



Existence, Stability And Scalability Of Orthogonal Convolutional Neural Networks

El Mehdi Achour, François Malgouyres, Franck Mamalet

► To cite this version:

El Mehdi Achour, François Malgouyres, Franck Mamalet. Existence, Stability And Scalability Of Orthogonal Convolutional Neural Networks. 2021. hal-03315801v1

HAL Id: hal-03315801

<https://hal.science/hal-03315801v1>

Preprint submitted on 9 Aug 2021 (v1), last revised 10 Jan 2023 (v3)

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.

Existence, Stability And Scalability Of Orthogonal Convolutional Neural Networks

El Mehdi Achour^{a*}, François Malgouyres^a, Franck Mamalet^b

^aInstitut de Mathématiques de Toulouse ; UMR 5219

Université de Toulouse ; CNRS

UPS IMT F-31062 Toulouse Cedex 9, France

^bInstitut de Recherche Technologique Saint Exupéry, Toulouse, France

*Corresponding author: El Mehdi Achour; El_mehdi.achour@math.univ-toulouse.fr

Abstract

Imposing orthogonal transformations between layers of a neural network has been considered for several years now. This facilitates their learning, by limiting the explosion/vanishing of the gradient; decorrelates the features; improves the robustness. In this framework, this paper studies theoretical properties of orthogonal convolutional layers.

More precisely, we establish necessary and sufficient conditions on the layer architecture guaranteeing the existence of an orthogonal convolutional transform. These conditions show that orthogonal convolutional transforms exist for almost all architectures used in practice.

Recently, a regularization term imposing the orthogonality of convolutional layers has been proposed. We make the link between this regularization term and orthogonality measures. In doing so, we show that this regularization strategy is stable with respect to numerical and optimization errors and remains accurate when the size of the signals/images is large. This holds for both row and column orthogonality.

Finally, we confirm these theoretical results with experiments, and also empirically study the landscape of the regularization term.

1 Introduction

1.1 On Orthogonal Convolutional Neural Networks

Orthogonality constraint have first been considered for fully connected neural networks [3]. For Convolutional Neural Networks (CNN) [22, 21, 46], the introduction of the orthogonality constraint is a way to improve the neural network in several regards. First, despite well established solutions [13, 17], the training of very deep convolutional networks remains difficult. This is in particular due to vanishing/exploding gradients problems [14, 5]. As a result, the expressive capacity of convolutional layers is not fully exploited [17]. This can lead to lower performances on machine learning tasks. Also, the absence of constraint on the convolutional layer often leads to irregular predictions that are prone to adversarial attacks [38, 28]. For these reasons, some authors have introduced Lipschitz [38, 30, 10, 40, 33] and orthogonality constraints to convolutional layers [43, 7, 16, 45, 25, 12, 29, 41, 39, 19, 24, 15, 19, 4]. Beside the above motivations for considering Lipschitz and orthogonality constraints, these constraints are commonly used : - in Generative Adversarial Networks (GAN) [27] and Wasserstein-GAN [2, 11]; - in Recurrent Neural Networks [3, 15].

This article focuses on theoretical properties of orthogonal convolutional networks. We will consider the architecture of a convolutional layer as characterized by (M, C, k, S) , where M is the number of output channels, C of input channels, convolution kernels are of size $k \times k$ and the stride parameter is S . Unless we specify otherwise, we consider convolutions with circular boundary conditions: For input channels of size $SN \times SN$, the M output channels are of size $N \times N$. We denote by $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$ the kernel tensor and by $\mathcal{K} \in \mathbb{R}^{MN^2 \times CS^2N^2}$ the matrix that applies the convolutional layer of architecture (M, C, k, S) to C vectorized channels of size $SN \times SN$.

We will first answer this important question: What is a necessary and sufficient condition on (M, C, k, S) and N such that there exists an orthogonal convolutional layer for this architecture?

Besides, we will rely on recently published papers [41, 29] which characterizes orthogonal convolutional layers as the zero level set of a particular function¹ that (in [41]) is called L_{orth} (see Sect. 1.3.2 for details). Formally, \mathcal{K} is orthogonal if and only if $L_{orth}(\mathbf{K}) = 0$. They use L_{orth} as a regularization term and obtain impressive performances on several machine learning tasks.

In the present paper, we investigate the following theoretical questions:

- In what sense \mathcal{K} still has good ‘approximate orthogonality properties’ when $L_{orth}(\mathbf{K})$ is small but non zero?
- Remarking that, for a given kernel tensor \mathbf{K} , the layer transform matrix \mathcal{K} depends on the parameter N defining the size of the channels: Does \mathcal{K} remains approximately orthogonal when N grows?
- Does the landscape of L_{orth} lend itself to global optimization?

We describe the related works in Section 1.2 and give the main elements of context in Section 1.3. The theorems constituting the main contribution of the article are in Section 2. Experiments illustrating the theorems as well as experiments on the landscape of L_{orth} are in Section 3. The code will be made available in *ANONYMIZED* library.

For clarity, we only consider convolutional layers applied to images (2D) in the introduction and the experiments. But we emphasize that the theorems in Section 2 and their proofs in the supplementary materials are provided both for signals (1D) or images (2D).

1.2 Related works

Orthogonal matrices form the Stiefel Manifold and were studied in [8]. In particular, the Stiefel Manifold is compact, smooth and of known dimension. It is made of several connected components. This can be a numerical issue, since most algorithms have difficulty changing connected component during optimization. The Stiefel Manifold has many other nice properties that make it suitable to (local) Riemannian optimization [23, 24]. Orthogonal convolutional layers are a subpart of this Stiefel Manifold. To the best of our knowledge, the understanding of orthogonal convolutional layers is weak. There is no paper focusing on the theoretical properties of orthogonal convolutional layers.

Many articles [44, 7, 37, 19, 32, 10, 9] focus on Lipschitz and orthogonality constraints of the neural networks layers, from a statistical point of view, in particular in the context of adversarial attacks.

Many recent papers have investigated the numerical problem of optimizing a kernel tensor \mathbf{K} under the constraint that \mathcal{K} is orthogonal or approximately orthogonal. They also provide modeling arguments and experiments in favor of this constraint. We can distinguish two main strategies: **kernel orthogonality** [43, 7, 16, 45, 12, 39, 19, 24, 15, 19, 4] and **convolutional layer orthogonality** [25, 29, 41, 39]. The latter has been introduced more recently.

Denoting the input of the layer $X \in \mathbb{R}^{C \times SN \times SN}$, its output $Y = \text{conv}(\mathbf{K}, X) \in \mathbb{R}^{M \times N \times N}$, $x = \text{Vect}(X) \in \mathbb{R}^{CS^2N^2}$ and $y = \text{Vect}(Y) \in \mathbb{R}^{MN^2}$, the convolutional layer orthogonality relies on the fact that $y = \mathcal{K}x$ and enforces the orthogonality of \mathcal{K} .

The kernel orthogonality strategy relies on the formulation $Y = \text{Reshape}(\overline{\mathbf{K}}U(X))$, where the columns of $U(X) \in \mathbb{R}^{Ck^2 \times N^2}$ contain the patches of X required to compute the corresponding

¹The situation is more complex in [41, 29]. One of the contributions of the present paper is to clarify the situation. We describe here the clarified statement.

entries of Y . The kernel matrix $\bar{\mathbf{K}} \in \mathbb{R}^{M \times Ck^2}$ is obtained by reshaping \mathbf{K} (see, for instance, [7, 41] for more details). The kernel orthogonality strategies constrain $\bar{\mathbf{K}}$ to be orthogonal (sometimes with a normalization factor). They, however, do not truly impose \mathcal{K} to be orthogonal. In a nutshell, the problem is that the composition of an orthogonal embedding² and an orthogonal projection has no reason to be orthogonal. This phenomenon has been observed empirically in [25] and [19]. The authors of [41] and [29] also argue that, when \mathcal{K} has more columns than rows (row orthogonality), the orthogonality of $\bar{\mathbf{K}}$ is necessary but not sufficient to guarantee \mathcal{K} orthogonal.

To the best of our knowledge, the first attempt that truly imposes convolutional layer orthogonality is [25]. It does so by constructing a parameterization of a subpart of all the orthogonal convolutional layers. Then, the authors of [25] optimize the parameters for the machine learning task. In the same way, the authors of [39] construct a parameterization of a subpart of the orthogonal convolutional layers based on the Cayley transform. They use the Fourier transform to efficiently compute the inversion required by the Cayley transform. The authors of [29, 41] introduce the idea of minimizing L_{orth} (see Sect. 1.3.2 for details) as a surrogate to $\|\mathcal{K}\mathcal{K}^T - \text{Id}_{MN^2}\|_F^2$ and $\|\mathcal{K}^T\mathcal{K} - \text{Id}_{CS^2N^2}\|_F^2$. In [41], the authors provide very impressive classification experiments on CIFAR100 and ImageNet, including in a semi-supervised setting, on image inpainting, image generation and robustness.

Orthogonal convolutional layers involving a stride have been considered for the first time in [41]. This is an alternative to the isometric activation functions and pooling layers [6, 34, 1].

Finally, [33] describes an algorithm to compute the singular values of \mathcal{K} , when $S = 1$. We remind that \mathcal{K} is orthogonal when all its singular values equal 1.

1.3 Context

1.3.1 Orthogonality

In order to define orthogonal matrices, we need to distinguish two cases:

- **Row Orthogonality (RO).** When the size of the input space of $\mathcal{K} \in \mathbb{R}^{MN^2 \times CS^2N^2}$ is larger than the size of its output space, i.e. $M \leq CS^2$, \mathcal{K} is orthogonal if and only if its rows are normalized and mutually orthogonal. Denoting the identity matrix $\text{Id}_{MN^2} \in \mathbb{R}^{MN^2 \times MN^2}$, this is written

$$\mathcal{K}\mathcal{K}^T = \text{Id}_{MN^2}. \quad (1)$$

In this case, the mapping \mathcal{K} performs a dimensionality reduction.

- **Column orthogonality (CO).** When $M \geq CS^2$, \mathcal{K} is orthogonal if and only if its columns are normalized and mutually orthogonal:

$$\mathcal{K}^T\mathcal{K} = \text{Id}_{CS^2N^2}. \quad (2)$$

In this case, the mapping \mathcal{K} is an embedding.

Both the RO and CO cases are encountered in practice. When $M = CS^2$, the matrix \mathcal{K} is square and both (1) and (2) hold. The matrix \mathcal{K} is orthogonal in the usual sense and is isometric.

1.3.2 The function $L_{orth}(\mathbf{K})$

In this section, we define a variant of the function $L_{orth} : \mathbb{R}^{M \times C \times k \times k} \rightarrow \mathbb{R}$ defined in [41, 29]. The purpose of the variant is to unify the definition for the RO and CO cases.

Reminding that $k \times k$ is the size of the convolution kernel, for any $h, g \in \mathbb{R}^{k \times k}$ and any $P \in \mathbb{N}$, we define $\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1) \in \mathbb{R}^{(2P+1) \times (2P+1)}$ as the convolution³ between h and the zero padding of g (see Figure 4, in the supplementary material, Appendix A.2.1). Formally, for all $i, j \in \llbracket 0, 2P \rrbracket$,

$$[\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_{i,j} = \sum_{i',j'=0}^{k-1} h_{i',j'} \bar{g}_{i+i',j+j'}$$

²Up to a re-scaling, when considering circular boundary conditions, the mapping U is orthogonal.

³As is common in machine learning, we do not flip h .

where $\bar{g} \in \mathbb{R}^{(k+2P) \times (k+2P)}$ is defined, for all $(i, j) \in \llbracket 0, k+2P-1 \rrbracket^2$, by

$$\bar{g}_{i,j} = \begin{cases} g_{i-P,j-P} & \text{if } (i, j) \in \llbracket P, P+k-1 \rrbracket^2, \\ 0 & \text{otherwise.} \end{cases}$$

We define $\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S) \in \mathbb{R}^{(\lfloor 2P/S \rfloor + 1) \times (\lfloor 2P/S \rfloor + 1)}$, for all integer $S \geq 1$ and all $i, j \in \llbracket 0, \lfloor 2P/S \rfloor \rrbracket$, by

$$[\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S)]_{i,j} = [\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_{Si, Sj}.$$

We denote by $\mathbf{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) \in \mathbb{R}^{M \times M \times (\lfloor 2P/S \rfloor + 1) \times (\lfloor 2P/S \rfloor + 1)}$ the fourth-order tensor such that, for all $m, l \in \llbracket 1, M \rrbracket$,

$$\begin{aligned} \mathbf{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S)_{m,l,:,:) &= \\ &= \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S). \end{aligned}$$

where, for all $m \in \llbracket 1, M \rrbracket$ and $c \in \llbracket 1, C \rrbracket$, $\mathbf{K}_{m,c} \in \mathbb{R}^{k \times k}$. Finally, we denote by $\|\cdot\|_F$ the Euclidean norm in high-order tensor spaces.

Definition 1 (L_{orth}). We denote by $P = \lfloor \frac{k-1}{S} \rfloor S$ and $I_{r0} \in \mathbb{R}^{M \times M \times (2P/S+1) \times (2P/S+1)}$ the tensor whose entries are all zero except its central $M \times M$ entry which is equal to an identity matrix: $[I_{r0}]_{::,P/S,P/S} = \text{Id}_M$. We define $L_{\text{orth}} : \mathbb{R}^{M \times C \times k \times k} \rightarrow \mathbb{R}_+$ as follows

- In the RO case, $M \leq CS^2$:

$$L_{\text{orth}}(\mathbf{K}) = \|\mathbf{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2. \quad (3)$$

- In the CO case, $M \geq CS^2$:

$$L_{\text{orth}}(\mathbf{K}) = \|\mathbf{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2 - (M - CS^2).$$

When $M = CS^2$, the two definitions trivially coincide. In the definition, the padding parameter P is the largest multiple of S strictly lower than k . The difference with the definitions of L_{orth} in [41, 29] is in the CO case. In this case with $S = 1$, [29, 41] use (3) with \mathbf{K}^T instead of \mathbf{K} and $I_{c0} \in \mathbb{R}^{C \times C \times (2P/S+1) \times (2P/S+1)}$ instead of I_{r0} . In [41], the authors also indicate that, whatever S , we can use that $\|\mathcal{K}^T \mathcal{K} - \text{Id}_{CS^2 N^2}\|_F^2$ and $\|\mathcal{K} \mathcal{K}^T - \text{Id}_{MN^2}\|_F^2$ differ by a constant. We explicit this constant in Definition 1 to obtain:

$$L_{\text{orth}}(\mathbf{K}) = 0 \quad \Longleftrightarrow \quad \mathcal{K} \text{ orthogonal,}$$

both in the RO case and the CO case. Once adapted to our notations, the authors in [41, 29] propose to regularize convolutional layers parameterized by $(\mathbf{K}_l)_l$ by optimizing

$$L_{\text{task}} + \lambda \sum_l L_{\text{orth}}(\mathbf{K}_l) \quad (4)$$

where L_{task} is the original objective function of a machine learning task. The function $L_{\text{orth}}(\mathbf{K})$ does not depend on N and can be implemented in a few lines of code with Neural Network frameworks. Its gradient is then computed using automatic differentiation.

Of course, when doing so, even if the optimization is efficient, we expect $L_{\text{orth}}(\mathbf{K}_l)$ to be different from 0 but less than ε , for a small ε .

Definition 2. Let $\varepsilon > 0$, we say that a kernel tensor $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$ is $L_{\text{orth}}\text{-}\varepsilon$ if and only if

$$L_{\text{orth}}(\mathbf{K}) \leq \varepsilon.$$

We investigate, in the sequel, whether an $L_{\text{orth}}\text{-}\varepsilon$ kernel tensor \mathbf{K} , for $\varepsilon \ll 1$, still satisfies useful orthogonality properties. To do so, we investigate relations between $L_{\text{orth}}\text{-}\varepsilon$ property and the approximate orthogonality criteria and approximate isometry property defined in the next two sections.

1.3.3 Approximate orthogonality

Perfect orthogonality is an idealization that, due to floating point arithmetic, numerical and optimization errors never happens. In order to measure how \mathcal{K} deviates from being orthogonal, we define the **orthogonality residual** by $\mathcal{K}\mathcal{K}^T - \text{Id}_{MN^2}$, in the RO case, and $\mathcal{K}^T\mathcal{K} - \text{Id}_{CS^2N^2}$, in the CO case. We consider both the Frobenius norm $\|\cdot\|_F$ of the orthogonality residual or its spectral norm $\|\cdot\|_2$. We have two criteria:

$$\text{err}_N^F(\mathbf{K}) = \begin{cases} \|\mathcal{K}\mathcal{K}^T - \text{Id}_{MN^2}\|_F & , \text{ in the RO case,} \\ \|\mathcal{K}^T\mathcal{K} - \text{Id}_{CS^2N^2}\|_F & , \text{ in the CO case,} \end{cases}$$

and

$$\text{err}_N^s(\mathbf{K}) = \begin{cases} \|\mathcal{K}\mathcal{K}^T - \text{Id}_{MN^2}\|_2 & , \text{ in the RO case,} \\ \|\mathcal{K}^T\mathcal{K} - \text{Id}_{CS^2N^2}\|_2 & , \text{ in the CO case.} \end{cases}$$

When $M = CS^2$, the definitions using the RO condition and the CO condition coincide. The two criteria are of course related since for any matrix $A \in \mathbb{R}^{a \times b}$, the Frobenius and spectral norms are such that

$$\|A\|_F \leq \sqrt{\min(a, b)}\|A\|_2 \quad \text{and} \quad \|A\|_2 \leq \|A\|_F. \quad (5)$$

Finally, in the RO case, if $\text{err}_N^F(\mathbf{K}) \leq \varepsilon$, then for any i, j with $i \neq j$, we have $|\mathcal{K}_{i,:} \mathcal{K}_{j,:}^T| \leq \varepsilon$, where $\mathcal{K}_{i,:}$ is the i^{th} line of \mathcal{K} . In other words, when ε is small, the features computed by \mathcal{K} are mostly uncorrelated.

1.3.4 Approximate Isometry Property

We denote the Euclidean norm of a vector by $\|\cdot\|$. In the applications, one key property of orthogonal operators is their connection to isometries. It is the property that prevents the gradient from exploding and vanishing [7, 42, 24, 15]. This property also enables to keep the examples well separated [29], like the batch normalization does, and to have a Lipschitz forward pass and therefore improve robustness [41, 7, 25, 39, 19].

In this section, we define the ‘ ε -Approximate Isometry Property’ (ε -AIP), interpret its benefits from the machine learning perspective and give the connection between the ε -AIP and the error err_N^s .

Definition 3. A layer transform matrix $\mathcal{K} \in \mathbb{R}^{MN^2 \times CS^2N^2}$ satisfies the ε -Approximate Isometry Property if and only if

- RO case, $M \leq CS^2$:

$$\begin{cases} \forall x \in \mathbb{R}^{CS^2N^2} & \|\mathcal{K}x\|^2 \leq (1 + \varepsilon)\|x\|^2 \\ \forall y \in \mathbb{R}^{MN^2} & (1 - \varepsilon)\|y\|^2 \leq \|\mathcal{K}^Ty\|^2 \leq (1 + \varepsilon)\|y\|^2 \end{cases}$$
- CO case, $M \geq CS^2$:

$$\begin{cases} \forall x \in \mathbb{R}^{CS^2N^2} & (1 - \varepsilon)\|x\|^2 \leq \|\mathcal{K}x\|^2 \leq (1 + \varepsilon)\|x\|^2 \\ \forall y \in \mathbb{R}^{MN^2} & \|\mathcal{K}^Ty\|^2 \leq (1 + \varepsilon)\|y\|^2 \end{cases}$$

We remind that a kernel tensor \mathbf{K} can define a convolutional layer or a deconvolution layer. Deconvolution layers are, for instance, used to define layers of the decoder of an auto-encoder or variational auto-encoder [20]. In the convolutional case, \mathcal{K} is applied during the forward pass and \mathcal{K}^T is applied during the backward pass. In a deconvolution layer, \mathcal{K}^T is applied during the forward pass and \mathcal{K} during the backward pass. Depending on whether we have $M < CS^2$, $M > CS^2$ or $M = CS^2$, either \mathcal{K}^T , \mathcal{K} or both preserve distances. We summarize in Table 1 the properties of the layer, depending on the situation.

The following proposition makes the link between $\text{err}_N^s(\mathbf{K})$ and AIP. It shows that minimizing $\text{err}_N^s(\mathbf{K})$ enhances the AIP property.

Proposition 1. Let N be such that $SN \geq k$. We have, both in the RO and CO case, \mathcal{K} is $\text{err}_N^s(\mathbf{K})$ -AIP.

This statement actually holds for any matrix (not only layer transform matrix) and is already stated in [4, 12]. For completeness, we provide a proof, in the supplementary material, in Appendix F.

In Proposition 1 and in the Theorem 1 of the next section, the condition $SN \geq k$ only states that the input width and height are larger than the size of the kernels. This is always the case in practice.

Table 1: Properties of a ε -AIP layer (when $\varepsilon \ll 1$), depending on whether \mathbf{K} defines a convolutional or deconvolutional layer. The red crosses indicate when the forward or backward pass performs a dimensionality reduction.

		Forward pass		Backward pass	
		Lipschitz	Keep examples	Prevent	Prevent
		Forward pass	separated	grad. expl.	grad. vanish.
Convolutional layer	$M < CS^2$	✓	✗	✓	✓
	$M > CS^2$	✓	✓	✓	✗
Deconvolution layer	$M < CS^2$	✓	✓	✓	✗
	$M > CS^2$	✓	✗	✓	✓
Conv. & Deconv.	$M = CS^2$	✓	✓	✓	✓

2 Theoretical analysis of orthogonal convolutional layers

In all the theorems in this section, the considered convolutional layers are either applied to a signal, when $d = 1$, or an image, when $d = 2$.

We remind that the architecture of the layer is characterized by (M, C, k, S) where: M is the number of output channels; C is the number of input channels; $k \geq 1$ is odd and the convolution kernels are of size k , when $d = 1$, and $k \times k$, when $d = 2$; the stride parameter is S .

All input channels are of size SN , when $d = 1$, $SN \times SN$, when $d = 2$. The output channels are of size N and $N \times N$, respectively when $d = 1$ and 2. The definitions of L_{orth} and the errors err in the $d = 1$ case are, in the supplementary material, in Appendix A.2.

We want to highlight that the following theorems are for convolution operators defined with circular boundary conditions. Preliminary results highlighting restrictions for the ‘valid’ and ‘same’ boundary conditions are reported, in the supplementary material, in Appendix G.

2.1 Existence of orthogonal convolutional layers

The next proposition gives a necessary and sufficient condition on the architecture of a convolutional layer (M, C, k, S) and N for an orthogonal convolutional layers to exist. To simplify notations, we denote, for $d = 1$ or 2, the space of all the kernel tensors by

$$\mathbb{K}_d = \begin{cases} \mathbb{R}^{M \times C \times k} & \text{when } d = 1, \\ \mathbb{R}^{M \times C \times k \times k} & \text{when } d = 2. \end{cases}$$

We also denote, for $d = 1$ or 2,

$$\mathbb{K}_d^\perp = \{\mathbf{K} \in \mathbb{K}_d \mid \mathbf{K} \text{ is orthogonal}\}.$$

Theorem 1. *Let N be such that $SN \geq k$ and $d = 1$ or 2.*

- *RO case, i.e. $M \leq CS^d$: $\mathbb{K}_d^\perp \neq \emptyset$ if and only if $M \leq Ck^d$.*
- *CO case, i.e. $M \geq CS^d$: $\mathbb{K}_d^\perp \neq \emptyset$ if and only if $S \leq k$.*

Theorem 1 is proved, in the supplementary material, in Appendix C. Again, the conditions coincide when $M = CS^d$.

When $S \leq k$, which is by far the most common situation, there exist orthogonal convolutional layers in both the CO and the RO cases. Indeed, in the RO case, when $S \leq k$, we have $M \leq CS^d \leq Ck^d$.

However, skip-connection (also called shortcut connection) with stride, in Resnet [13] architecture for instance, usually have an architecture $(M, C, k, S) = (2C, C, 1, 2)$, where C is the number of input channels, and a 1×1 kernel. In that case, $M \leq CS^d$ and $M > Ck^d$. Theorem 1 says that there is no orthogonal convolutional layer for this type of layers.

To conclude, the main consequence of Theorem 1 is that, with circular boundary conditions and for most of the architecture used in practice (with an exception for the skip-connections with stride), there exist orthogonal convolutional layers.

2.2 Frobenius norm scalability

The following proposition mostly formalizes the reasoning leading to the regularization with L_{orth} , in [41]. It connects $L_{orth-\varepsilon}$ and the Frobenius criteria err_N^F , of Section 1.3.3.

Theorem 2. *Let N be such that $SN \geq 2k - 1$ and $d = 1$ or 2 . We have, both in the RO and CO case,*

$$(\text{err}_N^F(\mathbf{K}))^2 = N^d L_{orth}(\mathbf{K}) .$$

Theorem 2 is proved, in the supplementary material, in Appendix D. We remind that $L_{orth}(\mathbf{K})$ is independent of N .

Using Theorem 2, we find that (4) becomes

$$L_{task} + \sum_l \frac{\lambda}{N_l^d} (\text{err}_{N_l}^F(\mathbf{K}_l))^2 .$$

Once the parameter λ is made dependent of the input size of layer l , the regularization term λL_{orth} is equal to the Frobenius norm of the orthogonality residual. This justifies the use of L_{orth} as a regularizer.

Theorem 2 also guarantees that, as already said in [41, 29], when $L_{orth}(\mathbf{K}) = 0$, \mathcal{K} is orthogonal, independently of N . We have this property in both the RO and CO cases.

The downside of Theorem 2 appears when we consider \mathbf{K} fixed and not exactly orthogonal. This will practically always occur because of numerical inaccuracy or imperfect optimization, leading to $L_{orth}(\mathbf{K}) \neq 0$. In this case, considering another signal size N' and applying Theorem 2 with the size N and N' , we find

$$(\text{err}_{N'}^F(\mathbf{K}))^2 = \frac{(N')^d}{N^d} (\text{err}_N^F(\mathbf{K}))^2 .$$

To the best of our knowledge, this equality is new. This could be of importance in situations when N varies. For instance when the neural network is learned on a dataset containing signals/images of a given size, but the inference is done for signals/images of varying size [31, 26, 18].

Finally, using (5) and Proposition 1, \mathcal{K} is ϵ -AIP with ϵ scaling like the square root of the signal/image size. This might not be satisfactory. We prove in the next section that it is actually not the case.

2.3 Spectral norm scalability

In this section, we formalize that the regularization with L_{orth} provides an upper-bound on err_N^s , independently of the signal/image size.

Theorem 3. *Let N be such that $SN \geq 2k - 1$ and $d = 1$ or 2 . We have,*

$$(\text{err}_N^s(\mathbf{K}))^2 \leq \alpha L_{orth}(\mathbf{K})$$

with:

$$\alpha = \begin{cases} \left(2 \lfloor \frac{k-1}{S} \rfloor + 1\right)^d M & \text{in the RO case } (M \leq CS^d), \\ (2k-1)^d C & \text{in the CO case } (M \geq CS^d). \end{cases}$$

Theorem 3 is proved, in the supplementary material, in Appendix E. When $M = CS^d$, the two inequalities hold and it is possible to take the minimum of the two α values.

As we can see from Theorem 3, unlike with the Frobenius norm, the spectral norm of the orthogonality residual is bounded by a quantity which does not depend on N . Moreover, $\sqrt{\alpha}$ is usually moderately large. For instance, with $(M, C, k, S) = (128, 128, 3, 2)$, for images, $\sqrt{\alpha} \leq 34$. For usual architectures, $\sqrt{\alpha}$ is smaller than 200. As will be confirmed in the next Section, this ensures that, independently of N , we have a tight control of the AIP, when $L_{orth}(\mathbf{K}) \ll 1$, both in the RO and CO cases.

3 Experiments

We conduct several experiments to complement the theorems of Section 2. In order to avoid interaction with other objectives, we train a single 2D convolutional layer with circular padding. We

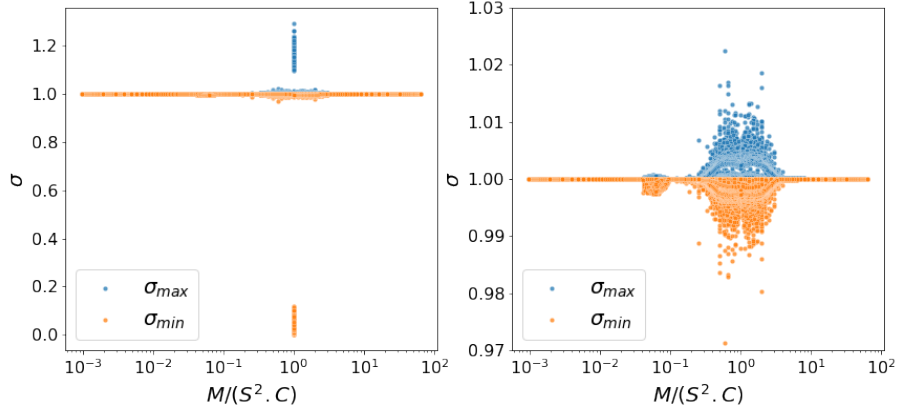


Figure 1: **Optimization of L_{orth} .** Each experiment corresponds to two dots: a blue dot for σ_{max} and an orange dot for σ_{min} . The x -axis is M/CS^2 in log scale. (left) All experiments for which $\mathbb{K}_2^\perp \neq \emptyset$; (right) All experiments for which $\mathbb{K}_2^\perp \neq \emptyset$ and $M \neq CS^2$.

explore all the architectures such that $\mathbb{K}_2^\perp \neq \emptyset$, for $C \in \llbracket 1, 64 \rrbracket$, $M \in \llbracket 1, 64 \rrbracket$, $S \in \{1, 2, 4\}$, and $k \in \{1, 3, 5, 7\}$, leading to 44924 (among 49152) configurations. The model is trained using a *Glorot uniform* initializer and a *Adam* optimizer with learning rate 0.01 on a null loss ($L_{task} = 0$) and the L_{orth} regularization (see Definition 1) during 3000 steps.

After training, we evaluate the singular values (σ) of \mathcal{K} for different input sizes $SN \times SN$. When $S = 1$, we can compute all the singular values with the algorithm in [33]. For convolutions with stride, $S > 1$, there is no practical algorithm to compute the singular values and we simply apply the well known power iteration algorithm, to retrieve the smallest and largest singular values ($\sigma_{min}, \sigma_{max}$) of \mathcal{K} (see Appendix H). We remind that, when \mathcal{K} is orthogonal, we have $\sigma_{min} = \sigma_{max} = 1$.

3.1 Optimization landscape

We plot $(\sigma_{min}, \sigma_{max})$, for \mathcal{K} such that $SN \times SN = 64 \times 64$, on Figure 1. Each experiment is represented by two points: σ_{max} , in blue, and σ_{min} , in orange. For each point, the x -axis corresponds to $\frac{M}{CS^2}$, the y -axis to the singular value. On the left of $x = 10^0$ are the architectures of the RO type, on its right those for the CO type.

Fig. 1-b shows that all configurations where $M \neq CS^2$ are trained very accurately to near perfect orthogonal convolutions. These configurations represent the vast majority of cases found in practice. However, Fig. 1-a points out that some architectures, with $M = CS^2$, might not fully benefit of regularization with L_{orth} . These architectures, corresponding to a square \mathcal{K} , can mostly be found when $M = C$ and $S = 1$, for instance in VGG [36] and Resnet [13].

3.2 Analysis of the $M = CS^2$ cases

Since we know that $\mathbb{K}_2^\perp \neq \emptyset$, the explanation for the failure cases (when σ_{max} or σ_{min} significantly differ from 1) is that the optimization was not successful. We tried many learning rate schemes, number of iterations and obtained similar results. This suggests that, in the failure cases, the landscape of L_{orth} does not lend itself to global optimization. We also ran 100 training experiments, with independent initialization, for each configuration when $M = CS^2$ ($M \in \llbracket 1, 64 \rrbracket$ and $k \in \{1, 3, 5, 7\}$). In average, at convergence, we found $\sigma_{min} \sim 1 \sim \sigma_{max}$ in 14% of runs, proving that the minimizer can be reached.

We display on Figure 2 the singular values of \mathcal{K} defined for $N \times N = 64 \times 64$ for two experiments where $M = C$. In the experiment on the left, the optimization is successful and the singular values are very accurately concentrated around 1. On the right, we see that only a few of the singular values

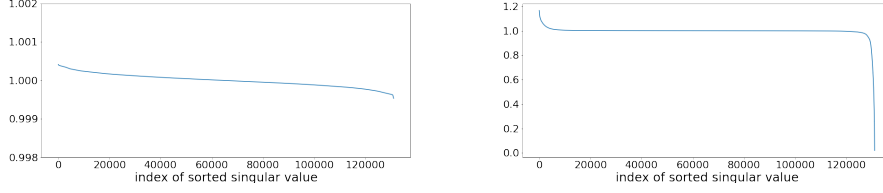


Figure 2: **Singular values of \mathcal{K}** , when $C = M$ and $S = 1$ and optimization is (Left) successful, L_{orth} small (Right) Unsuccessful, L_{orth} large.

significantly differ from 1. These experiments are representative of the many similar experiments we have conducted. This suggests that the landscape problem might have little practical impact.

3.3 Stability of $(\sigma_{min}, \sigma_{max})$ when N varies

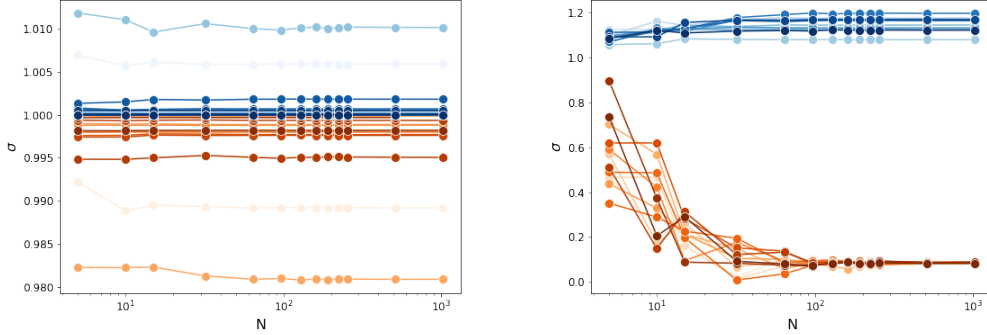


Figure 3: **Evolution of σ_{min} and σ_{max} according to input image size** (x-axis: N in log-scale) (Left) successful training, L_{orth} small, (Right) unsuccessful training, L_{orth} large

In this experiment, we evaluate how the singular values $(\sigma_{min}, \sigma_{max})$ of \mathcal{K} vary when the parameter N defining the size $SN \times SN$ of the input channels varies, for \mathbf{K} fixed. This is important for applications [35, 18, 31] using fully convolutional networks, or for transfer learning using pre-learned convolutional feature extractor.

To do so, we randomly select 50 experiments for which the optimization was successful and display, on Fig. 3, $(\sigma_{min}, \sigma_{max})$ as orange and blue dots, when $N \in \{5, 12, 15, 32, 64, 128, 256, 512, 1024\}$. The dots corresponding to the same \mathbf{K} are linked by a line. We see that for successful experiments (L_{orth} small), the singular values are very stable, when N varies. This corresponds to the behavior described in Theorem 3 and Proposition 1, when L_{orth} is small. We also display on Fig. 3 the same experiment when the optimization is unsuccessful (L_{orth} large). In this case, σ_{min} (resp. σ_{max}) values decrease (resp. increase) rapidly when N increases.

4 Conclusion

This paper provides a necessary and sufficient condition on the architecture for the existence of an orthogonal convolutional layer with circular padding. We would like also to highlight that preliminary results, reported in the Appendix G, suggest that the situation is less favorable with ‘valid’ and ‘same’ paddings. We also prove that the minimization of the surrogate L_{orth} enables to construct orthogonal convolutional layers in a stable manner, that also scales well with the input size N . The experiments confirm that this is practically the case for most of the configurations, except when $M = CS^2$ for which interrogations remain.

References

- [1] Cem Anil, James Lucas, and Roger Grosse. Sorting out lipschitz function approximation. In International Conference on Machine Learning, pages 291–301, 2019.
- [2] Martin Arjovsky and Léon Bottou. Towards principled methods for training generative adversarial networks. In International Conference on Learning Representations (ICLR 2017), 2017.
- [3] Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In International Conference on Machine Learning, pages 1120–1128, 2016.
- [4] Nitin Bansal, Xiaohan Chen, and Zhangyang Wang. Can we gain more from orthogonality regularizations in training deep networks? Advances in Neural Information Processing Systems, 31:4261–4271, 2018.
- [5] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. IEEE transactions on neural networks, 5(2):157–166, 1994.
- [6] Y Lan Boureau, Jean Ponce, and Yann Lecun. A theoretical analysis of feature pooling in visual recognition. In 27th International Conference on Machine Learning, ICML 2010, pages 111–118, 2010.
- [7] Moustapha Cisse, Piotr Bojanowski, Edouard Grave, Yann Dauphin, and Nicolas Usunier. Parseval networks: improving robustness to adversarial examples. In Proceedings of the 34th International Conference on Machine Learning-Volume 70, pages 854–863, 2017.
- [8] Alan Edelman, Tomás A Arias, and Steven T Smith. The geometry of algorithms with orthogonality constraints. SIAM journal on Matrix Analysis and Applications, 20(2):303–353, 1998.
- [9] Farzan Farnia, Jesse Zhang, and David Tse. Generalizable adversarial training via spectral normalization. In International Conference on Learning Representations, 2018.
- [10] Henry Gouk, Eibe Frank, Bernhard Pfahringer, and Michael J Cree. Regularisation of neural networks by enforcing lipschitz continuity. Machine Learning, 110(2):393–416, 2021.
- [11] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved training of wasserstein gans. In Proceedings of the 31st International Conference on Neural Information Processing Systems, pages 5769–5779, 2017.
- [12] Pei-Chang Guo and Qiang Ye. On the regularization of convolutional kernel tensors in neural networks. Linear and Multilinear Algebra, 0(0):1–13, 2020.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
- [14] Sepp Hochreiter. Untersuchungen zu dynamischen neuronalen netzen. Diploma, Technische Universität München, 91(1), 1991.
- [15] Lei Huang, Li Liu, Fan Zhu, Diwen Wan, Zehuan Yuan, Bo Li, and Ling Shao. Controllable orthogonalization in training dnns. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 6429–6438, 2020.
- [16] Lei Huang, Xianglong Liu, Bo Lang, Adams Yu, Yongliang Wang, and Bo Li. Orthogonal weight normalization: Solution to optimization over multiple dependent stiefel manifolds in deep neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 32, 2018.
- [17] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning, pages 448–456. PMLR, 2015.

- [18] Xu Ji, João F Henriques, and Andrea Vedaldi. Invariant information clustering for unsupervised image classification and segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9865–9874, 2019.
- [19] Kui Jia, Shuai Li, Yuxin Wen, Tongliang Liu, and Dacheng Tao. Orthogonal deep neural networks. IEEE transactions on pattern analysis and machine intelligence, 2019.
- [20] D. P. Kingma and M. Welling. Auto-encoding variational bayes. In International Conference on Learning Representations, 2014.
- [21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25:1097–1105, 2012.
- [22] Yann LeCun and Yoshua Bengio. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10):1995, 1995.
- [23] Mario Lezcano-Casado and David Martinez-Rubio. Cheap orthogonal constraints in neural networks: A simple parametrization of the orthogonal and unitary group. In International Conference on Machine Learning, pages 3794–3803. PMLR, 2019.
- [24] Jun Li, Fuxin Li, and Sinisa Todorovic. Efficient riemannian optimization on the stiefel manifold via the cayley transform. In International Conference on Learning Representations, 2019.
- [25] Qiyang Li, Saminul Haque, Cem Anil, James Lucas, Roger B Grosse, and Jörn-Henrik Jacobsen. Preventing gradient attenuation in lipschitz constrained convolutional networks. In Advances in neural information processing systems, pages 15390–15402, 2019.
- [26] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015.
- [27] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. In International Conference on Learning Representations, 2018.
- [28] Anh Nguyen, Jason Yosinski, and Jeff Clune. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 427–436, 2015.
- [29] Haozhi Qi, Chong You, Xiaolong Wang, Yi Ma, and Jitendra Malik. Deep isometric learning for visual recognition. In International Conference on Machine Learning, pages 7824–7835, 2020.
- [30] Haifeng Qian and Mark N Wegman. L2-nonexpansive neural networks. In International Conference on Learning Representations, 2018.
- [31] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: towards real-time object detection with region proposal networks. In Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 1, pages 91–99, 2015.
- [32] Kevin Scaman and Aladin Virmaux. Lipschitz regularity of deep neural networks: analysis and efficient estimation. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, pages 3839–3848, 2018.
- [33] Hanie Sedghi, Vineet Gupta, and Philip M Long. The singular values of convolutional layers. In International Conference on Learning Representations, 2018.
- [34] Wenling Shang, Kihyuk Sohn, Diogo Almeida, and Honglak Lee. Understanding and improving convolutional neural networks via concatenated rectified linear units. In International Conference on Machine Learning, pages 2217–2225. PMLR, 2016.
- [35] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4):640–651, 2017.

- [36] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations, 2015.
- [37] Jure Sokolić, Raja Giryes, Guillermo Sapiro, and Miguel RD Rodrigues. Robust large margin deep neural networks. IEEE Transactions on Signal Processing, 65(16):4265–4280, 2017.
- [38] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. In International Conference on Learning Representations, 2014.
- [39] Asher Trockman and J Zico Kolter. Orthogonalizing convolutional layers with the cayley transform. In International Conference on Learning Representations, 2021.
- [40] Yusuke Tsuzuku, Issei Sato, and Masashi Sugiyama. Lipschitz-margin training: scalable certification of perturbation invariance for deep neural networks. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, pages 6542–6551, 2018.
- [41] Jiayun Wang, Yubei Chen, Rudrasis Chakraborty, and Stella X Yu. Orthogonal convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11505–11515, 2020.
- [42] Lechao Xiao, Yasaman Bahri, Jascha Sohl-Dickstein, Samuel Schoenholz, and Jeffrey Pennington. Dynamical isometry and a mean field theory of cnns: How to train 10,000-layer vanilla convolutional neural networks. In International Conference on Machine Learning, pages 5393–5402. PMLR, 2018.
- [43] Di Xie, Jiang Xiong, and Shiliang Pu. All you need is beyond a good init: Exploring better solution for training extremely deep convolutional neural networks with orthonormality and modulation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 6176–6185, 2017.
- [44] Huan Xu and Shie Mannor. Robustness and generalization. Machine learning, 86(3):391–423, 2012.
- [45] Guoqiang Zhang, Kenta Niwa, and W Bastiaan Kleijn. Approximated orthonormal normalisation in training neural networks. arXiv preprint arXiv:1911.09445, 2019.
- [46] Xiang Zhang, Junbo Zhao, and Yann Lecun. Character-level convolutional networks for text classification. Advances in Neural Information Processing Systems, 2015:649–657, 2015.

A Notation

First, we start with setting notation.

A.1 Standard math definition

Recall that $\|\cdot\|_F$ denotes the norm which, to any tensor of any order, associates the square root of the sum of the squares of all its elements (e.g., for a matrix it corresponds to the Frobenius norm). For a vector $x = (x_0, \dots, x_{n-1})^T \in \mathbb{R}^n$, we recall the classic norm definitions, $\|x\|_1 = \sum_{i=0}^{n-1} |x_i|$, and $\|x\|_2 = \sqrt{\sum_{i=0}^{n-1} x_i^2}$. For $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = x^T y$ denotes the scalar product between x and y . We denote by 0_s the null vector of \mathbb{R}^s . For a matrix $A \in \mathbb{R}^{m \times n}$, $\|\cdot\|_2$ denotes the spectral norm defined by $\|A\|_2 = \sigma_{\max}(A)$, where $\sigma_{\max}(A)$ denotes the largest singular value of A . We also have $\|A\|_1 = \max_{0 \leq j \leq n-1} \sum_{i=0}^{m-1} |A_{i,j}|$ and $\|A\|_\infty = \max_{0 \leq i \leq m-1} \sum_{j=0}^{n-1} |A_{i,j}|$. We denote by $Id_n \in \mathbb{R}^{n \times n}$ the identity matrix of size n .

The floor of a real number will be denoted by $\lfloor \cdot \rfloor$. For two integers a and b , $\llbracket a, b \rrbracket$ denotes the set of integers n such that $a \leq n \leq b$. We also denote by $a \% b$ the rest of the euclidean division of a by b , and $\llbracket a, b \rrbracket \% n = \{x \% n | x \in \llbracket a, b \rrbracket\}$. We denote by $\delta_{i=j}$, the Kronecker symbol, which is equal to 1 if $i = j$, and 0 if $i \neq j$.

Recall that S is the stride paramater, $k = 2r + 1$ is the size of the 1D kernels.

For a vector space \mathcal{E} , we denote by $\mathcal{B}(\mathcal{E})$ its canonical basis. We set

$$\begin{cases} (e_i)_{i=0..k-1} = \mathcal{B}(\mathbb{R}^k) \\ (f_i)_{i=0..SN-1} = \mathcal{B}(\mathbb{R}^{SN}) \\ (E_{a,b})_{a=0..N-1, b=0..SN-1} = \mathcal{B}(\mathbb{R}^{N \times SN}) \\ (\overline{E}_{a,b})_{a=0..SN-1, b=0..N-1} = \mathcal{B}(\mathbb{R}^{SN \times N}) \\ (F_{a,b})_{a=0..SN-1, b=0..SN-1} = \mathcal{B}(\mathbb{R}^{SN \times SN}) \\ (G_{a,b})_{a=0..N-1, b=0..N-1} = \mathcal{B}(\mathbb{R}^{N \times N}) \end{cases} \quad (6)$$

Note that the indices start with 0, thus we have for example $e_0 = \begin{bmatrix} 1 \\ 0_{k-1} \end{bmatrix}$, $e_{k-1} = \begin{bmatrix} 0_{k-1} \\ 1 \end{bmatrix}$, and for

$$\text{all } i \in \llbracket 1, k-2 \rrbracket, e_i = \begin{bmatrix} 0_i \\ 1 \\ 0_{k-i-1} \end{bmatrix}.$$

To simplify the calculations, the definitions are extended for a, b outside the usual intervals, it is done by periodization. Hence, for all $a, b \in \mathbb{Z}$, denoting by $\bar{a} = a \% SN$ and $\tilde{a} = a \% N$, we set

$$\begin{cases} e_a = e_{a \% k} \\ f_a = f_{\bar{a}} \\ E_{a,b} = E_{\bar{a}, \tilde{b}} \\ \overline{E}_{a,b} = \overline{E}_{\bar{a}, \tilde{b}} \\ F_{a,b} = F_{\bar{a}, \tilde{b}} \\ G_{a,b} = G_{\bar{a}, \tilde{b}} \end{cases} \quad (7)$$

Therefore, for all $a, b, c, d \in \mathbb{Z}$, we have

$$\begin{cases} E_{a,b} F_{c,d} = \delta_{\bar{b}=\tilde{c}} E_{a,d} \\ E_{a,b} \overline{E}_{c,d} = \delta_{\bar{b}=\tilde{c}} G_{a,d} \\ \overline{E}_{a,b} E_{c,d} = \delta_{\bar{b}=\tilde{c}} F_{a,d} \\ F_{a,b} \overline{E}_{c,d} = \delta_{\bar{b}=\tilde{c}} \overline{E}_{a,d} \end{cases} \quad (8)$$

Note also that

$$E_{a,b}^T = \overline{E}_{b,a} \quad (9)$$

A.2 Corresponding 1D definitions

In this section, we give the 1D definitions of corresponding objects defined in the introduction, for the 2D case.

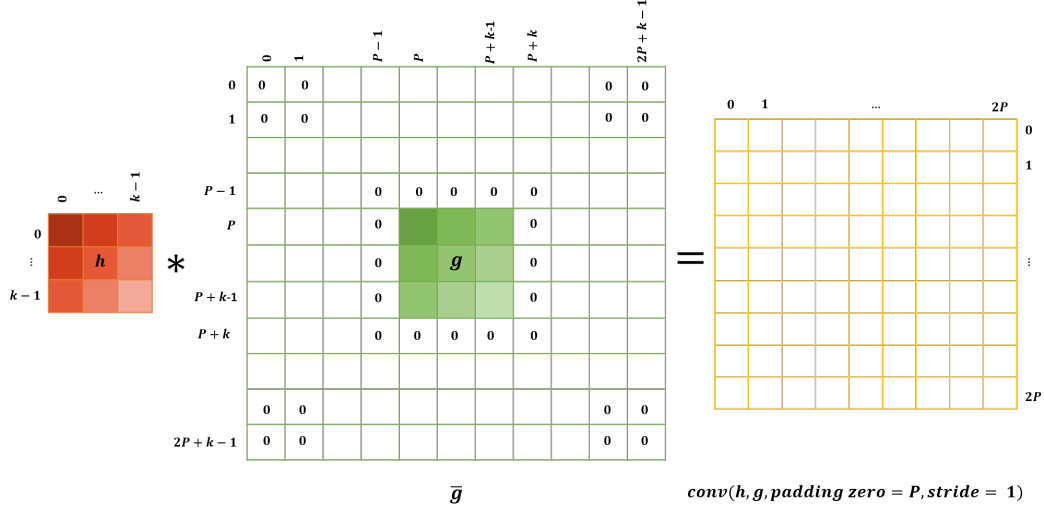


Figure 4: Illustration of $\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)$, in the 2D case.

A.2.1 Orthogonality

As before (see Section 1.3.1), we denote by $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$ the kernel tensor and $\mathcal{K} \in \mathbb{R}^{MN \times CSN}$ the matrix that applies the convolutional layer of architecture (M, C, k, S) to C vectorized channels of size SN .

RO: When $M \leq CS$, \mathcal{K} is orthogonal if and only if

$$\mathcal{K}\mathcal{K}^T = Id_{MN}.$$

CO: When $M \geq CS$, \mathcal{K} is orthogonal if and only if

$$\mathcal{K}^T\mathcal{K} = Id_{CSN}.$$

Note that, in the 1D case, we need to compare M with CS instead of CS^2 .

A.2.2 The function L_{orth}

We define L_{orth} similarly to the 2D case (see Section 1.3.2 and Figure 4). Formally, for $P \in \mathbb{N}$, and $h, g \in \mathbb{R}^k$, we define

$$\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1) \in \mathbb{R}^{2P+1} \quad (10)$$

such that for all $i \in \llbracket 0, 2P \rrbracket$,

$$[\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_i = \sum_{i'=0}^{k-1} h_{i'} \bar{g}_{i'+i}, \quad (11)$$

where \bar{g} is defined for $i \in \llbracket 0, 2P+k-1 \rrbracket$ as follows

$$\bar{g}_i = \begin{cases} g_{i-P} & \text{if } i \in \llbracket P, P+k-1 \rrbracket, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Beware not to confuse the notation \bar{g} , for $g \in \mathbb{R}^k$, with \bar{a} , for $a \in \mathbb{Z}$.

Note that, for $P' \leq P$, we have, for all $i \in \llbracket 0, 2P' \rrbracket$,

$$[\text{conv}(h, g, \text{padding zero} = P', \text{stride} = 1)]_i = [\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_{i+P-P'}. \quad (13)$$

The strided version will be denoted by $\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S) \in \mathbb{R}^{\lfloor 2P/S \rfloor + 1}$ and is defined as follows: For all $i \in \llbracket 0, \lfloor 2P/S \rfloor \rrbracket$

$$[\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S)]_i = [\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_{Si}. \quad (14)$$

Finally, reminding that for all $m \in \llbracket 1, M \rrbracket$ and $c \in \llbracket 1, C \rrbracket$, $\mathbf{K}_{m,c} \in \mathbb{R}^k$, we denote by

$$\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) \in \mathbb{R}^{M \times M \times (\lfloor 2P/S \rfloor + 1)}$$

the third-order tensor such that, for all $m, l \in \llbracket 1, M \rrbracket$,

$$\begin{aligned} \text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S)_{m,l,:} \\ = \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S). \end{aligned} \quad (15)$$

We denote by $P = \lfloor \frac{k-1}{S} \rfloor S$ and $I_{r0} \in \mathbb{R}^{M \times M \times (2P/S+1)}$ the tensor whose entries are all zero except its central $M \times M$ entry which is equal to an identity matrix: $[I_{r0}]_{:, :, P/S} = Id_M$. Put differently, we have for all $m, l \in \llbracket 1, M \rrbracket$,

$$[I_{r0}]_{m,l,:} = \delta_{m=l} \begin{bmatrix} 0_{P/S} \\ 1 \\ 0_{P/S} \end{bmatrix}. \quad (16)$$

And L_{orth} for 1D convolutions is defined as follows:

- In the RO case:

$$L_{orth}(\mathbf{K}) = \|\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2.$$

- In the CO case:

$$L_{orth}(\mathbf{K}) = \|\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2 - (M - CS).$$

Finally, the orthogonality errors are defined by

$$\text{err}_N^F(\mathbf{K}) = \begin{cases} \|\mathcal{K}\mathcal{K}^T - Id_{MN}\|_F & , \text{ in the RO case,} \\ \|\mathcal{K}^T\mathcal{K} - Id_{CSN}\|_F & , \text{ in the CO case,} \end{cases}$$

and

$$\text{err}_N^s(\mathbf{K}) = \begin{cases} \|\mathcal{K}\mathcal{K}^T - Id_{MN}\|_2 & , \text{ in the RO case,} \\ \|\mathcal{K}^T\mathcal{K} - Id_{CSN}\|_2 & , \text{ in the CO case.} \end{cases}$$

B The convolutional layer as a matrix-vector product

In this section, we write the convolutional layer as a matrix-vector product. In other words, we explicit \mathcal{K} and the ingredients composing it. Note that this is already known and can be found for example in [33].

B.1 1D case

B.1.1 Preliminaries

We denote by $S_N \in \mathbb{R}^{N \times SN}$ the sampling matrix (i.e., for $x = (x_0, \dots, x_{SN-1})^T \in \mathbb{R}^{SN}$, we have for all $m \in \llbracket 0, N-1 \rrbracket$, $(S_N x)_m = x_{Sm}$).

More precisely, we have

$$S_N = \sum_{i=0}^{N-1} E_{i, Si}. \quad (17)$$

Also, note that, using (8) and (9), we have

$$S_N S_N^T = Id_N \quad (18)$$

and

$$S_N^T S_N = \sum_{i=0}^{N-1} F_{Si, Si}. \quad (19)$$

For a vector $x = (x_0, \dots, x_{n-1})^T \in \mathbb{R}^n$, we denote by $C(x) \in \mathbb{R}^{n \times n}$ the circulant matrix defined by

$$C(x) = \begin{pmatrix} x_0 & x_{n-1} & \cdots & x_2 & x_1 \\ x_1 & x_0 & x_{n-1} & & x_2 \\ \vdots & x_1 & x_0 & \ddots & \vdots \\ x_{n-2} & & \ddots & \ddots & x_{n-1} \\ x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \end{pmatrix}. \quad (20)$$

In other words, for $x \in \mathbb{R}^n$ and $X \in \mathbb{R}^{n \times n}$, we have

$$X = C(x) \iff \forall m, l \in \llbracket 0, n-1 \rrbracket, X_{m,l} = x_{(m-l) \% n}. \quad (21)$$

We also denote by $\tilde{x} \in \mathbb{R}^n$ the vector such that for all $i \in \llbracket 0, n-1 \rrbracket$, $\tilde{x}_i = x_{(-i) \% n}$ (Again, the notation \tilde{x} , for $x \in \mathbb{R}^n$, should not be confused with \tilde{a} , for $a \in \mathbb{Z}$). We have

$$C(x)^T = C(\tilde{x}). \quad (22)$$

Also, for $x, y \in \mathbb{R}^n$, we have

$$C(x)C(y) = C(x * y), \quad (23)$$

where $x * y \in \mathbb{R}^n$, is such that for all $j \in \llbracket 0, n-1 \rrbracket$,

$$[x * y]_j = \sum_{i=0}^{n-1} x_i y_{(j-i) \% n}. \quad (24)$$

Remark that $x * y$ is n -periodic. Note also that $x * y = y * x$ and therefore

$$C(x)C(y) = C(y)C(x). \quad (25)$$

Throughout the article, the size of a filter is smaller than the size of the signal ($k = 2r + 1 \leq SN$). Therefore, for $n \geq k$, we introduce an operator P_n which associates to each $h = (h_0, \dots, h_{2r})^T \in \mathbb{R}^k$ the corresponding vector

$$P_n(h) = (h_r, \dots, h_1, h_0, 0, \dots, 0, h_{2r}, \dots, h_{r+1})^T \in \mathbb{R}^n.$$

Setting $[P_n(h)]_i = [P_n(h)]_{i \% n}$ for all $i \in \mathbb{Z}$, we have the following formula for P_n : for $i \in \llbracket -r, -r + n - 1 \rrbracket$,

$$[P_n(h)]_i = \begin{cases} h_{r-i} & \text{if } i \in \llbracket -r, r \rrbracket \\ 0 & \text{otherwise.} \end{cases} \quad (26)$$

B.1.2 The convolutional layer as a matrix-vector product

Single-channel case: Let $x = (x_0, \dots, x_{SN-1})^T \in \mathbb{R}^{SN}$ be a 1D signal. We denote by $\text{Circular_Conv}(h, x, \text{stride} = 1)$ the result of the circular convolution of x with the kernel $h = (h_0, \dots, h_{2r})^T \in \mathbb{R}^k$. We have

$$\text{Circular_Conv}(h, x, \text{stride} = 1) = \left(\sum_{i'=0}^{k-1} h_{i'} x_{(i'+i-r) \% SN} \right)_{i=0 \dots SN-1}.$$

Written as a matrix-vector product, this becomes

$$\begin{aligned}
& \begin{pmatrix} h_0 & \cdots & h_{2r} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & \cdots & h_{2r} \end{pmatrix} \begin{pmatrix} x_{SN-r} \\ \vdots \\ x_{SN-1} \\ x_0 \\ \vdots \\ x_{SN-1} \\ x_0 \\ \vdots \\ x_{r-1} \end{pmatrix} \in \mathbb{R}^{SN} \\
& = \begin{pmatrix} h_r & h_{r+1} & \cdots & h_{2r} & 0 & \cdots & 0 & h_0 & \cdots & h_{r-1} \\ h_{r-1} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & h_0 \\ h_0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & h_{2r} \\ h_{2r} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & h_{r+1} \\ h_{r+1} & \cdots & h_{2r} & 0 & \cdots & 0 & h_0 & \cdots & h_{r-1} & h_r \end{pmatrix} x \\
& = C(P_{SN}(h))x .
\end{aligned}$$

The strided convolution is

$$\text{Circular_Conv}(h, x, \text{stride} = S) = S_N C(P_{SN}(h))x \in \mathbb{R}^N . \quad (27)$$

Notice that $S_N C(P_{SN}(h)) \in \mathbb{R}^{N \times SN}$.

Multi-channel convolution: Let $X \in \mathbb{R}^{C \times SN}$ be a multi-channel 1D signal. We denote by $\text{Circular_Conv}(\mathbf{K}, X, \text{stride} = S)$ the result of the strided circular convolutional layer of kernel $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$ applied to X . Using (27) for all the input-output channel correspondances, we have $Y = \text{Circular_Conv}(\mathbf{K}, X, \text{stride} = S) \in \mathbb{R}^{M \times N}$ if and only if

$$\text{Vect}(Y) = \begin{pmatrix} S_N C(P_{SN}(\mathbf{K}_{1,1})) & \cdots & S_N C(P_{SN}(\mathbf{K}_{1,C})) \\ \vdots & \vdots & \vdots \\ S_N C(P_{SN}(\mathbf{K}_{M,1})) & \cdots & S_N C(P_{SN}(\mathbf{K}_{M,C})) \end{pmatrix} \text{Vect}(X) ,$$

where $\mathbf{K}_{i,j} = \mathbf{K}_{i,j,:} \in \mathbb{R}^k$. Therefore,

$$\mathcal{K} = \begin{pmatrix} S_N C(P_{SN}(\mathbf{K}_{1,1})) & \cdots & S_N C(P_{SN}(\mathbf{K}_{1,C})) \\ \vdots & \vdots & \vdots \\ S_N C(P_{SN}(\mathbf{K}_{M,1})) & \cdots & S_N C(P_{SN}(\mathbf{K}_{M,C})) \end{pmatrix} \in \mathbb{R}^{MN \times CSN} \quad (28)$$

is the layer transform matrix associated to kernel \mathbf{K} .

B.2 2D case

Notice that, since they are very similar, the proofs and notation are detailed in the 1D case, but we only provide a sketch of the proof and the main equations in 2D.

B.2.1 Preliminaries

We denote by $\mathcal{S}_N \in \mathbb{R}^{N^2 \times S^2 N^2}$ the sampling matrix in the 2D case (i.e., for a matrix $x \in \mathbb{R}^{SN \times SN}$, if we denote by $z \in \mathbb{R}^{N \times N}$, such that for all $i, j \in \llbracket 0, N-1 \rrbracket$, $z_{i,j} = x_{Si, Sj}$, then $\text{Vect}(z) = \mathcal{S}_N \text{Vect}(x)$).

For a matrix $x \in \mathbb{R}^{n \times n}$, we denote by $\mathcal{C}(x) \in \mathbb{R}^{n^2 \times n^2}$ the doubly-block circulant matrix defined by

$$\mathcal{C}(x) = \begin{pmatrix} C(x_{0,:}) & C(x_{n-1,:}) & \cdots & C(x_{2,:}) & C(x_{1,:}) \\ C(x_{1,:}) & C(x_{0,:}) & C(x_{n-1,:}) & & C(x_{2,:}) \\ \vdots & C(x_{1,:}) & C(x_{0,:}) & \ddots & \vdots \\ C(x_{n-2,:}) & & \ddots & \ddots & C(x_{n-1,:}) \\ C(x_{n-1,:}) & C(x_{n-2,:}) & \cdots & C(x_{1,:}) & C(x_{0,:}) \end{pmatrix}. \quad (29)$$

For $n \geq k = 2r + 1$, we introduce the operator \mathcal{P}_n which associates to a matrix $h \in \mathbb{R}^{k \times k}$ the corresponding matrix

$$\mathcal{P}_n(h) = \begin{pmatrix} h_{r,r} & \cdots & h_{r,0} & 0 & \cdots & 0 & h_{r,2r} & \cdots & h_{r,r+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{0,r} & \cdots & h_{0,0} & 0 & \cdots & 0 & h_{0,2r} & \cdots & h_{0,r+1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ h_{2r,r} & \cdots & h_{2r,0} & 0 & \cdots & 0 & h_{2r,2r} & \cdots & h_{2r,r+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{r+1,r} & \cdots & h_{r+1,0} & 0 & \cdots & 0 & h_{r+1,2r} & \cdots & h_{r+1,r+1} \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Setting $[\mathcal{P}_n(h)]_{i,j} = [\mathcal{P}_n(h)]_{i \% n, j \% n}$ for all $i, j \in \mathbb{Z}$, we have the following formula for \mathcal{P}_n : for $(i, j) \in \llbracket -r, -r + n - 1 \rrbracket^2$,

$$[\mathcal{P}_n(h)]_{i,j} = \begin{cases} h_{r-i, r-j} & \text{if } (i, j) \in \llbracket -r, r \rrbracket^2 \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

B.2.2 The convolutional layer as a matrix-vector product

Single-channel case: Let $x \in \mathbb{R}^{SN \times SN}$ be a 2D image. We denote by $\text{Circular_Conv}(h, x, \text{stride} = 1)$ the result of the circular convolution of x with the kernel $h \in \mathbb{R}^{k \times k}$. As in the 1D case, we have

$$y = \text{Circular_Conv}(h, x, \text{stride} = 1) \iff \text{Vect}(y) = \mathcal{C}(\mathcal{P}_{SN}(h)) \text{Vect}(x)$$

and the strided circular convolution

$$y = \text{Circular_Conv}(h, x, \text{stride} = S) \iff \text{Vect}(y) = \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(h)) \text{Vect}(x).$$

Notice that $\mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(h)) \in \mathbb{R}^{N^2 \times S^2 N^2}$.

Multi-channel convolution : Let $X \in \mathbb{R}^{C \times SN \times SN}$ be a multi-channel 2D image. We denote by $\text{Circular_Conv}(\mathbf{K}, X, \text{stride} = S)$ the result of the strided circular convolutional layer of kernel $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$ applied to X . We have $Y = \text{Circular_Conv}(\mathbf{K}, X, \text{stride} = S) \in \mathbb{R}^{M \times N \times N}$ if and only if

$$\text{Vect}(Y) = \begin{pmatrix} \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{1,1})) & \cdots & \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{1,C})) \\ \vdots & \vdots & \vdots \\ \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{M,1})) & \cdots & \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{M,C})) \end{pmatrix} \text{Vect}(X),$$

where $\mathbf{K}_{i,j} = \mathbf{K}_{i,j,:,:} \in \mathbb{R}^{k \times k}$. Therefore,

$$\mathcal{K} = \begin{pmatrix} \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{1,1})) & \cdots & \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{1,C})) \\ \vdots & \vdots & \vdots \\ \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{M,1})) & \cdots & \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(\mathbf{K}_{M,C})) \end{pmatrix} \in \mathbb{R}^{MN^2 \times CS^2 N^2} \quad (31)$$

is the layer transform matrix associated to kernel \mathbf{K} .

C Proof of Theorem 1

As the proof is very similar in 1D and 2D case, we give the full proof in the 1D case and we only give a sketch of the proof in the 2D case.

C.1 Proof of Theorem 1, for 1D convolutional layers

We start by stating and proving three intermediate lemmas. Recall that $k = 2r + 1$ and from (6), that $(e_i)_{i=0..k-1} = \mathcal{B}(\mathbb{R}^k)$ and $(E_{a,b})_{a=0..N-1, b=0..SN-1} = \mathcal{B}(\mathbb{R}^{N \times SN})$.

Lemma 1. *Let $j \in \llbracket 0, k - 1 \rrbracket$. We have*

$$S_N C(P_{SN}(e_j)) = \sum_{i=0}^{N-1} E_{i, Si+j-r}.$$

Proof. Let $j \in \llbracket 0, k - 1 \rrbracket$. Using (26), (20), (6) and (7), we have

$$\begin{aligned} C(P_{SN}(e_j)) &= C(f_{r-j}) \\ &= \sum_{i=0}^{SN-1} F_{i, i-(r-j)} \\ &= \sum_{i=0}^{SN-1} F_{i, i+j-r}. \end{aligned}$$

Using (17) and (8), we have

$$\begin{aligned} S_N C(P_{SN}(e_j)) &= \left(\sum_{i=0}^{N-1} E_{i, Si} \right) \left(\sum_{i'=0}^{SN-1} F_{i', i'+j-r} \right) \\ &= \sum_{i=0}^{N-1} E_{i, Si+j-r}. \end{aligned}$$

□

Lemma 2. *Let $k_S = \min(k, S)$ and $j, l \in \llbracket 0, k_S - 1 \rrbracket$. We have*

$$S_N C(P_{SN}(e_j)) C(P_{SN}(e_l))^T S_N^T = \delta_{j=l} Id_N.$$

Proof. Let $j, l \in \llbracket 0, k_S - 1 \rrbracket$. Since $k_S \leq k$, using Lemma 1 and (9),

$$\begin{aligned} S_N C(P_{SN}(e_j)) C(P_{SN}(e_l))^T S_N^T &= \left(\sum_{i=0}^{N-1} E_{i, Si+j-r} \right) \left(\sum_{i'=0}^{N-1} E_{i', Si'+l-r} \right)^T \\ &= \left(\sum_{i=0}^{N-1} E_{i, Si+j-r} \right) \left(\sum_{i'=0}^{N-1} \overline{E}_{Si'+l-r, i'} \right). \end{aligned} \quad (32)$$

We know from (8) that $E_{i, Si+j-r} \overline{E}_{Si'+l-r, i'} = \delta_{\overline{Si+j-r} = \overline{Si'+l-r}} G_{i, i'}$. But for $i, i' \in \llbracket 0, N - 1 \rrbracket$ and $j, l \in \llbracket 0, k_S - 1 \rrbracket$, since $k_S \leq S$, we have

$$-r \leq Si + j - r \leq S(N - 1) + k_S - 1 - r \leq SN - 1 - r.$$

Similarly, $Si' + l - r \in \llbracket -r, SN - 1 - r \rrbracket$. Therefore, $Si + j - r$ and $Si' + l - r$ lie in the same interval of size SN , hence

$$\begin{aligned} \overline{Si + j - r} = \overline{Si' + l - r} &\iff Si + j - r = Si' + l - r \\ &\iff Si + j = Si' + l. \end{aligned}$$

If $Si + j = Si' + l$, then

$$S \geq k_S > |j - l| = |S(i - i')|.$$

Therefore, since $|i - i'| \in \mathbb{N}$, this implies $i = i'$ and, as a consequence, $j = l$. Finally,

$$\overline{Si + j - r} = \overline{Si' + l - r} \iff i = i' \text{ and } j = l.$$

Hence, using (8), the equality (32) becomes

$$\begin{aligned} S_N C(P_{SN}(e_j)) C(P_{SN}(e_l))^T S_N^T &= \delta_{j=l} \sum_{i=0}^{N-1} G_{i,i} \\ &= \delta_{j=l} Id_N. \end{aligned}$$

□

Lemma 3. *Let $S \leq k$. We have*

$$\sum_{z=0}^{S-1} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) = Id_{SN}.$$

Proof. Using Lemma 1, (9) and (8), we have for all $z \in \llbracket 0, S-1 \rrbracket$,

$$\begin{aligned} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) &= \left(\sum_{i=0}^{N-1} E_{i, Si+z-r} \right)^T \left(\sum_{i'=0}^{N-1} E_{i', Si'+z-r} \right) \\ &= \left(\sum_{i=0}^{N-1} \bar{E}_{Si+z-r, i} \right) \left(\sum_{i'=0}^{N-1} E_{i', Si'+z-r} \right) \\ &= \sum_{i=0}^{N-1} F_{Si+z-r, Si+z-r}. \end{aligned}$$

Hence

$$\sum_{z=0}^{S-1} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) = \sum_{z=0}^{S-1} \sum_{i=0}^{N-1} F_{Si+z-r, Si+z-r}.$$

But, for $z \in \llbracket 0, S-1 \rrbracket$ and $i \in \llbracket 0, N-1 \rrbracket$, we have $Si+z-r$ traverses $\llbracket -r, SN-1-r \rrbracket$. Therefore, using (7)

$$\sum_{z=0}^{S-1} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) = \sum_{i=-r}^{SN-1-r} F_{i,i} = \sum_{i=0}^{SN-1} F_{i,i} = Id_{SN}.$$

□

Proof of Theorem 1. Let N be a positive integer such that $SN \geq k$.

We start by proving the theorem in the RO case.

Suppose $CS \geq M$ and $M \leq Ck$:

Let us exhibit $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$ such that $\mathcal{K}\mathcal{K}^T = Id_{MN}$.

Let $k_S = \min(k, S)$. Since $M \leq CS$ and $M \leq Ck$, we have $M \leq Ck_S$. Therefore, there exist a unique couple $(i_{max}, j_{max}) \in \llbracket 0, k_S-1 \rrbracket \times \llbracket 1, C \rrbracket$ such that $M = i_{max}C + j_{max}$. We define the kernel tensor $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$ as follows: For all $(i, j) \in \llbracket 0, k_S-1 \rrbracket \times \llbracket 1, C \rrbracket$ such that $iC + j \leq M$, we set $\mathbf{K}_{iC+j, j} = e_i$, and $\mathbf{K}_{u, v} = 0$ for all the other indices. Put differently, if we write \mathbf{K} as a 3rd order tensor (where the rows represent the first dimension, the columns the second one, and the $\mathbf{K}_{i, j} \in \mathbb{R}^k$ are in the third dimension) we have :

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{1,1} & \cdots & \mathbf{K}_{1,C} \\ \vdots & \vdots & \vdots \\ \mathbf{K}_{M,1} & \cdots & \mathbf{K}_{M,C} \end{pmatrix} = \begin{bmatrix} e_0 & & \\ & \ddots & \\ & & 0 \\ & & e_0 \\ e_1 & & \\ 0 & \ddots & 0 \\ & & e_1 \\ & & \vdots \\ e_{i_{max}} & & \\ 0 & \ddots & 0 \end{bmatrix} \in \mathbb{R}^{M \times C \times k},$$

where $e_{i_{max}}$ appears j_{max} times. Therefore, using (28), we have

$$\mathcal{K} = \begin{bmatrix} S_NC(P_{SN}(e_0)) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_0)) \\ S_NC(P_{SN}(e_1)) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_1)) \\ & & \vdots & \\ S_NC(P_{SN}(e_{i_{max}})) & & & \\ & 0 & \ddots & 0 \end{bmatrix} \in \mathbb{R}^{MN \times CSN},$$

where $S_NC(P_{SN}(e_{i_{max}}))$ appears j_{max} times. We have $\mathcal{K} = D_{1:MN,:}$, where we set

$$D = \begin{bmatrix} S_NC(P_{SN}(e_0)) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_0)) \\ S_NC(P_{SN}(e_1)) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_1)) \\ & & \vdots & \\ S_NC(P_{SN}(e_{k_S-1})) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_{k_S-1})) \end{bmatrix} \in \mathbb{R}^{k_S CN \times CSN}.$$

But, for $j, l \in \llbracket 0, k_S - 1 \rrbracket$, the (j, l) -th block of size (CN, CN) of DD^T is :

$$\begin{bmatrix} S_NC(P_{SN}(e_j)) & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_j)) \end{bmatrix} \begin{bmatrix} C(P_{SN}(e_l))^T S_N^T & & & \\ & 0 & \ddots & 0 \\ & & & C(P_{SN}(e_l))^T S_N^T \end{bmatrix},$$

which is equal to

$$\begin{bmatrix} S_NC(P_{SN}(e_j))C(P_{SN}(e_l))^T S_N^T & & & \\ & 0 & \ddots & 0 \\ & & & S_NC(P_{SN}(e_j))C(P_{SN}(e_l))^T S_N^T \end{bmatrix}.$$

Using Lemma 2, this is equal to $\delta_{j=l} Id_{CN}$. Hence, $DD^T = Id_{k_S CN}$, and therefore,

$$\mathcal{K}\mathcal{K}^T = D_{1:MN,:}(D_{1:MN,:})^T = (DD^T)_{1:MN,1:MN} = Id_{MN}.$$

This proves the first implication in the RO case, i.e., if $M \leq Ck$, then $\mathbb{K}_1^\perp \neq \emptyset$.

Suppose $CS \geq M$ and $M > Ck$:

We need to prove that for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, we have $\mathcal{K}\mathcal{K}^T \neq Id_{MN}$.

Since for all $(i, j) \in \llbracket 1, M \rrbracket \times \llbracket 1, C \rrbracket$, each of the N rows of $S_NC(P_{SN}(\mathbf{K}_{i,j}))$ has at most k non-zero elements, the number of non-zero columns of $S_NC(P_{SN}(\mathbf{K}_{i,j}))$ is less than or equal to kN . Also, for all $i, i' \in \llbracket 1, M \rrbracket$, the columns of $S_NC(P_{SN}(\mathbf{K}_{i,j}))$ which can be non-zero are the same as those of $S_NC(P_{SN}(\mathbf{K}_{i',j}))$. Hence, we have for all j , the number of non-zero columns of

$\begin{bmatrix} S_NC(P_{SN}(\mathbf{K}_{1,j})) \\ \vdots \\ S_NC(P_{SN}(\mathbf{K}_{M,j})) \end{bmatrix}$ is less than or equal to kN . Therefore, the number of non-zero columns of

\mathcal{K} is less than or equal to CkN . Hence, since $Ck < M$, we have $\text{rk}(\mathcal{K}\mathcal{K}^T) \leq \text{rk}(\mathcal{K}) \leq CkN < MN = \text{rk}(Id_{MN})$. Therefore, for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, we have

$$\mathcal{K}\mathcal{K}^T \neq Id_{MN}.$$

This proves that if $CS \geq M$ and $M > Ck$, then $\mathbb{K}_1^\perp = \emptyset$. This concludes the proof for the RO case.

Suppose $M \geq CS$ and $S \leq k$:

Let us exhibit $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$ such that $\mathcal{K}^T \mathcal{K} = Id_{CSN}$.

For all $(i, j) \in \llbracket 0, S-1 \rrbracket \times \llbracket 1, C \rrbracket$, we set $\mathbf{K}_{iC+j, j} = e_i$, and $\mathbf{K}_{u, v} = 0$ for all the other indices. Put differently, if we write \mathbf{K} as a 3rd order tensor, we have

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{1,1} & \cdots & \mathbf{K}_{1,C} \\ \vdots & \vdots & \vdots \\ \mathbf{K}_{M,1} & \cdots & \mathbf{K}_{M,C} \end{pmatrix} = \begin{bmatrix} e_0 & & \\ & \ddots & 0 \\ & & e_0 \\ e_1 & & \\ 0 & \ddots & 0 \\ & & e_1 \\ & \vdots & \\ e_{S-1} & & \\ 0 & \ddots & 0 \\ & & e_{S-1} \\ & O & \end{bmatrix} \in \mathbb{R}^{M \times C \times k},$$

where $O = 0_{(M-CS) \times C \times k}$ denotes the null tensor. Therefore, using (28), we have

$$\mathcal{K} = \begin{bmatrix} S_N C(P_{SN}(e_0)) & & \\ 0 & \ddots & 0 \\ & & S_N C(P_{SN}(e_0)) \\ S_N C(P_{SN}(e_1)) & & \\ 0 & \ddots & 0 \\ & & S_N C(P_{SN}(e_1)) \\ & \vdots & \\ S_N C(P_{SN}(e_{S-1})) & & \\ 0 & \ddots & 0 \\ & & S_N C(P_{SN}(e_{S-1})) \\ & O & \end{bmatrix} \in \mathbb{R}^{MN \times CSN},$$

where $O = 0_{(MN-CSN) \times CSN}$ denotes the null matrix. Hence,

$$\begin{aligned} \mathcal{K}^T \mathcal{K} &= \begin{bmatrix} \sum_{z=0}^{S-1} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) & & 0 \\ & \ddots & \\ 0 & & \sum_{z=0}^{S-1} C(P_{SN}(e_z))^T S_N^T S_N C(P_{SN}(e_z)) \end{bmatrix}. \end{aligned}$$

Using Lemma 3, we obtain

$$\mathcal{K}^T \mathcal{K} = Id_{CSN}.$$

This proves that in the CO case, if $S \leq k$, then $\mathbb{K}_1^\perp \neq \emptyset$.

Suppose $M \geq CS$ and $S > k$:

We need to prove that for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, we have $\mathcal{K}^T \mathcal{K} \neq Id_{CSN}$.

Following the same reasoning as for the case $CS \geq M$ and $M > Ck$, we have that the number of non-zero columns of \mathcal{K} is less than or equal to CkN . So, since $k < S$, we have $\text{rk}(\mathcal{K}^T \mathcal{K}) \leq \text{rk}(\mathcal{K}) \leq CkN < CSN = \text{rk}(Id_{CSN})$. Therefore, for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, we have

$$\mathcal{K}^T \mathcal{K} \neq Id_{CSN}.$$

This proves that in the CO case, if $k < S$, then $\mathbb{K}_1^\perp = \emptyset$. This concludes the proof. \square

C.2 Sketch of the proof of Theorem 1, for 2D convolutional layers

We first set $(e_{i,j})_{i=0..k-1,j=0..k-1} = \mathcal{B}(\mathbb{R}^{k \times k})$. As in the 1D case, we have the following two lemmas

Lemma 4. Let $k_S = \min(k, S)$ and $j, j', l, l' \in \llbracket 0, k_S - 1 \rrbracket$. We have

$$\mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(e_{j,j'})) \mathcal{C}(\mathcal{P}_{SN}(e_{l,l'}))^T \mathcal{S}_N^T = \delta_{j=l} \delta_{j'=l'} Id_{N^2} .$$

Lemma 5. Let $S \leq k$. We have

$$\sum_{z=0}^{S-1} \sum_{z'=0}^{S-1} \mathcal{C}(\mathcal{P}_{SN}(e_{z,z'}))^T \mathcal{S}_N^T \mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(e_{z,z'})) = Id_{S^2 N^2} .$$

For $CS^2 \geq M$ and $M \leq Ck^2$:

We set $\bar{e}_{i+kj} = e_{i,j}$ for $i, j \in \llbracket 0, k-1 \rrbracket$.

Let $i_{max}, j_{max} \in \llbracket 0, k_S^2 - 1 \rrbracket \times \llbracket 1, C \rrbracket$ such that $i_{max}C + j_{max} = M$. We set

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{1,1} & \cdots & \mathbf{K}_{1,C} \\ \vdots & \vdots & \vdots \\ \mathbf{K}_{M,1} & \cdots & \mathbf{K}_{M,C} \end{pmatrix} = \begin{bmatrix} \bar{e}_0 & & \\ 0 & \ddots & 0 \\ & & \bar{e}_0 \\ \bar{e}_1 & & \\ 0 & \ddots & 0 \\ & & \bar{e}_1 \\ & & \vdots \\ \bar{e}_{i_{max}} & & \\ 0 & \ddots & 0 \end{bmatrix} \in \mathbb{R}^{M \times C \times k \times k} ,$$

where $\bar{e}_{i_{max}}$ appears j_{max} times. Then we proceed as in the 1D case.

For $CS^2 \geq M$ and $M > Ck^2$:

Using the same argument as in 1D, we can conclude that the number of non-zero columns of \mathcal{K} is less than or equal to $Ck^2 N^2$. Hence, $\text{rk}(\mathcal{K}) \leq Ck^2 N^2 < MN^2$. Therefore, for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$, we have

$$\mathcal{K} \mathcal{K}^T \neq Id_{MN^2} .$$

For $M \geq CS^2$ and $S \leq k$:

Denoting by $O \in \mathbb{R}^{(M-S^2C) \times C \times k \times k}$ the null 4th order tensor of size $(M - S^2C) \times C \times k \times k$, we set

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{1,1} & \cdots & \mathbf{K}_{1,C} \\ \vdots & \vdots & \vdots \\ \mathbf{K}_{M,1} & \cdots & \mathbf{K}_{M,C} \end{pmatrix} = \begin{bmatrix} e_{0,0} & & \\ 0 & \ddots & 0 \\ & & e_{0,0} \\ e_{1,0} & & \\ 0 & \ddots & 0 \\ & & e_{1,0} \\ & & \vdots \\ e_{S-1,S-1} & & \\ 0 & \ddots & 0 \\ & & e_{S-1,S-1} \\ & & O \end{bmatrix} \in \mathbb{R}^{M \times C \times k \times k} .$$

Then we proceed as in 1D.

For $M \geq CS^2$ and $S > k$:

By the same reasoning as in 1D, we have that the number of non-zero columns of \mathcal{K} is less than

or equal to Ck^2N^2 . So, since $k < S$, we have $\text{rk}(\mathcal{K}) \leq Ck^2N^2 < CS^2N^2$. Therefore, for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$, we have

$$\mathcal{K}^T \mathcal{K} \neq \text{Id}_{CS^2N^2} .$$

D Proof of Theorem 2

As in the previous section, we give the full proof in the 1D case and a sketch of proof for 2D.

D.1 Proof of Theorem 2, for 1D convolutional layers

Before proving Theorem 2, we first present three intermediate lemmas.

Lemma 6. *Let $x \in \mathbb{R}^{SN}$. We have*

$$S_N C(x) S_N^T = C(S_N x) .$$

Proof. Let $x \in \mathbb{R}^{SN}$, $X = C(x)$ and $Y = S_N X S_N^T \in \mathbb{R}^{N \times N}$. The matrix Y is formed by sampling X , i.e., for all $m, n \in \llbracket 0, N-1 \rrbracket$,

$$Y_{m,n} = X_{Sm,Sn} .$$

Hence, using (21), $Y_{m,n} = x_{(Sm-Sn)\%SN} = x_{S((m-n)\%N)}$. Setting $y = S_N x$, we have $y_l = x_{Sl}$ for all $l \in \llbracket 0, N-1 \rrbracket$. Therefore, $Y_{m,n} = y_{(m-n)\%N}$, and using (21), we obtain

$$Y = C(y) .$$

Hence we conclude that

$$S_N C(x) S_N^T = C(S_N x) .$$

This completes the proof of the lemma. \square

For N such that $SN \geq 2k-1$, and $P = \lfloor \frac{k-1}{S} \rfloor S$, we introduce the operator $Q_{S,N}$ which associates to a vector $x = (x_0, \dots, x_{2\frac{P}{S}})^T \in \mathbb{R}^{2\frac{P}{S}+1}$, the vector

$$Q_{S,N}(x) = (x_{\frac{P}{S}}, \dots, x_{2\frac{P}{S}}, 0, \dots, 0, x_0, x_1, \dots, x_{\frac{P}{S}-1})^T \in \mathbb{R}^N . \quad (33)$$

Lemma 7. *Let $S, k = 2r+1$ and N be positive integers such that $SN \geq 2k-1$. Let $h, g \in \mathbb{R}^k$ and $P = \lfloor \frac{k-1}{S} \rfloor S$, we have*

$$S_N C(P_{SN}(h)) C(P_{SN}(g))^T S_N^T = C(Q_{S,N}(\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S))) .$$

Proof. Let N be such that $SN \geq 2k-1$, and $P = \lfloor \frac{k-1}{S} \rfloor S$. Recall that by definition $P_{SN}(h)$ is SN -periodic: $[P_{SN}(h)]_i = [P_{SN}(h)]_{i\%SN}$ for all $i \in \mathbb{Z}$. Using (22), (23), and (24), we have

$$\begin{aligned} C(P_{SN}(h)) C(P_{SN}(g))^T &= C(P_{SN}(h)) C(\widetilde{P_{SN}(g)}) \\ &= C \left(\left(\sum_{i=0}^{SN-1} [P_{SN}(h)]_i [\widetilde{P_{SN}(g)}]_{j-i} \right)_{j=0..SN-1} \right) \\ &= C \left(\left(\sum_{i=0}^{SN-1} [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \right)_{j=0..SN-1} \right) . \end{aligned}$$

Setting

$$b^{(SN)}[h, g] = \left(\sum_{i=0}^{SN-1} [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \right)_{j=0..SN-1} , \quad (34)$$

we have

$$C(P_{SN}(h)) C(P_{SN}(g))^T = C(b^{(SN)}[h, g]) . \quad (35)$$

To simplify the forthcoming calculations, we temporarily denote by $b := b^{(SN)}[h, g]$. Notice that by definition, b is SN -periodic. From the definition of P_{SN} in (26), we have, for $i \in \llbracket -r, -r + SN - 1 \rrbracket$,

$$[P_{SN}(h)]_i = \begin{cases} h_{r-i} & \text{if } i \in \llbracket -r, r \rrbracket \\ 0 & \text{if } i \in \llbracket r+1, -r + SN - 1 \rrbracket. \end{cases} \quad (36)$$

Hence, since $P_{SN}(h)$ and $P_{SN}(g)$ are periodic, we have

$$\begin{aligned} b_j &= \sum_{i=0}^{SN-1} [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \\ &= \sum_{i=-r}^{SN-1-r} [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \\ &= \sum_{i=-r}^r [P_{SN}(h)]_i [P_{SN}(g)]_{i-j}. \end{aligned} \quad (37)$$

The signal b is SN -periodic. Therefore, we can restrict the study to an interval of size SN . We consider $j \in \llbracket -2r, SN - 2r - 1 \rrbracket$. For any j , the set of indices $i \in \llbracket -r, r \rrbracket$ such that $[P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \neq 0$ is included in $\llbracket -r, r \rrbracket \cap \{i \mid (i-j) \% SN \in \llbracket -r, r \rrbracket\}$.

Let $j \in \llbracket -2r, SN - 2r - 1 \rrbracket$: We have $-r \leq i \leq r$ and $-2r \leq j \leq SN - 2r - 1$, then $-SN + r + 1 \leq i - j \leq 3r$, but by hypothesis, $SN \geq 2k - 1 = 4r + 1$, hence $3r < SN - r$ and so $-SN + r < i - j < SN - r$. Therefore, for $i \in \llbracket -r, r \rrbracket$ and $j \in \llbracket -2r, SN - 2r - 1 \rrbracket$

$$(i-j) \% SN \in (\llbracket -r, r \rrbracket \% SN) \iff i-j \in \llbracket -r, r \rrbracket \iff i \in \llbracket -r+j, r+j \rrbracket.$$

Hence, for $j \in \llbracket -2r, SN - 2r - 1 \rrbracket$, the set of indices $i \in \llbracket -r, r \rrbracket$ such that $[P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \neq 0$ is included in $\llbracket -r, r \rrbracket \cap \llbracket -r+j, r+j \rrbracket$.

Using (10), we denote by

$$a = \text{conv}(h, g, \text{padding zero} = k-1, \text{stride} = 1) \in \mathbb{R}^{2k-1}.$$

We have from (11), for $j \in \llbracket 0, 2k-2 \rrbracket$,

$$a_j = \sum_{i=0}^{k-1} h_i \bar{g}_{i+j}.$$

Using (12) and keeping the indices $i \in \llbracket 0, k-1 \rrbracket$ for which $\bar{g}_{i+j} \neq 0$, i.e. such that $i+j \in \llbracket k-1, 2k-2 \rrbracket$, we obtain

$$\begin{cases} a_j = \sum_{i=k-1-j}^{k-1} h_i g_{i+j-(k-1)} & \text{if } j \in \llbracket 0, k-2 \rrbracket, \\ a_j = \sum_{i=0}^{2k-2-j} h_i g_{i+j-(k-1)} & \text{if } j \in \llbracket k-1, 2k-2 \rrbracket. \end{cases} \quad (38)$$

In the sequel, we will connect b with a by distinguishing several cases depending on the value of j . We distinguish $j \in \llbracket 0, 2r \rrbracket$, $j \in \llbracket -2r, -1 \rrbracket$ and $j \in \llbracket 2r+1, -2r + SN - 1 \rrbracket$. Recall that $k = 2r + 1$.

If $j \in \llbracket 0, 2r \rrbracket$: then $\llbracket -r, r \rrbracket \cap \llbracket -r+j, r+j \rrbracket = \llbracket -r+j, r \rrbracket$. Using (36), the equality (37) becomes

$$\begin{aligned} b_j &= \sum_{i=-r+j}^r [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \\ &= \sum_{i=-r+j}^r h_{r-i} g_{r-i+j}. \end{aligned}$$

By changing the variable $l = r - i$, and using $k = 2r + 1$, we find

$$\begin{aligned} b_j &= \sum_{l=0}^{2r-j} h_l g_{l+j} \\ &= \sum_{l=0}^{k-1-j} h_l g_{l+j} \\ &= \sum_{l=0}^{2k-2-(k-1+j)} h_l g_{l+(k-1+j)-(k-1)}. \end{aligned}$$

When $j \in \llbracket 0, 2r \rrbracket = \llbracket 0, k-1 \rrbracket$, we have $k-1+j \in \llbracket k-1, 2k-2 \rrbracket$, therefore using (38), we obtain

$$b_j = a_{k-1+j} . \quad (39)$$

If $j \in \llbracket -2r, -1 \rrbracket$: then $\llbracket -r, r \rrbracket \cap \llbracket -r+j, r+j \rrbracket = \llbracket -r, r+j \rrbracket$. Using (36), the equality (37) becomes

$$\begin{aligned} b_j &= \sum_{i=-r}^{r+j} [P_{SN}(h)]_i [P_{SN}(g)]_{i-j} \\ &= \sum_{i=-r}^{r+j} h_{r-i} g_{r-i+j} . \end{aligned}$$

By changing the variable $l = r - i$, and using $k = 2r + 1$, we find

$$\begin{aligned} b_j &= \sum_{l=-j}^{2r} h_l g_{l+j} \\ &= \sum_{l=-j}^{k-1} h_l g_{l+j} \\ &= \sum_{l=k-1-(k-1+j)}^{k-1} h_l g_{l+(k-1+j)-(k-1)} . \end{aligned}$$

When $j \in \llbracket -2r, -1 \rrbracket = \llbracket -(k-1), -1 \rrbracket$, we have $k-1+j \in \llbracket 0, k-2 \rrbracket$, and using (38), we obtain

$$b_j = a_{k-1+j} . \quad (40)$$

If $j \in \llbracket 2r+1, SN-2r-1 \rrbracket$: then $\llbracket -r, r \rrbracket \cap \llbracket -r+j, r+j \rrbracket = \emptyset$. The equality (37) becomes

$$b_j = 0. \quad (41)$$

Therefore, we summarize (39), (40) and (41): For all $j \in \llbracket -(k-1), -(k-1) + SN - 1 \rrbracket$,

$$b_j = \begin{cases} a_{k-1+j} & \text{if } j \in \llbracket -(k-1), k-1 \rrbracket, \\ 0 & \text{if } j \in \llbracket k, SN-k \rrbracket. \end{cases} \quad (42)$$

Recall that $P = \lfloor \frac{k-1}{S} \rfloor S \leq k-1$, and let $i \in \llbracket 0, 2P \rrbracket$. Therefore $i - P \in \llbracket -P, P \rrbracket \subset \llbracket -(k-1), k-1 \rrbracket$, hence using (13) and (42), we have

$$[\text{conv}(h, g, \text{padding zero} = P, \text{stride} = 1)]_i = a_{k-1+i-P} = b_{i-P} .$$

Therefore, using (14),

$$\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S) = \begin{bmatrix} b_{-\lfloor \frac{k-1}{S} \rfloor S} \\ \vdots \\ b_{-2S} \\ b_{-S} \\ b_0 \\ b_S \\ b_{2S} \\ \vdots \\ b_{\lfloor \frac{k-1}{S} \rfloor S} \end{bmatrix} \in \mathbb{R}^{2P/S+1} . \quad (43)$$

Using the definition of $Q_{S,N}$ in (33), we obtain

$$Q_{S,N}(\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S)) = \begin{bmatrix} b_0 \\ b_S \\ b_{2S} \\ \vdots \\ b_{\lfloor \frac{k-1}{S} \rfloor S} \\ 0 \\ \vdots \\ 0 \\ b_{- \lfloor \frac{k-1}{S} \rfloor S} \\ \vdots \\ b_{-2S} \\ b_{-S} \end{bmatrix} \in \mathbb{R}^N .$$

But, using (41), and since $\lfloor \frac{k-1}{S} \rfloor S$ is the largest multiple of S less than or equal to $k-1$, we have

$$S_N b = \begin{bmatrix} b_0 \\ b_S \\ b_{2S} \\ \vdots \\ b_{\lfloor \frac{k-1}{S} \rfloor S} \\ 0 \\ \vdots \\ 0 \\ b_{SN - \lfloor \frac{k-1}{S} \rfloor S} \\ \vdots \\ b_{SN-2S} \\ b_{SN-S} \end{bmatrix} = \begin{bmatrix} b_0 \\ b_S \\ b_{2S} \\ \vdots \\ b_{\lfloor \frac{k-1}{S} \rfloor S} \\ 0 \\ \vdots \\ 0 \\ b_{- \lfloor \frac{k-1}{S} \rfloor S} \\ \vdots \\ b_{-2S} \\ b_{-S} \end{bmatrix} \in \mathbb{R}^N .$$

Finally, we have

$$S_N b = Q_{S,N}(\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S)) .$$

Using (35) and Lemma 6, we conclude that

$$\begin{aligned} S_N C(P_{SN}(h)) C(P_{SN}(g))^T S_N^T &= S_N C(b^{(SN)}[h, g]) S_N^T \\ &= C(S_N b^{(SN)}[h, g]) \\ &= C(Q_{S,N}(\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S))) . \end{aligned}$$

□

Lemma 8. Let $M, C, S, k = 2r + 1$ be positive integers, and let $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$. Let N be such

that $SN \geq 2k - 1$, and $P = \lfloor \frac{k-1}{S} \rfloor S$. We denote by $z_{P/S} = \begin{bmatrix} 0_{P/S} \\ 1 \\ 0_{P/S} \end{bmatrix} \in \mathbb{R}^{2P/S+1}$. We have

$$\mathcal{K}\mathcal{K}^T - Id_{MN} = \begin{pmatrix} C(Q_{S,N}(x_{1,1})) & \dots & C(Q_{S,N}(x_{1,M})) \\ \vdots & \ddots & \vdots \\ C(Q_{S,N}(x_{M,1})) & \dots & C(Q_{S,N}(x_{M,M})) \end{pmatrix} ,$$

where for all $m, l \in \llbracket 1, M \rrbracket$,

$$x_{m,l} = \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S} \in \mathbb{R}^{2P/S+1} . \quad (44)$$

Proof. We have from (28),

$$\mathcal{K} = \begin{pmatrix} S_N C(P_{SN}(\mathbf{K}_{1,1})) & \dots & S_N C(P_{SN}(\mathbf{K}_{1,C})) \\ \vdots & \ddots & \vdots \\ S_N C(P_{SN}(\mathbf{K}_{M,1})) & \dots & S_N C(P_{SN}(\mathbf{K}_{M,C})) \end{pmatrix} \in \mathbb{R}^{MN \times CSN}.$$

Hence, we have that the block $(m, l) \in \llbracket 1, M \rrbracket^2$ of size (N, N) of $\mathcal{K}\mathcal{K}^T$ is equal to :

$$\begin{aligned} & \begin{pmatrix} S_N C(P_{SN}(\mathbf{K}_{m,1})) & \dots & S_N C(P_{SN}(\mathbf{K}_{m,C})) \end{pmatrix} \begin{pmatrix} C(P_{SN}(\mathbf{K}_{l,1}))^T S_N^T \\ \vdots \\ C(P_{SN}(\mathbf{K}_{l,C}))^T S_N^T \end{pmatrix} \\ &= \sum_{c=1}^C S_N C(P_{SN}(\mathbf{K}_{m,c})) C(P_{SN}(\mathbf{K}_{l,c}))^T S_N^T. \end{aligned}$$

We denote by $A_{m,l}$ the block $(m, l) \in \llbracket 1, M \rrbracket^2$ of size (N, N) of $\mathcal{K}\mathcal{K}^T - Id_{MN}$. We want to prove that $A_{m,l} = C(Q_{S,N}(x_{m,l}))$ where $x_{m,l}$ is defined in (44). Using (6), (20), and (33), we have $Id_N = C\left(\begin{bmatrix} 1 \\ 0_{N-1} \end{bmatrix}\right) = C(Q_{S,N}(z_{P/S}))$, and therefore,

$$A_{m,l} = \sum_{c=1}^C S_N C(P_{SN}(\mathbf{K}_{m,c})) C(P_{SN}(\mathbf{K}_{l,c}))^T S_N^T - \delta_{m=l} C(Q_{S,N}(z_{P/S})).$$

Using Lemma 7, this becomes

$$A_{m,l} = \sum_{c=1}^C C(Q_{S,N}(\text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S))) - \delta_{m=l} C(Q_{S,N}(z_{P/S})).$$

By linearity of C and $Q_{S,N}$, we obtain

$$A_{m,l} = C\left(Q_{S,N}\left(\sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S}\right)\right).$$

□

Proof of Theorem 2. Let $M, C, S, k = 2r + 1$ be positive integers, and let $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$. Let N be such that $SN \geq 2k - 1$, and $P = \lfloor \frac{k-1}{S} \rfloor S$. For all $m, l \in \llbracket 1, M \rrbracket$, we denote by $A_{m,l} \in \mathbb{R}^{N \times N}$ the block (m, l) of $\mathcal{K}\mathcal{K}^T - Id_{MN}$. Using Lemma 8, we have

$$A_{m,l} = C\left(Q_{S,N}\left(\sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S}\right)\right).$$

Hence, from (20) and (33), using the fact that for all $x \in \mathbb{R}^N$, $\|C(x)\|_F^2 = N\|x\|_2^2$, and for all $x \in \mathbb{R}^{2P/S+1}$, $\|Q_{S,N}(x)\|_2^2 = \|x\|_2^2$, we have

$$\begin{aligned} & \|\mathcal{K}\mathcal{K}^T - Id_{MN}\|_F^2 \\ &= \sum_{m=1}^M \sum_{l=1}^M \|A_{m,l}\|_F^2 \\ &= \sum_{m=1}^M \sum_{l=1}^M \left\| C\left(Q_{S,N}\left(\sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S}\right)\right) \right\|_F^2 \\ &= \sum_{m=1}^M \sum_{l=1}^M N \left\| Q_{S,N}\left(\sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S}\right) \right\|_2^2 \\ &= N \sum_{m=1}^M \sum_{l=1}^M \left\| \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S} \right\|_2^2. \end{aligned}$$

Therefore, using (16) and (15), we obtain for any $M, C, S, k = 2r + 1$ and $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$,

$$\|\mathcal{K}\mathcal{K}^T - Id_{MN}\|_F^2 = N \|\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2. \quad (45)$$

This concludes the proof in the RO case.

In order to prove the theorem in the CO case we use Lemma 1 in [41],

$$\|\mathcal{K}^T \mathcal{K} - Id_{CSN}\|_F^2 = \|\mathcal{K}\mathcal{K}^T - Id_{MN}\|_F^2 + CSN - MN.$$

Therefore, using that (45) holds for all M, C and S , we have

$$\|\mathcal{K}^T \mathcal{K} - Id_{CSN}\|_F^2 = N (\|\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2 - (M - CS)) \quad (46)$$

Hence, using the definitions of err_N^F and L_{orth} in Section A.2.1, (45) and (46) lead to

$$(\text{err}_N^F(\mathbf{K}))^2 = NL_{orth}(\mathbf{K}).$$

This concludes the proof. \square

D.2 Sketch of the proof of Theorem 2, for 2D convolutional layers

We start by stating intermediate lemmas. First we introduce a slight abuse of notation, for a vector $x \in \mathbb{R}^{N^2}$, we denote by $\mathcal{C}(x) = \mathcal{C}(X)$, where $X \in \mathbb{R}^{N \times N}$ such that $\text{Vect}(X) = x$.

Lemma 9. *Let $X \in \mathbb{R}^{SN \times SN}$. We have*

$$\mathcal{S}_N \mathcal{C}(X) \mathcal{S}_N^T = \mathcal{C}(\mathcal{S}_N \text{Vect}(X)).$$

Let $\mathcal{Q}_{S,N}$ be the operator which associates to a matrix $x \in \mathbb{R}^{(2P/S+1) \times (2P/S+1)}$ the matrix

$$\begin{pmatrix} x_{P/S,P/S} & \cdots & x_{P/S,2P/S} & 0 & \cdots & 0 & x_{P/S,0} & \cdots & x_{P/S,P/S-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{2P/S,P/S} & \cdots & x_{2P/S,2P/S} & 0 & \cdots & 0 & x_{2P/S,0} & \cdots & x_{2P/S,P/S-1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ x_{0,P/S} & \cdots & x_{0,2P/S} & 0 & \cdots & 0 & x_{0,0} & \cdots & x_{0,P/S-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{P/S-1,P/S} & \cdots & x_{P/S-1,2P/S} & 0 & \cdots & 0 & x_{P/S-1,0} & \cdots & x_{P/S-1,P/S-1} \end{pmatrix} \in \mathbb{R}^{N \times N}.$$

Lemma 10. *Let N be such that $SN \geq 2k - 1$, $h, g \in \mathbb{R}^{k \times k}$ and $P = \lfloor \frac{k-1}{S} \rfloor S$, we have*

$$\mathcal{S}_N \mathcal{C}(\mathcal{P}_{SN}(h)) \mathcal{C}(\mathcal{P}_{SN}(g))^T \mathcal{S}_N^T = \mathcal{C}(\mathcal{Q}_{S,N}(\text{conv}(h, g, \text{padding zero} = P, \text{stride} = S))).$$

Lemma 11. *Let $M, C, S, k = 2r + 1$ be positive integers, and let $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$. Let N be such that $SN \geq 2k - 1$, and $P = \lfloor \frac{k-1}{S} \rfloor S$. We set $z_{P/S,P/S} \in \mathbb{R}^{2P/S+1 \times 2P/S+1}$ such that for all $i, j \in \llbracket 0, 2P/S \rrbracket$, $[z_{P/S,P/S}]_{i,j} = \delta_{i=P/S} \delta_{j=P/S}$. We have*

$$\mathcal{K}\mathcal{K}^T - Id_{MN^2} = \begin{pmatrix} \mathcal{C}(\mathcal{Q}_{S,N}(x_{1,1})) & \cdots & \mathcal{C}(\mathcal{Q}_{S,N}(x_{1,M})) \\ \vdots & \ddots & \vdots \\ \mathcal{C}(\mathcal{Q}_{S,N}(x_{M,1})) & \cdots & \mathcal{C}(\mathcal{Q}_{S,N}(x_{M,M})) \end{pmatrix},$$

where for all $m, l \in \llbracket 1, M \rrbracket$,

$$x_{m,l} = \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S,P/S}.$$

Then we proceed as in the 1D case.

E Proof of Theorem 3

E.1 Proof of Theorem 3, for 1D convolutional layers

Let N be a positive integer such that $SN \geq k$.

RO case ($M \leq CS$): We denote by $P = \lfloor \frac{k-1}{S} \rfloor S$. From Lemma 8, we have

$$\mathcal{K}\mathcal{K}^T - Id_{MN} = \begin{pmatrix} C(Q_{S,N}(x_{1,1})) & \dots & C(Q_{S,N}(x_{1,M})) \\ \vdots & \ddots & \vdots \\ C(Q_{S,N}(x_{M,1})) & \dots & C(Q_{S,N}(x_{M,M})) \end{pmatrix}, \quad (47)$$

where for all $m, l \in \llbracket 1, M \rrbracket$,

$$x_{m,l} = \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S} \in \mathbb{R}^{2P/S+1}. \quad (48)$$

We set

$$B = \mathcal{K}\mathcal{K}^T - Id_{MN}.$$

Since B is symmetric and due to the well-known properties of matrix norms, we have $\|B\|_1 = \|B\|_\infty$ and $\|B\|_2^2 \leq \|B\|_1 \|B\|_\infty$. Hence, using the definition of $\|B\|_1$, we have

$$\begin{aligned} \|B\|_2^2 &\leq \|B\|_1 \|B\|_\infty \\ &= \|B\|_1^2 \\ &= \left(\max_{1 \leq l \leq MN} \sum_{m=1}^{MN} |B_{m,l}| \right)^2. \end{aligned}$$

Using (47), and (20), we obtain

$$\|B\|_2^2 \leq \max_{1 \leq l \leq M} \left(\sum_{m=1}^M \|Q_{S,N}(x_{m,l})\|_1 \right)^2.$$

Given the definition of $Q_{S,N}$ in (33), we have for all $x \in \mathbb{R}^{2P/S+1}$, $\|Q_{S,N}(x)\|_1 = \|x\|_1$, therefore,

$$\|B\|_2^2 \leq \max_{1 \leq l \leq M} \left(\sum_{m=1}^M \|x_{m,l}\|_1 \right)^2.$$

We set $l_0 = \arg \max_{1 \leq l \leq M} \left(\sum_{m=1}^M \|x_{m,l}\|_1 \right)^2$. Using that for all $x \in \mathbb{R}^n$, $\|x\|_1 \leq \sqrt{n}\|x\|_2$, we have

$$\begin{aligned} \|B\|_2^2 &\leq \left(\sum_{m=1}^M \|x_{m,l_0}\|_1 \right)^2 \\ &\leq (2P/S + 1) \left(\sum_{m=1}^M \|x_{m,l_0}\|_2 \right)^2. \end{aligned}$$

Using Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|B\|_2^2 &\leq (2P/S + 1) M \sum_{m=1}^M \|x_{m,l_0}\|_2^2 \\ &\leq (2P/S + 1) M \sum_{m=1}^M \sum_{l=1}^M \|x_{m,l}\|_2^2. \end{aligned}$$

Using (48), then (16) and (15), we obtain

$$\begin{aligned}
& \|B\|_2^2 \\
& \leq (2P/S + 1)M \sum_{m=1}^M \sum_{l=1}^M \left\| \sum_{c=1}^C \text{conv}(\mathbf{K}_{m,c}, \mathbf{K}_{l,c}, \text{padding zero} = P, \text{stride} = S) - \delta_{m=l} z_{P/S} \right\|_2^2 \\
& = (2P/S + 1)M \sum_{m=1}^M \sum_{l=1}^M \left\| [\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}]_{m,l,:} \right\|_2^2 \\
& = (2P/S + 1)M \|\text{conv}(\mathbf{K}, \mathbf{K}, \text{padding zero} = P, \text{stride} = S) - I_{r0}\|_F^2 \\
& = (2P/S + 1)ML_{\text{orth}}(\mathbf{K}).
\end{aligned}$$

This proves the inequality in the RO case.

CO case ($M \geq CS$): First, for $n \geq 2k - 1$, let R_n be the operator that associates to $x \in \mathbb{R}^{2k-1}$, the vector

$$R_n(x) = (x_{k-1}, \dots, x_{2k-2}, 0, \dots, 0, x_0, \dots, x_{k-2})^T \in \mathbb{R}^n. \quad (49)$$

Note that, when $S' = 1$, $N' = SN$, we have in (33), $P' = k - 1$ and

$$Q_{1,SN} = R_{SN}. \quad (50)$$

Recall from (6) that $(f_i)_{i=0..SN-1}$ is the canonical basis of \mathbb{R}^{SN} . Let

$$\Lambda_j = C(f_j) \in \mathbb{R}^{SN \times SN}$$

be the permutation matrix which shifts down (cyclically) any vector by $j \in \llbracket 0, SN - 1 \rrbracket$: for all $x \in \mathbb{R}^{SN}$, for $i \in \llbracket 0, SN - 1 \rrbracket$, $(\Lambda_j x)_i = x_{(i-j)\%SN}$. Note that, using (20), we have for all $x \in \mathbb{R}^{SN}$,

$$[C(x)]_{:,j} = \Lambda_j x. \quad (51)$$

Recall that $k = 2r + 1$, and for all $h \in \mathbb{R}^k$,

$$P_{SN}(x) = (h_r, \dots, h_0, 0, \dots, 0, h_{2r}, \dots, h_{r+1})^T \in \mathbb{R}^{SN}.$$

For $j \in \llbracket 0, SN - 1 \rrbracket$, for $x \in \mathbb{R}^k$, we denote by

$$P_{SN}^{(j)}(x) = \Lambda_j P_{SN}(x) \quad (52)$$

and for $x \in \mathbb{R}^{2k-1}$, we denote by

$$R_{SN}^{(j)}(x) = \Lambda_j R_{SN}(x). \quad (53)$$

By assumption $SN \geq 2k - 1$, hence $R_{SN}(x)$ is well-defined and we have for all $j \in \llbracket 0, SN - 1 \rrbracket$, for all $x \in \mathbb{R}^{2k-1}$,

$$\begin{cases} \|R_{SN}^{(j)}(x)\|_1 = \|x\|_1, \\ \|R_{SN}^{(j)}(x)\|_2 = \|x\|_2. \end{cases} \quad (54)$$

We first start by introducing the following Lemma.

Lemma 12. *Let $h, g \in \mathbb{R}^k$. There exist S vectors $x_0, \dots, x_{S-1} \in \mathbb{R}^{2k-1}$ such that for all N satisfying $SN \geq 2k - 1$, we have for all $j \in \llbracket 0, SN - 1 \rrbracket$,*

$$[C(P_{SN}(h))^T S_N^T S_N C(P_{SN}(g))]_{:,j} = R_{SN}^{(j)}(x_{j\%S}).$$

Proof. Recall that from (17) and (19), we have $S_N = \sum_{i=0}^{N-1} E_{i,Si}$ and $A_N := S_N^T S_N = \sum_{i=0}^{N-1} F_{Si,Si}$. When applied to a vector $x \in \mathbb{R}^{SN}$, A_N keeps the elements of x with index multiple of S unchanged, while the other elements of Ax equal to zero. We know from (51) and (52) that, for $j \in \llbracket 0, SN - 1 \rrbracket$, the j -th column of $C(P_{SN}(g))$ is equal to $P_{SN}^{(j)}(g)$. Therefore, when applying

A_N , this becomes $A_N P_{SN}^{(j)}(g) = P_{SN}^{(j)}(g^j)$, where $g^j \in \mathbb{R}^k$ is formed from g by putting zeroes in the place of the elements that have been replaced by 0 when applying A_N . But since A_N keeps the elements of index multiple of S , we have that the j -th column of $A_N C(P_{SN}(g))$ has the same elements as its $j\%S$ -th column, shifted down $(j - j\%S)$ indices. This implies that $g^j = g^{j\%S}$. Note that we can also derive the exact formula of g^j , in fact for all $i \in \llbracket 0, 2r \rrbracket$,

$$[g^j]_i = \begin{cases} g_i & \text{if } (i - r - j)\%S = 0 \\ 0 & \text{otherwise.} \end{cases}$$

We again can see that $g^j = g^{j\%S}$. Therefore, $[A_N C(P_{SN}(g))]_{:,j} = P_{SN}^{(j)}(g^{j\%S})$. Using (51) and (52), this becomes $[A_N C(P_{SN}(g))]_{:,j} = [C(P_{SN}(g^{j\%S}))]_{:,j}$. Therefore, we have, for all $j \in \llbracket 0, SN - 1 \rrbracket$,

$$\begin{aligned} [C(P_{SN}(h))^T A_N C(P_{SN}(g))]_{:,j} &= C(P_{SN}(h))^T P_{SN}^{(j)}(g^{j\%S}) \\ &= \left[C(P_{SN}(h))^T C(P_{SN}(g^{j\%S})) \right]_{:,j}. \end{aligned}$$

Using the fact that the transpose of a circulant matrix is a circulant matrix and that two circulant matrices commute with each other (see (22) and (25)), we conclude that the transpose of any circulant matrix commutes with any circulant matrix, therefore

$$[C(P_{SN}(h))^T A_N C(P_{SN}(g))]_{:,j} = \left[C(P_{SN}(g^{j\%S})) C(P_{SN}(h))^T \right]_{:,j}.$$

Using Lemma 7 with $S' = 1$ and $N' = SN$, and noting that, when $S' = 1$, the sampling matrix $S_{N'}$ is equal to the identity, we have

$$\begin{aligned} &C(P_{SN}(g^{j\%S})) C(P_{SN}(h))^T \\ &= Id_{N'} C(P_{N'}(g^{j\%S})) C(P_{N'}(h))^T Id_{N'}^T \\ &= C(Q_{S',N'}(\text{conv}(g^{j\%S}, h, \text{padding zero} = \lfloor \frac{k-1}{S'} \rfloor S', \text{stride} = S'))) \\ &= C(Q_{1,SN}(\text{conv}(g^{j\%S}, h, \text{padding zero} = k-1, \text{stride} = 1))) \end{aligned}$$

To simplify, we denote by $x_{j\%S} = \text{conv}(g^{j\%S}, h, \text{padding zero} = k-1, \text{stride} = 1) \in \mathbb{R}^{2k-1}$. Using (50), we obtain

$$\begin{aligned} C(P_{SN}(g^{j\%S})) C(P_{SN}(h))^T &= C(Q_{1,SN}(x_{j\%S})) \\ &= C(R_{SN}(x_{j\%S})). \end{aligned}$$

Using (51) and (53), we obtain

$$\begin{aligned} [C(P_{SN}(h))^T A_N C(P_{SN}(g))]_{:,j} &= [C(R_{SN}(x_{j\%S}))]_{:,j} \\ &= \Lambda_j R_{SN}(x_{j\%S}) \\ &= R_{SN}^{(j)}(x_{j\%S}). \end{aligned}$$

Therefore, we have for all $j \in \llbracket 0, SN - 1 \rrbracket$,

$$[C(P_{SN}(h))^T S_N^T S_N C(P_{SN}(g))]_{:,j} = R_{SN}^{(j)}(x_{j\%S}).$$

This concludes the proof of the lemma. \square

Let us go back to the main proof.

Using (28), we have that the block $(c, c') \in \llbracket 1, C \rrbracket^2$ of size (SN, SN) of $\mathcal{K}^T \mathcal{K}$ is equal to :

$$\begin{aligned} &\left(C(P_{SN}(\mathbf{K}_{1,c}))^T S_N^T \quad \dots \quad C(P_{SN}(\mathbf{K}_{M,c}))^T S_N^T \right) \begin{pmatrix} S_N C(P_{SN}(\mathbf{K}_{1,c'})) \\ \vdots \\ S_N C(P_{SN}(\mathbf{K}_{M,c'})) \end{pmatrix} \\ &= \sum_{m=1}^M C(P_{SN}(\mathbf{K}_{m,c}))^T S_N^T S_N C(P_{SN}(\mathbf{K}_{m,c'})). \end{aligned} \tag{55}$$

For any $(m, c, c') \in \llbracket 1, M \rrbracket \times \llbracket 1, C \rrbracket^2$, we denote by $(x_{m,c,c',s})_{s=0..S-1} \in \mathbb{R}^{2k-1}$ the vectors obtained when applying Lemma 12 with $h = \mathbf{K}_{m,c}$, and $g = \mathbf{K}_{m,c'}$. Hence, we have, for all $j \in \llbracket 0, SN-1 \rrbracket$,

$$[C(P_{SN}(\mathbf{K}_{m,c}))^T S_N^T S_N C(P_{SN}(\mathbf{K}_{m,c'}))]_{:,j} = R_{SN}^{(j)}(x_{m,c,c',j\%S}). \quad (56)$$

Let $\bar{f}_{k-1} = \begin{bmatrix} 0_{k-1} \\ 1 \\ 0_{k-1} \end{bmatrix} \in \mathbb{R}^{2k-1}$. For all $s \in \llbracket 0, S-1 \rrbracket$, we denote by

$$x_{c,c',s} = \sum_{m=1}^M x_{m,c,c',s} - \delta_{c=c'} \bar{f}_{k-1} \in \mathbb{R}^{2k-1}. \quad (57)$$

Note that, from (6), (49), and (53), we have for all $j \in \llbracket 0, SN-1 \rrbracket$, $f_j = R_{SN}^{(j)}(\bar{f}_{k-1})$. Therefore, $Id_{SN} = (f_0, \dots, f_{SN-1}) = (R_{SN}^{(0)}(\bar{f}_{k-1}), \dots, R_{SN}^{(SN-1)}(\bar{f}_{k-1}))$. We denote by $A_{c,c'}^N \in \mathbb{R}^{SN \times SN}$ the block $(c, c') \in \llbracket 1, C \rrbracket^2$ of $\mathcal{K}^T \mathcal{K} - Id_{CSN}$. Using (55), (56), and (57), we have, for all $j \in \llbracket 0, SN-1 \rrbracket$,

$$\begin{aligned} [A_{c,c'}^N]_{:,j} &= \left[\sum_{m=1}^M C(P_{SN}(\mathbf{K}_{m,c}))^T S_N^T S_N C(P_{SN}(\mathbf{K}_{m,c'})) - \delta_{c=c'} Id_{SN} \right]_{:,j} \\ &= \sum_{m=1}^M R_{SN}^{(j)}(x_{m,c,c',j\%S}) - \delta_{c=c'} R_{SN}^{(j)}(\bar{f}_{k-1}) \\ &= R_{SN}^{(j)}(x_{c,c',j\%S}). \end{aligned} \quad (58)$$

We set

$$B_N = \mathcal{K}^T \mathcal{K} - Id_{CSN}.$$

We then proceed in the same way as in the RO case. Since B_N is clearly symmetric, we have

$$\begin{aligned} \|B_N\|_2^2 &\leq \|B_N\|_1 \|B_N\|_\infty \\ &= \|B_N\|_1^2 \\ &= \left(\max_{1 \leq j \leq CSN} \sum_{i=1}^{CSN} |(B_N)_{i,j}| \right)^2 \\ &= \max_{1 \leq c' \leq C, 0 \leq j \leq SN-1} \left(\sum_{c=1}^C \| [A_{c,c'}^N]_{:,j} \|_1 \right)^2. \end{aligned}$$

Using (58) and (54), this becomes

$$\begin{aligned} \|B_N\|_2^2 &\leq \max_{1 \leq c' \leq C, 0 \leq j \leq SN-1} \left(\sum_{c=1}^C \|R_{SN}^{(j)}(x_{c,c',j\%S})\|_1 \right)^2 \\ &= \max_{1 \leq c' \leq C, 0 \leq s \leq S-1} \left(\sum_{c=1}^C \|x_{c,c',s}\|_1 \right)^2. \end{aligned}$$

We set $(c'_0, s_0) = \arg \max_{1 \leq c' \leq C, 0 \leq s \leq S-1} \left(\sum_{c=1}^C \|x_{c,c',s}\|_1 \right)^2$. Using that for all $x \in \mathbb{R}^n$, $\|x\|_1 \leq \sqrt{n}\|x\|_2$, we have

$$\begin{aligned} \|B_N\|_2^2 &\leq \left(\sum_{c=1}^C \|x_{c,c'_0,s_0}\|_1 \right)^2 \\ &\leq (2k-1) \left(\sum_{c=1}^C \|x_{c,c'_0,s_0}\|_2 \right)^2. \end{aligned}$$

Using Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\|B_N\|_2^2 &\leq (2k-1)C \sum_{c=1}^C \|x_{c,c'_0,s_0}\|_2^2 \\ &\leq (2k-1)C \sum_{c=1}^C \sum_{c'=1}^C \sum_{s=0}^{S-1} \|x_{c,c',s}\|_2^2.\end{aligned}$$

Using (54) in the particular case of $N = 2k - 1$, we obtain

$$\begin{aligned}\|B_N\|_2^2 &\leq (2k-1)C \sum_{c=1}^C \sum_{c'=1}^C \sum_{s=0}^{S-1} \|R_{S(2k-1)}(x_{c,c',s})\|_2^2 \\ &= C \sum_{c=1}^C \sum_{c'=1}^C \sum_{s=0}^{S-1} (2k-1) \|R_{S(2k-1)}(x_{c,c',s})\|_2^2 \\ &= C \sum_{c=1}^C \sum_{c'=1}^C \sum_{j=0}^{S(2k-1)-1} \|R_{S(2k-1)}^{(j)}(x_{c,c',j \% S})\|_2^2.\end{aligned}$$

Using (58) for $N = 2k - 1$, we obtain

$$\begin{aligned}\|B_N\|_2^2 &\leq C \sum_{c=1}^C \sum_{c'=1}^C \sum_{j=0}^{S(2k-1)-1} \left\| \left[A_{c,c'}^{2k-1} \right]_{:,j} \right\|_2^2 \\ &= C \|B_{2k-1}\|_F^2.\end{aligned}$$

Using Theorem 2 for $N = 2k - 1$, we have $\|B_{2k-1}\|_F^2 = (2k-1)L_{orth}(\mathbf{K})$ and we obtain

$$\|B_N\|_2^2 \leq (2k-1)CL_{orth}(\mathbf{K}).$$

Therefore, we conclude that, in the CO case

$$(\text{err}_N^s(\mathbf{K}))^2 \leq (2k-1)CL_{orth}(\mathbf{K}).$$

This concludes the proof.

E.2 Sketch of the proof of Theorem 3, for 2D convolutional layers

In the RO case, we proceed as in the 1D case.

In the CO case, we first prove a lemma similar to Lemma 12, then we proceed as in the 1D case.

F Proof of Proposition 1

Below, we prove Proposition 1 for a general matrix $M \in \mathbb{R}^{a \times b}$. In order to obtain the statement for a convolutional layer $\mathcal{K} \in \mathbb{R}^{MN \times CSN}$:

In the RO case ($M \leq CS$): we take $M = \mathcal{K}^T$, $a = CSN$, $b = MN$.

In the CO case ($M \geq CS$): we take $M = \mathcal{K}$, $a = MN$, $b = CSN$.

Let $M \in \mathbb{R}^{a \times b}$ such that $a \geq b$. We denote by $\varepsilon = \|M^T M - Id_b\|_2$. Let $x \in \mathbb{R}^b$, we have

$$\begin{aligned}|\|Mx\|^2 - \|x\|^2| &= |x^T M^T M x - x^T x| = |x^T (M^T M - Id_b) x| \leq \|x^T\| \|M^T M - Id_b\|_2 \|x\| \\ &\leq \varepsilon \|x\|^2.\end{aligned}$$

Hence, for all $x \in \mathbb{R}^b$,

$$(1 - \varepsilon)\|x\|^2 \leq \|Mx\|^2 \leq (1 + \varepsilon)\|x\|^2.$$

This also implies $\sigma_{max}(M)^2 \leq 1 + \varepsilon$. But we know that $\sigma_{max}(M^T) = \sigma_{max}(M)$, hence $\sigma_{max}(M^T)^2 \leq 1 + \varepsilon$ and therefore, for all $x \in \mathbb{R}^a$,

$$\|M^T x\|^2 \leq (1 + \varepsilon)\|x\|^2.$$

This proves Proposition 1.

G Boundary conditions comparison

We consider in this section an unstrided convolution, and we state results about other paddings and their limitations.

Proposition 2. *Let $N \geq 2k - 1$. With the padding 'valid', there exist no orthogonal convolutional layer in the CO case.*

Note that this result still holds in the 2D case, we give a proof only in the 1D.

Proof. For a single-channel convolution of kernel $h \in \mathbb{R}^k$ with 'valid' padding, the matrix applying the transformation on a signal $x \in \mathbb{R}^N$ has the following form:

$$A_N(h) := \begin{pmatrix} h_0 & \cdots & h_{2r} & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & h_0 & \cdots & h_{2r} \end{pmatrix} \in \mathbb{R}^{(N-k+1) \times N}.$$

Hence, for $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, the layer transform matrix is:

$$\mathcal{K} = \begin{pmatrix} A_N(\mathbf{K}_{1,1}) & \cdots & A_N(\mathbf{K}_{1,C}) \\ \vdots & \vdots & \vdots \\ A_N(\mathbf{K}_{M,1}) & \cdots & A_N(\mathbf{K}_{M,C}) \end{pmatrix} \in \mathbb{R}^{M(N-k+1) \times CN}.$$

Suppose that $\mathcal{K}^T \mathcal{K} = Id_{CN}$, therefore the columns of \mathcal{K} are of norm 1. Let us focus on the columns corresponding to the first channel. To simplify, for $m \in \llbracket 1, M \rrbracket$ we denote by $a^{(m)} := \mathbf{K}_{m,1} \in \mathbb{R}^k$, hence

$$\begin{pmatrix} A_N(\mathbf{K}_{1,1}) \\ \vdots \\ A_N(\mathbf{K}_{M,1}) \end{pmatrix} = \begin{pmatrix} a_0^{(1)} & \cdots & a_{2r}^{(1)} & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & a_0^{(1)} & \cdots & a_{2r}^{(1)} \\ a_0^{(2)} & \cdots & a_{2r}^{(2)} & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & a_0^{(2)} & \cdots & a_{2r}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_0^{(M)} & \cdots & a_{2r}^{(M)} & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & a_0^{(M)} & \cdots & a_{2r}^{(M)} \end{pmatrix}.$$

Since $\mathcal{K}^T \mathcal{K} = Id_{CN}$, we have in particular

$$\begin{cases} \sum_{m=1}^M \left(a_0^{(m)}\right)^2 = 1 \\ \sum_{m=1}^M \left(a_{2r}^{(m)}\right)^2 = 1 \\ \sum_{i=0}^{2r} \sum_{m=1}^M \left(a_i^{(m)}\right)^2 = 1. \end{cases}$$

This is impossible. Therefore, for all $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, we have $\mathcal{K}^T \mathcal{K} \neq Id_{CN}$. □

Recall that $k = 2r + 1$, and that $(e_i)_{i=0..k-1}$ is the canonical basis of \mathbb{R}^k (see (6)). For the zero padding 'same', we have the following lemma.

Proposition 3. Let $N \geq k$. For $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, with the zero padding 'same' and $S = 1$, both in the RO and CO case, if \mathcal{K} is orthogonal then there exist $(\alpha_{m,c})_{m=1..M, c=1..C}$ such that for all $(m, c) \in \llbracket 1, M \rrbracket \times \llbracket 1, C \rrbracket$, $\mathbf{K}_{m,c} = \alpha_{m,c} e_r$. This is equivalent to

$$\mathcal{K} = \begin{pmatrix} \alpha_{1,1} Id_N & \dots & \alpha_{1,C} Id_N \\ \vdots & \vdots & \vdots \\ \alpha_{M,1} Id_N & \dots & \alpha_{M,C} Id_N \end{pmatrix} \in \mathbb{R}^{MN \times CN}.$$

Note that this result still holds in the 2D case.

Proof. For a single-channel convolution of kernel $h \in \mathbb{R}^k$ with zero padding 'same', the matrix applying the transformation on a signal $x \in \mathbb{R}^N$ has the following form:

$$A_N(h) := \begin{pmatrix} h_r & \dots & h_{2r} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & h_{2r} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_0 & \dots & h_r \end{pmatrix} \in \mathbb{R}^{N \times N}.$$

Hence, for $\mathbf{K} \in \mathbb{R}^{M \times C \times k}$, the matrix that applies the convolutional layer is :

$$\mathcal{K} = \begin{pmatrix} A_N(\mathbf{K}_{1,1}) & \dots & A_N(\mathbf{K}_{1,C}) \\ \vdots & \vdots & \vdots \\ A_N(\mathbf{K}_{M,1}) & \dots & A_N(\mathbf{K}_{M,C}) \end{pmatrix} \in \mathbb{R}^{MN \times CN}.$$

Suppose $M \leq C$ (RO case): If \mathcal{K} is orthogonal, then $\mathcal{K}\mathcal{K}^T = Id_{MN}$. Let us fix $m \in \llbracket 1, M \rrbracket$. Since $\mathcal{K}\mathcal{K}^T = Id_{MN}$, we have, the first row, the last row and the row of index r of the m -th block matrix of size $N \times CN$ are of norm equal to 1, i.e.

$$\begin{cases} \|\mathcal{K}_{(m-1)N,:}\|_2^2 = 1 \\ \|\mathcal{K}_{mN-1,:}\|_2^2 = 1 \\ \|\mathcal{K}_{(m-1)N+r,:}\|_2^2 = 1. \end{cases}$$

To simplify the notation, for $c \in \llbracket 1, C \rrbracket$, we denote by $a^{(c)} := \mathbf{K}_{m,c} \in \mathbb{R}^k$, hence the previous system is equivalent to

$$\begin{cases} \sum_{i=r}^{2r} \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 1 \\ \sum_{i=0}^r \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 1 \\ \sum_{i=0}^{2r} \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 1. \end{cases}$$

Subtracting the first equality from the third one, and the second equality from the third one, we obtain

$$\begin{cases} \sum_{i=0}^{r-1} \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 0 \\ \sum_{i=r+1}^{2r} \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 0 \\ \sum_{i=0}^{2r} \sum_{c=1}^C \left(a_i^{(c)}\right)^2 = 1. \end{cases}$$

This implies that for all $c \in \llbracket 1, C \rrbracket$, for all $i \in \llbracket 0, 2r \rrbracket \setminus \{r\}$, $a_i^{(c)} = 0$.

As a conclusion, for any $m \in \llbracket 1, M \rrbracket$, any $c \in \llbracket 1, C \rrbracket$, and any $i \in \llbracket 0, 2r \rrbracket \setminus \{r\}$, we have

$$\mathbf{K}_{m,c,i} = 0.$$

This proves the result in the RO case.

The proof of the CO case is similar, and we have the same conclusion. \square

To recapitulate, we have proven that with padding 'valid', no orthogonal convolution can be built in the CO case, and that for zero padding 'same', the orthogonal convolutions lead to trivial transformation.

H Computing the singular values of \mathcal{K}

In this appendix, we describe methods for computing singular values of a 2D layer transform matrix, with or without stride. The codes are provided in *ANONYMIZED* library.

H.1 Computing the singular values of \mathcal{K} when $S = 1$

For convolutional layers without stride, $S = 1$, we use the algorithm described in [33]. We describe the algorithm for 2D convolutional layers in Algorithm 1. The algorithm provides the full list of singular values.

Algorithm 1 Computing the list of singular values of \mathcal{K} , when $S = 1$, [33].

Input: kernel tensor: $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$, channel size: $N \geq k$

Output: list of the singular values of \mathcal{K} : σ

```
transforms = FFT2(K, (N,N), axes=[0, 1])
σ = linalg.svd(transforms, compute_uv=False)
```

H.2 Computing the smallest and the largest singular value of \mathcal{K} for any stride S

For convolutions with stride, $S > 1$, there is no known practical algorithm to compute the list of singular values σ . In this configuration, we use the well known power iteration algorithm and a spectral shift to compute the smallest and the largest singular value $(\sigma_{min}, \sigma_{max})$ of \mathcal{K} . We give the principle of the algorithm in Algorithm 2. For clarity, in Algorithm 2, we assume a function ' $\sigma = \text{power_iteration}(M)$ ', that applies the power iteration algorithm to a square matrix M and returns its largest singular value $\sigma \geq 0$. In practice, of course, we cannot construct M and the implementation must use the usual functions that apply \mathcal{K} and \mathcal{K}^T . A detailed python implementation is provided in *ANONYMIZED* library.

Algorithm 2 Computing $(\sigma_{min}, \sigma_{max})$, for any $S \geq 1$.

Input: kernel tensor: $\mathbf{K} \in \mathbb{R}^{M \times C \times k \times k}$, channel size: $N \geq k$, stride parameter: $S \geq 1$

Output: the smallest and the largest singular value of \mathcal{K} : $(\sigma_{min}, \sigma_{max})$

```
if CS2 ≥ M then
    # RO case
    σmax = sqrt( power_iteration( K KT ) )
    λ = 1.1 * σmax * σmax
    σmin = sqrt( λ - power_iteration( λ IdMN2 - K KT ) )
else
    # CO case
    σmax = sqrt( power_iteration( KT K ) )
    λ = 1.1 * σmax * σmax
    σmin = sqrt( λ - power_iteration( λ IdCS2N2 - KT K ) )
end if
```
