://people.rennes.inria.fr/Cedric.Herzet, 2020.

[11] Zai Yang, Lihua Xie, and Petre Stoica, "Vandermonde decomposition of multilevel toeplitz matrices with application to multidimensional super-resolution," *IEEE Trans. Information Theory*, vol. 62, pp. 3685–3701, 2016.