Fourier-based and Rational Graph Filters for Spectral Processing

Giuseppe Patané

Abstract—Data are represented as graphs in a wide range of applications, such as Computer Vision (e.g., images) and Graphics (e.g., 3D meshes), network analysis (e.g., social networks), and bio-informatics (e.g., molecules). In this context, our overall goal is the definition of novel Fourier-based and graph filters induced by rational polynomials for graph processing, which generalise polynomial filters and the Fourier transform to non-Euclidean domains. For the efficient evaluation of discrete spectral Fourier-based and wavelet operators, we introduce a spectrum-free approach, which requires the solution of a small set of sparse, symmetric, well-conditioned linear systems and is oblivious of the evaluation of the Laplacian or kernel spectrum. Approximating arbitrary graph filters with rational polynomials provides a more accurate and numerically stable alternative with respect to polynomials. To achieve these goals, we also study the link between spectral operators, wavelets, and filtered convolution with integral operators induced by spectral kernels. According to our tests, main advantages of the proposed approach are (i) its generality with respect to the input data (e.g., graphs, 3D shapes), applications (e.g., signal reconstruction and smoothing, shape correspondence), and filters (e.g., polynomial, rational polynomial), and (ii) a spectrum-free computation with a generally low computational cost and storage overhead.

Index Terms—Laplacian spectrum, kernels, graphs, spectral graph processing, frequency filtering, graph Fourier transform, heat kernel, Chebyshev rational polynomials

1 INTRODUCTION

Data are represented as graphs in a wide range of applications, such as Computer Vision (e.g., images) and Graphics (e.g., 3D meshes), network analysis (e.g., social networks), and bio-informatics (e.g., molecules). Spectral graph processing represents the input signal on a graph in terms of the eigenvectors of a graph operator (e.g., the graph Laplacian, a kernel matrix) in order to define its Fourier transform and convolution with another signal. According to this simple approach, whose discrete counterpart is based mainly on numerical linear algebra, spectral graph processing has been successfully applied to the characterisation of geometric and topological properties of graphs and to dimensionality reduction through the projection of the input data on low-dimensional subspaces generated by a small set of Laplacian or kernel eigenfunctions.

Signal processing on graphs also supports smooth signal interpolation and the definition of diffusion wavelets. Main applications of spectral processing are graph embedding through frequency analysis and random walks, data representation and classification. Furthermore, the modelling and training of convolutional neural networks, and the design of fast localised convolutional filters in high-dimensional irregular domains (e.g., graphs) have been recently addressed in the spectral domain. Finally, we recall the generalisation of spectral graph theory to 3D data through geometric deep learning on non-Euclidean domains.

In this context, our overall goal is the definition of novel Fourier-based and rational graph filters for graph processing. Firstly (Sect. 2), we show that the convolution of signals on a non-Euclidean space is uniquely determined by the linearity and commutativity of the translation operator with respect to convolution. Indeed, the definition of the convolution operator, which is commonly used in the context of spectral graph processing, is unique.

Through spectral filtering, we define spectral operators that map 1D filter functions to signals defined on an arbitrary domain. The spectral operator induces a special class of Fourier-based spectral operators, which generalise the notion of Fourier transform to non-Euclidean domains. In particular, we show that the main properties of the 1D Fourier transform, such as dilation, translation, scaling, derivation, localisation, and Parseval’s equality, still apply to a signal defined on an arbitrary domain. Combining the spectral operator with convolution, we introduce the filtered convolution operator, which is used to show the link between spectral operators, wavelets, and filtered convolution with integral operators induced by spectral kernels (Sect. 3).

The filtered convolution operator reduces to well-known Laplacian spectral operators (e.g., harmonic, bi-harmonic, diffusion, wave operators) for specific filters.

To efficiently compute spectral operators, it is necessary to apply fast polynomial approximations that are oblivious of the evaluation of the Laplacian spectrum, which is computationally expensive and numerically unstable. Main examples include polynomials (e.g., Chebyshev polynomials) and algebraic filters, which are evaluated through recursive relations and without diagonalising the Laplacian matrix associated with the input data. To overcome the time-consuming computation of the filter coefficients in terms of the Chebyshev polynomials through the evaluation of
integrals, we approximate the $L^2(\mathbb{R})$ scalar product with a discrete scalar product defined in terms of the Chebyshev nodes. In this way, we guarantee a high approximation accuracy and a linear computational cost for the evaluation of the Chebyshev coefficients of the input filter.

Generalising these results, we propose a novel class of Laplacian spectral wavelets induced by rational polynomial filters, whose evaluation is recursively expressed in terms of Chebyshev rational polynomials and requires the solution of a set of Laplace equations (Sect. 4). Rational polynomials are then applied to approximate arbitrary filters within a given tolerance, thus providing a more accurate and numerically stable alternative with respect to polynomials. In fact, rational polynomials are a richer class of functions with respect to polynomials, improve the approximation accuracy of polynomials, are more stable with respect to oscillations, as the errors in the numerator and denominator compensate each other [21], [22]. Furthermore, rational polynomials have been computed analytically for filters (e.g., sin/cos, exponential, logarithm) commonly used in spectral graph processing.

The definition of the spectral, filtered convolution, and Fourier-based wavelets through the filtering of the Laplacian spectrum faces a high computational cost for the evaluation of the spectrum in case of large graphs and numerical instabilities, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing. Even though multiple or close eigenvalues are quite common in numerical stability, which are associated with multiple or close eigenvalues for spectral graph processing.

2 Spectral and Fourier-based operators

For the definition of spectral operators, we introduce the Laplace-Beltrami operator and its spectrum, which provide a generalisation of the Fourier basis to non-Euclidean domains. Let $\mathcal{M}$ be an input domain (e.g., a manifold, a graph) and $\mathcal{F}(\mathcal{M})$ the space of signals defined on $\mathcal{M}$ (e.g., the space $L^2(\mathcal{M})$ of square integrable functions or the space $C^0(\mathcal{M})$ of continuous functions) equipped with the inner product $(f, g) := \int_\mathcal{M} f g d\mu$ and the corresponding norm $\|f\|$.

The Laplace-Beltrami operator $\Delta$ is self-adjoint, positive semidefinite, and admits the Laplacian orthonormal eigensystem ($\lambda_n, \phi_n$)$_{n=0}^\infty$, $\Delta \phi_n = \lambda_n \phi_n$, $\lambda_0 = 0$, $\lambda_n \leq \lambda_{n+1}$, in $L^2(\mathcal{M})$.

We show that the convolution of signals defined on a non-Euclidean space is uniquely determined by the linearity and commutativity of the translation operator with respect to convolution (Sect. 2.1). Then, we define a spectral operator that extend 1D filter functions to signals through the filtering of the spectrum and generalises well-known Laplacian spectral operators, such as the harmonic, bi-harmonic, and diffusion operators (Sect. 2.2). Considering the continuous Fourier transform of 1D filters, the spectral operator induces a special class of Fourier-based spectral operators that extend the notion of Fourier transform to non-Euclidean domains (Sect. 2.3). Finally, we show that the main properties of the 1D Fourier transform, such as dilation, translation, scaling, Fourier transform, derivation, localisation, and Parseval’s equality, still apply to the Fourier-based operator.

2.1 Convolution operator

Applying the linearity and commutativity with respect to convolution, we show that the spectral representation of the convolution operator, which is commonly used in the context of spectral graph processing, is unique. According to the continuous definition of the convolution operator

$$\begin{align*}
(f \ast g)(p) &= \langle f, T_p g \rangle, \\
\langle f, T_p q \rangle &= \langle f, T_q p \rangle, \quad \forall p, q \in \mathcal{M}.
\end{align*}
$$

in terms of the $L^2(\mathcal{M})$ scalar product and of the translation operator $T_p$, the convolution operator is uniquely defined by the translation operator. Indeed, it is enough to derive the spectral representation of the translation operator from its properties (i.e., linearity, commutativity). Noting that the translation operator is commutative with respect to convolution (i.e., $T_p f \ast g = f \ast T_p g$), and applying the definition, we get that

$$\begin{align*}
\langle T_p f, T_q g \rangle &= \langle f, T_p g \rangle, \quad \forall p, q \in \mathcal{M}.
\end{align*}
$$

Expressing the functions $f$, $T_p f$, and $T_q g$ in terms of the eigenfunctions ($\phi_n$)$_{n=0}^\infty$ as

$$f = \sum_{n=0}^{\infty} \langle f, \phi_n \rangle \phi_n, \quad T_p f = \sum_{n=0}^{\infty} \langle T_p f, \phi_n \rangle \phi_n, \quad T_q g = \sum_{n=0}^{\infty} \langle T_q g, \phi_n \rangle \phi_n.$$

2 Integrals of the rational approximation of spectral wavelets with respect to spectral approximations (e.g., low pass filters).

Finally, rational filters, and in particular rational Chebyshev polynomials, are particularly useful to enlarge the class of learning networks, as a generalisation of networks based on polynomial filters (e.g., PolyNet [23], ChebNet [24], CayleyNet [25]) in order to improve the discriminative capabilities of networks for 3D geometric deep learning.
Eq. (2) is equivalent to
\[
\langle T_p f, \phi_n \rangle_2 \langle T_q g, \phi_n \rangle_2 = \langle f, \phi_n \rangle_2 \langle T_q g, \phi_n \rangle_2. \tag{3}
\]
Choosing \( f = g = \phi_m \), the previous relation reduces to
\[
\langle T_p \phi_m, \phi_n \rangle_2 = \delta_{mn}, \langle T_q \phi_m, \phi_n \rangle_2 = \delta_{mn}. \tag{4}
\]

The resulting identity \( \langle T_p \phi_m, \phi_n \rangle_2 = 0, m \neq n \), implies that \( T_p \phi_m = \alpha(p) \phi_n, \forall p \in M \). Selecting \( f = \phi_m, g = \delta_p, m = n \) in Eq. (5), we obtain that \( \alpha(p) \phi_n(q) = \alpha(q) \phi_n(p) \), i.e., \( \alpha = \phi_n \). Indeed, the action of the translation operator on the eigenfunctions is defined as \( T_p \phi_n = \alpha(p) \phi_n \) and is extended to an arbitrary function \( f \) by linearity as
\[
T_p f = \sum_{n=0}^{+\infty} \langle f, \phi_n \rangle_2 T_p \phi_n = \sum_{n=0}^{+\infty} \hat{f}(n) \phi_n(p) \phi_n(p),
\]
with \( \hat{f}(n) := \langle f, \phi_n \rangle_2 \) \( n \)-th Fourier coefficient of \( f \). Applying the spectral representation of the translation operator to Eq. (1), we obtain the identity
\[
(f * g)(p) = \langle f, T_p g \rangle_2 = \sum_{n=0}^{+\infty} \hat{f}(n) \hat{g}(n) \phi_n(p). \tag{6}
\]

### 2.2 Spectral and filtered convolution operators

Firstly, we introduce the definition and main properties of the spectral operator (Sect. 2.2.1), which is applied to interpret filtered operators as convolution operators (Sect. 2.2.2) and to extend the Fourier transform of 1D functions to signals defined on non-Euclidean domains (Sect. 2.3).

#### 2.2.1 Spectral operator

Let \( H := L^2(\mathbb{R}) \cap C^0(\mathbb{R}) \) be the filter space and \( \varphi : \mathbb{R} \to \mathbb{R} \) be a positive, continuous, and square integrable filter in \( H \). Then, we define the spectral operator
\[
\Phi : H \to \mathcal{F}(\mathcal{M}), \quad \varphi \mapsto \Phi_\varphi := \sum_{n=0}^{+\infty} \varphi(\lambda_n) \phi_n, \tag{7}
\]
where \( \Phi_\varphi \) is the spectral function induced by \( \varphi \). The spectral operator is linear and continuous, according to the following upper bound
\[
\|\Phi_\varphi\|_2^2 = \sum_{n=0}^{+\infty} |\varphi(\lambda_n)|^2 \leq \int_0^{+\infty} |\varphi(s)|^2 ds = \|\varphi\|_2^2. \tag{8}
\]

#### 2.2.2 Filtered convolution operator

We now show the link between spectral and filtered convolution operators with filtered Laplacian operators and integral operators induced by the spectral kernels. Introducing the filtered convolution operator \( \Psi_\varphi : \mathcal{F}(\mathcal{M}) \to \mathcal{F}(\mathcal{M}) \)
\[
\Psi_\varphi f := \Phi_\varphi \ast f = \sum_{n=0}^{+\infty} \varphi(\lambda_n) \langle f, \phi_n \rangle_2 \phi_n = \Phi_{\varphi_1} = \Psi_{\varphi_1} f, \tag{9}
\]
the continuity of the filter function allows us to evaluate its values \( (\varphi(\lambda_n))_{n=0}^{\infty} \) at the eigenvalues and the \( L^2(\mathbb{R}) \)-integrability of the filter ensures the well-posedness of the spectral operator. The properties of the spectral operator depends only on the behaviour of the filter in the spectral domain and on the Laplacian/kernel spectrum; increasing or decreasing the filter decay to zero encodes global or local properties of the input graph, respectively.

#### Scalar product and norm of spectral operators

The \( L^2(\mathcal{M}) \) scalar product between the spectral functions \( \Phi_{\varphi_1}, \Phi_{\varphi_2} \) reduces to the \( \ell_2 \) scalar product of the filtered eigenvalues and is bounded by the \( L^2(\mathbb{R}) \) scalar product of the corresponding filters \( \varphi_1, \varphi_2 \), i.e.,
\[
\langle \Phi_{\varphi_1}, \Phi_{\varphi_2} \rangle_2 = \sum_{n=0}^{+\infty} \varphi_1(\lambda_n) \varphi_2(\lambda_n) = \langle (\varphi_1(\lambda_n))_{n=0}^{+\infty}, (\varphi_2(\lambda_n))_{n=0}^{+\infty} \rangle_2 \leq \int_0^{+\infty} \varphi_1(s) \varphi_2(s) ds \leq \langle \varphi_1, \varphi_2 \rangle_2, \quad (\varphi_1, \varphi_2 \geq 0). \tag{10}
\]

The \( L^2(\mathcal{M}) \)-distance of the spectral functions is bounded by \( L^2(\mathbb{R}) \)-distance of the corresponding filter functions, i.e.,
\[
\|\Phi_{\varphi_1} - \Phi_{\varphi_2}\|_2 = \|\Phi_{\varphi_1} - \Phi_{\varphi_2}\|_2 = \|\varphi_1 - \varphi_2\|_2. \tag{11}
\]
Indeed, the approximation of \( \Phi_\varphi \) reduces to the approximation of the 1D filter \( \varphi \); this remark will be applied to the spectrum-free computation of spectral and wavelet operators with rational polynomials (Sect. 3).
with \( \varphi_1 := \varphi \hat{f} \), the convolution between the spectral function \( \Phi_\varphi \) and a function \( f \) is induced by the pointwise product \( \varphi \hat{f} \) between the filter \( \varphi \) and the Fourier transform of \( f \). In particular, the convolution of the spectral functions \( \Phi_{\varphi_1}, \Phi_{\varphi_2} \) is induced by their pointwise product \( \varphi_1 \varphi_2 \), i.e., \( \Phi_{\varphi_1} \star \Phi_{\varphi_2} = \Phi_{\varphi_1 \varphi_2} \). Assuming that the input filter admits the power-series representation \( \varphi(s) := \sum_{n=0}^{\infty} \alpha_n s^n \), we define the filtered Laplacian operator \( \varphi(\Delta) := \sum_{n=0}^{\infty} \alpha_n \Delta_n \), which is equal to the filtered convolution operator \( \varphi_\Delta := \varphi(\Delta) \); in fact, \( \varphi(\Delta) f = \sum_{n=0}^{\infty} \varphi(\lambda_n)f(\lambda_n)\Phi_{\lambda_n}f = \varphi_\Delta f \).

Defining \( K_\varphi(p,q) := \Psi_\varphi \delta_n = \sum_{n=0}^{\infty} \varphi(\lambda_n)\phi_n(p)\phi_n(q) \), as spectral kernel and noting that

\[
\Psi_\varphi f(p) = \langle K_\varphi(p, \cdot), f \rangle_2,
\]

the filtered convolution operator is equal to the integral operator induced by the spectral kernel, i.e., \( \Psi_\varphi f = (K_\varphi, f)_2 \).

Fig. 3 shows different spectral kernels centred at a seed point on a 3D surface represented as a triangle mesh. Indeed, the input graph is arbitrary and the nodes have a 3D embedding, i.e., the coordinates of the mesh vertices. The spectral kernel well encodes the local geometry, as confirmed by the shape and distribution of the level-sets at all scales.

**Approximation of filtered convolution operators**

The norm of the filtered convolution operator \( \Psi_\varphi \) is bounded in terms of the \( L^2(\mathbb{R}) \) and \( L^\infty(\mathbb{R}) \) norm of the input filter function as \( \| \Psi_\varphi \| \leq \| \varphi \|_{k, 2}, k = 2, \infty \). In fact, we have that

\[
\| \Psi_\varphi f \|^2 = \sum_{n=0}^{\infty} |\varphi(\lambda_n)|^2 |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]

Choosing the signal \( f = \sum_n \phi_n \), we get \( \| \Psi_\varphi f \|_2 = \| \varphi \|_2 \), i.e., \( \| \Psi_\varphi \| = \| \varphi \|_2 \). From the equality (10), it follows that

\[
\| \Psi_\varphi f \|^2 \leq \| \varphi \|_2^2 \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle_2|^2 \leq \| f \|^2 \| \varphi \|_2^2.
\]
3 SPECTRAL WAVELETS & INTEGRAL OPERATORS

We derive equivalent representations of the spectral wavelet operator (Sect. 3.1), and its link with the spectral and filtered convolution operators and with existing kernels (Sect. 3.2).

3.1 Continuous spectral wavelet

We define the continuous spectral wavelet \( \Psi_{\phi, p} : \mathcal{M} \to \mathbb{R} \) centred at \( p \) and associated with the filter \( \phi : \mathbb{R} \to \mathbb{R} \) as

\[
\Psi_{\phi, p} := \Psi_{\phi} \delta_p = \sum_{n=0}^{+\infty} \phi(n) \phi_n(p) \phi_n.
\]

(13)

Indeed, the filtered convolution operator \( \psi \) generalises the spectral wavelet \( \Psi \); in fact, the spectral convolution operator \( \Psi \) is achieved as the action of the spectral operator on the \( \delta \)-function at \( p \). Applying the identity \( (\Psi f)(p) = (\Psi_{\phi} f)(p) \) and Eq. (3), the spectral convolution operator is also interpreted as the integral operator induced by the continuous spectral wavelet. Finally, the spectral wavelet \( \psi \) is continuous and its \( L^2 \)-energy is bounded by the \( L^2(\mathbb{R}) \) norm of the filter, i.e.,

\[
\|\Psi_{\psi, p}\|^2 = \sum_{n=0}^{+\infty} |\phi(n)|^2 |\phi_n(p)|^2 \leq \|\phi\|^2.
\]

(14)

Comparing wavelets induced by the same filter

The \( L^2(\mathcal{M}) \) scalar product of the continuous wavelets induced by \( \phi \) and centred at \( p, q \) is bounded by the \( L^2(\mathbb{R}) \)-norm of \( \psi \), i.e.,

\[
|\langle \Psi_{\phi, p}, \Psi_{\phi, q} \rangle | = \sum_{n=0}^{+\infty} |\phi(n)|^2 \int_{\mathcal{M} \times \mathcal{M}} \phi_n(p) \phi_n(q) dp dq = \sum_{n=0}^{+\infty} |\phi(n)|^2 \left( \int_{\mathcal{M}} \phi_n(p) dp \right)^2 \leq \sum_{n=0}^{+\infty} |\phi(n)|^2 \leq \|\phi\|^2.
\]

Comparing wavelets induced by different filters

The variation of the spectral wavelets induced by the filters \( \psi_1, \psi_2 \) and centred at \( p \) is bounded as

\[
|\psi_{\phi_1, p}(q) - \psi_{\phi_2, p}(q)| \leq \sum_{n=0}^{+\infty} |\phi_1(n) - \phi_2(n)| |\phi_n(p)| |\phi_n(q)|
\]

\[
\leq C_\phi C_p C_q, C_\psi := \|\psi_1 - \psi_2\|, C_p := \left[ \sum_{n=0}^{+\infty} |\phi_n(p)|^2 \right]^{1/2},
\]

where we have applied the Cauchy-Schwartz inequality in the last relation. Indeed, the maximum variation of the spectral wavelets \( \psi_{\phi_1, p}, \psi_{\phi_2, p} \) is bounded by the maximum variation \( C_\psi \) of the corresponding filters and by the values of the function \( C(\cdot) \) at the center \( p \) and at the evaluation point \( q \). Finally, the \( L^2(\mathcal{M}) \)-distance between the wavelets \( \psi_{\phi_1, p} \) and \( \psi_{\phi_2, p} \) is bounded as

\[
\left\|\psi_{\phi_1, p} - \psi_{\phi_2, p}\right\|^2 = \sum_{n=0}^{+\infty} |\phi_1(n) - \phi_2(n)|^2 |\phi_n(p)|^2 \leq C_\psi^2 \sum_{n=0}^{+\infty} |\phi_n(p)|^2,
\]

and its integral over \( \mathcal{M} \)

\[
\int_{\mathcal{M}} \left\|\psi_{\phi_1, p} - \psi_{\phi_2, p}\right\|^2 dp = \sum_{n=0}^{+\infty} |\phi_1(n) - \phi_2(n)|^2 \leq \|\phi_1 - \phi_2\|_2^2
\]

is bounded by the \( L^2(\mathbb{R}) \)-distance of the corresponding filters. Indeed, the distance of the filters determines the maximum distance of the corresponding spectral wavelets, and close filters generally correspond to close spectral wavelets. Finally, we notice the analogy of Eq. (15) with the upper bounds related to the approximation of the spectral (7) and filtered convolution (12) operators.

3.2 Examples of spectral wavelet operators

We report main examples of spectral wavelets (13), defined by properly selecting the filter function, thus showing their link with filtered convolution and spectral kernels.

\textbf{Commutator wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)

\textbf{Bi-harmonic wavelet/kernel} induced by the spectral operator \( \Psi = \partial^2 \) and the filter \( \psi(s) = \exp(-s^2) \), i.e.,

\[
\Psi_{\phi, p}(q) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(p) \phi_n(q)
\]

(5)
The Chebyshev polynomials of the first kind are orthogonal by the identity $\cos(n\theta) = T_n(\cos \theta)$, e.g., $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$, etc. Equivalently, these polynomials are defined through the recursive relation

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

(16)

The computation of the coefficients of the Chebyshev representation is generally time-consuming, as it requires to evaluate several integrals for an arbitrary filter. To overcome this limitation, we propose to approximate the weighted $L^2(\mathbb{R})$ scalar product through the Chebyshev nodes in order to accurately compute the Chebyshev coefficients in linear time.

$$a_0 = \frac{1}{N} \sum_{k=0}^{N-1} f(x_k), \quad a_i = \frac{2}{N} \sum_{k=0}^{N-1} f(x_k)T_i(x_k).$$

4.2 Rational spectrum-free computation

Noting that $\|\Psi_\varphi - \Psi_\rho\|^2 \leq \|\varphi - \rho\|_2$ (c.f., Eq. (14)), the approximation of $\Psi_\varphi$ is reduced to the computation of a function $\rho : \mathbb{R} \to \mathbb{R}$ that well approximates the filter $\varphi$ with respect to the $L^2(\mathbb{R})$-norm, i.e., thus solving 1D approximation problem. The representation of $\rho$ must guarantee that $\Psi_\rho$ can be efficiently computed through numerically robust algorithms, e.g., direct/iterative solvers of sparse, symmetric, and well-conditioned linear systems.

Rational approximation

According to the Weierstrass approximation theorem, polynomial and rational polynomial approximations allow us to approximate any continuous function on an interval within an arbitrary tolerance. Even though polynomials are easily evaluated at arbitrary values through the Horner’s method, and support an easy computation of derivatives and integrals, polynomial approximations tend to be oscillatory as their degree increases. Since rational polynomials are a richer class of functions with respect to polynomials, as novel contribution, we propose to apply rational filters, which generally improve the approximation accuracy of arbitrary filters with respect to polynomials and are more stable with respect to oscillations. To approximate the input filter with a polynomials, we consider a rational polynomial approximation $R(s) := P_n(s)/Q_m(s)$, where $P_n, Q_m$ are polynomials of degree $n$ and $m$, respectively. In particular, we consider rational polynomials $R(s)$ of the form

$$R(s) = \frac{P_n(s)}{Q_m(s)} = \frac{\sum_{i=0}^{n} p_is^i}{1 + \sum_{i=1}^{m} q_is^i},$$

(17)

of degree $N := n + m$. If $q_i = 0$, $i = 1, \ldots, m$, then the rational polynomial $R$ reduces to a polynomial $P_n$ of degree $n$. The rational approximation is a generalisation of the polynomial expansion (i.e., $Q_m := 1$), and more generally, of the Taylor series. As rational polynomials generate polynomials, we expect that rational approximations of degree $(n, m)$ are as good as polynomial approximations of degree $(m + n)$.

**Padé approximant**

The Padé approximant is the best approximation of a function with a rational polynomial such that the power series of the rational polynomial is the Taylor polynomial of the input function. Indeed, the Padé approximation $R(s)$ of order $(n, m)$ of $\varphi : \mathbb{R} \to \mathbb{R}$ in a neighbour of 0 is such that $R^{(k)}(0) = \varphi^{(k)}(0)$, $k = 0, \ldots, m + n$. In particular, $\varphi(s) - R^{(k)}(s) = O(s^{n+m+k+1})$, $s \to 0$. The Padé approximant is uniquely defined if the constant term at the denominator has been set equal to 1 (c.f., Eq. (17)); otherwise, the approximant is defined up to a multiplication by a constant. For the computation of the coefficients of the polynomials $P_n, Q_m$, we can apply the Wylm’s epsilon algorithm or the extended Euclidean algorithm [25], and the sequence transformation [27]. According [21, 22], a rational approximation is more stable than a polynomial approximation, as the errors in the numerator and denominator of a rational polynomial compensate each other. Furthermore, rational polynomial approximations have been computed
analytically for commonly used filters (e.g., \( \sin/\cos \), exponential, logarithm). For the approximation and evaluation of the rational polynomial, we consider (i) the canonical polynomial basis and the Chebyshev polynomial basis, applied to \( P_n \) and \( Q_m \) (Sect. 4.2.1) and (ii) the Chebyshev rational polynomials of the first kind, applied to \( R \) (Sect. 4.2.2). Among these options, the recursive representation of the Chebyshev rational polynomials of the first kind allows us to reduce the computation cost for the evaluation of the approximation and resembles the polynomial case. Finally, we discuss the accuracy and convergence of the corresponding rational approximation (Sect. 4.2.3).

### 4.2.1 Rational approximation with polynomial basis

**Approximation through the canonical polynomial basis**

Applied to \( P_n \) and \( Q_m \), we discuss the computation of the coefficients of the rational polynomial in terms of the Taylor coefficients \( a_k := \varphi^{(k)}(0)/k! \) of the input filter. Considering the power series \( \varphi(s) := \sum_{n=0}^{+\infty} a_n s^n \), we have that

\[
\varphi(s) - R(s) = \frac{\varphi(s)Q_m(s) - P_n(s)}{Q_m(s)} = \sum_{i=0}^{+\infty} q_i s^i \sum_{j=0}^{N} q_j s^j - \sum_{i=0}^{N} p_i s^i = 0.
\]

Defining the vectors \((p_i)_{i=n+1}^{N} = 0, (q_i)_{i=1}^{N} = 0\), the previous relation is equivalent to \( \sum_{i=0}^{+\infty} a_i q_i s^{i+j} = \sum_{j=0}^{N} p_i s^j \), \( j = 0, \ldots, N \). Introducing the index \( k := i + j \), this last relation is rewritten as \( \sum_{i=0}^{+\infty} a_i q_{k-i} - p_k = 0, k = 0, \ldots, N \). Recalling that \( q_0 = 1 \), we get \( a_0 = p_0 \) and the remaining unknowns are computed as the solution to the linear system

\[
\begin{bmatrix}
  a_1 & a_0 & \cdots & 0 \\
  a_2 & a_1 & a_0 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  a_N & a_{N-1} & \cdots & a_0
\end{bmatrix}
\begin{bmatrix}
  q_1 \\
  q_2 \\
  \vdots \\
  q_N
\end{bmatrix}
= \begin{bmatrix}
  1 \\
  1 \\
  \vdots \\
  1
\end{bmatrix}
\begin{bmatrix}
  p_1 \\
  p_2 \\
  \vdots \\
  p_N
\end{bmatrix}.
\]

**Approximation through the Chebyshev polynomial basis**

Applied to \( P_n \) and \( Q_m \). Instead of expanding the numerator and denominator in terms of the monomial basis, we use the Chebyshev rational polynomials \( T_k \), through the representation

\[
R := \sum_{k=0}^{N+m} \frac{p_k T_k}{q_k T_k}, \quad N = n + m, \quad q_0 = 1.
\]

Expanding the input filter \( \varphi = \sum_{n=0}^{+\infty} a_n T_n \) in a series of Chebyshev polynomials, we get the relation

\[
\varphi - R = \sum_{n=0}^{+\infty} a_n T_n \sum_{k=0}^{N+m} \frac{q_k T_k}{p_k T_k} - \sum_{k=0}^{N+m} q_k T_k = \sum_{k=0}^{N+m} \gamma_k T_k
\]

where the coefficients \((p_i)_{i=0}^{m} = 0, (q_i)_{i=0}^{N} = 0\) are chosen in such a way that the numerator has zero coefficient for \((T_k)_{k=0}^{N+m}\), i.e.,

\[
\varphi - R = \sum_{n=0}^{+\infty} a_n T_n \sum_{k=0}^{N+m} \frac{q_k T_k}{p_k T_k} - \sum_{k=0}^{N+m} q_k T_k = \sum_{k=0}^{N+m} \gamma_k T_k.
\]

To this end, we apply the following relations

\[
T_i T_j = \frac{1}{2} \left[ T_{i+j} + T_{|i-j|} \right], \quad a_0 = \frac{1}{2} f_{-1}, \quad a_k = \frac{2}{f_{|k|}} f_{k} \frac{df}{dx}, \quad k \geq 1
\]

whose evaluation is generally faster than monomial basis functions.

---

**Fig. 7: Level-sets of rational graph filters centred at a seed point \( p \) (black dot on the elbow) induced by the Chebyshev rational basis \( R_k(\Delta)\delta_p \) with an increasing degree.**

**Computation**

Once the rational approximation \( P_n/Q_m \) of \( \varphi \) has been computed with one of the two previous approaches, any signal \( \varphi(\Delta)f \) is computed as the solution to the problem \( Q_m(\Delta)g = P_n(\Delta)f \). In case of the canonical basis \( \{\mathcal{P}\} \), we evaluate the right- and left-side terms of

\[
\sum_{k=0}^{m} q_k \Delta^k g = \sum_{k=0}^{m} p_k \Delta^k f.
\]

In case of the Chebyshev polynomials \( \{\mathcal{P}\} \), we evaluate the right- and left-side term of

\[
\sum_{k=0}^{m} q_k T_k(\Delta)g = \sum_{k=0}^{m} p_k T_k(\Delta)f,
\]

through the recursive relations [16].

### 4.2.2 Rational approximation with rational basis

**The Chebyshev rational polynomials** are defined as

\[
R_n(x) := T_n \left( \frac{x - 1}{x + 1} \right) = \frac{2x - 1}{x + 1} R_{n-1}(x) - R_{n-2}(x)
\]

on the interval \([0, +\infty)\). These rational polynomials are orthogonal with respect to the weighted scalar product induced by \( w(x) := x^{-1/2} (1 + x)^{-1} \), according to the relation

\[
\langle R_n, R_m \rangle = \int_0^\infty \frac{R_n(x) R_m(x)}{x^{1/2} (x + 1)} dx = \left\{ \begin{array}{ll}
0 & n \neq m; \\
\pi & n = m = 0; \\
\pi/2 & n = m > 0.
\end{array} \right.
\]

For an arbitrary function \( f \in \mathcal{L}^2(\mathbb{R}) \), the orthogonality of the Chebyshev rational polynomials allows us to apply the relation \( f(x) = \sum_{n=0}^{+\infty} F_n R_n(x) \), \( F_n := \langle f, R_n \rangle \). In Fig. [7] we show the level-sets of rational graph filters \( R_k(\Delta)\delta_p \) (c.f., Eq. (21)) centred at a seed point \( p \) (black dot) induced by Chebyshev rational polynomials with increasing degree.

**Computation**

Through the recursive relation (21), the rational graph filter is efficiently evaluated as

\[
f_n := R_n(\Delta)f = 2(\Delta + id)^{-1}(\Delta - id)f_{n-1} - f_{n-2} = 2g_{n-1} - f_{n-2}
\]

where \( g_{n-1} \) is the solution to \((\Delta + id)g_{n-1} = (\Delta - id)f_{n-1}\). We notice that \( g_{n-1} \) is uniquely defined as the operator \((\Delta + id)\) is positive-definite.

### 4.2.3 Approximation accuracy and convergence

Approximating the input filter with a rational polynomial \( \rho^{(r)} \) of degree \( (n, m) \) (e.g., \( m = n = r \)), the sequence \( (\Psi^{(r)} f)_{r=0}^{\infty} \)

\[
\Psi^{(r)} f := \sum_{n=0}^{+\infty} \rho^{(r)}(\lambda_n) \langle f, \phi_n \rangle \phi_n
\]

induced
by the rational polynomial approximation $\rho^{(r)}$ of $\varphi$, converges to $\Psi_\varphi f$; in fact,
\[
\|\Psi_{\rho^{(r)}} f - \Psi_\varphi f\|^2 \leq \|\rho^{(r)}\|_\infty \sum_{n=0}^{\infty} |\langle f, \phi_n \rangle|^2 = \sigma_r^2\|f\|^2,
\]
where $\sigma_r = O(s^{n+1})$, $s \to 0$, is the approximation error between $\varphi$ and $\rho^{(r)}$.

## 5 Numerical computation

Even though a central element in spectral graph processing is the evaluation of the spectrum, or equivalently, of the characteristic polynomial of the graph Laplacian, the fast computation of the characteristic polynomial and numerical instabilities associated with multiple or close eigenvalues has not been addressed in detail by previous work. As novel contributions, we focus on the fast approximation of the characteristic polynomial and on the definition of the pseudo-spectrum, which allows us to identify a subset of the eigenvalues that is robust with respect to data perturbation. To this end, we introduce the discrete spectral wavelets and kernels (Sect. 5.1); then, we define the pseudo-spectrum and spectral density for the approximation of the characteristic polynomial (Sect. 5.2). Finally, we discuss the evaluation of the rational basis, the numerical stability and computational cost of the proposed approach (Sect. 5.3).

### 5.1 Discrete spectral wavelets/kernels

**Graph Laplacian** For the discretisation of the spectral operators, wavelets, and filtered convolution operators, let us consider a graph $\mathcal{G}$ and a signal $f : \mathcal{G} \to \mathbb{R}$ identified with the vector $f := (f(p_i))_{i=1}^t$ of $f$-values at the nodes of the input graph. The graph Laplacian is defined as $L := D^{-1}W$, where $W := D - D$. Here, $W$ is the weight matrix whose entry $(i, j)$ is the weight associated with the corresponding edge, and $D$ is the diagonal matrix whose entries are the sum of the rows of $L$. Alternatively, we can consider the normalised graph Laplacian $\tilde{L} := D^{-1/2}LD^{-1/2}$. We notice that the graph Laplacian is $D$-adjoint, i.e., $\langle Lf, g \rangle_D = \langle f, Lg \rangle_D$. In our discussion, we focus on the graph Laplacian; analogous considerations can be derived for the normalised graph Laplacian. For the graph Laplacian, the spectral decomposition is $LX = DX\Lambda$, $X^TDX = I$, where $X := [x_1, \ldots, x_n]$ is the eigenvectors’ matrix and $\Lambda$ is the diagonal matrix of the eigenvalues $(\lambda_i)_{i=1}^n$.

**Discrete spectral wavelets and kernels** Noting that $(\varphi(\lambda_n), \phi_n)_{n=1}^\infty$ is the eigensystem of $\Psi_\varphi$, the spectral representation of the data-driven kernel is $K_\varphi$, such that
\[
\varphi(\lambda_i) = \langle \Psi_\varphi f, \phi_j \rangle = x_i^T K_\varphi x_j, \quad \forall i = 1, \ldots, n.
\]

Indeed, $K_\varphi = X \varphi(\Lambda) X^TD$ is the spectral kernel, which is a filtered version of the Laplacian matrix and $D$-adjoint. According to \cite{21}, the approximation of $K_\varphi$ with a new kernel $K^r_\varphi$ reduces to the approximation of the corresponding filters. The approximation $\varphi$ of $\rho$ is computed on the interval $[0, \lambda_{\max}(L)]$, where the maximum Laplacian eigenvalue is evaluated by the Arnoldi method \cite{21}, or is set equal to the upper bound \cite{28, 29}.

\[
\lambda_{\max}(L) \leq \min \{ \max \{ \sum_j \tilde{L}(i, j) \}, \max \{ \sum_i \tilde{L}(i, j) \} \}.\]

### 5.2 Polynomial filters: pseudo-spectrum and density

In spectral graph processing, the numerical stability of the spectrum and the evaluation of the characteristic polynomial of matrices associated with large graphs are two important aspects not addressed in detail by previous work. Indeed, we discuss the sensitivity of the computation of the spectrum with respect to the presence of multiple or close eigenvalues, the computation of the characteristic polynomials in case of multiple eigenvalues (Sect. 5.2.1), and the definition of the pseudo-spectrum as a way to compute a stable subset of the eigenvalues (Sect. 5.2.2). Then, we address the approximation of the characteristic polynomial of a large matrix through spectral densities, which reduces to the computation of the trace of Chebyshev polynomial matrices. This approximation of the characteristic polynomial is necessary to apply the Cayley-Hamilton theorem for the reduction of the degree of polynomial filters (Sect. 5.2.3).

#### 5.2.1 Characteristic polynomials: multiple eigenvalues

One of the main issues with spectral graph processing is the presence of multiple or close Laplacian eigenvalues, which makes the computation of the spectrum and the evaluation of the spectral filters numerically unstable. Even though these situations are quite common in applications (e.g., for symmetric graph), previous work has paid a little attention to this problem that we address by discussing the computation of the characteristic polynomial in case of multiple eigenvalues.

Given an arbitrary square matrix $A \in \mathbb{R}^{n \times n}$ (e.g., the Laplacian matrix, a kernel matrix $\varphi(L)$), or a kernel matrix, we consider its characteristic polynomial $P(s) := \det(A - sI)$. Indicating the eigenvalues of $A$ as $(\lambda_i)_{i=1}^n$, previous work computes the coefficient of the characteristic polynomial $P(s) = \sum_{i=0}^n \alpha_i s^i$ by imposing the interpolating conditions $P(\lambda_i) = 0, \forall i$, i.e., solving the homogeneous linear system $V \alpha = 0$, where $V$ is the Vandermonde matrix and $\alpha := (\alpha_i)_{i=1}^n$ is the unknown vector. In case of close or repeated eigenvalues, the Vandermonde coefficient matrix is singular, as it has multiple identical rows. In this case, for each repeated eigenvalue $\lambda_i$ of multiplicity $m$ the first ($m - 1$) derivatives of the characteristic polynomial vanishes at $\lambda_i$. Combining the interpolating conditions $f(\lambda_i) = R(\lambda_i)$ at all the eigenvalues with the constraints $\partial_s^k \varphi(\lambda_i) = \partial_s^k R(\lambda_i), k = 1, \ldots, m - 1$, on the derivatives

![Fig. 8: Colormap and level-sets of the diffusion wavelet at a source point computed with the P.-C. approximation and the truncatc approximation with 100 eigenfunctions, which is affected by undulations far from the source point and at small scales.](image)

| t | P.-C. approx. r := 7 | Trunc. approx. (100 eigs) |
|---|-----------------|-----------------|
| $t = 0.01$ | ![Image 1] | ![Image 2] |
| $t = 0.001$ | ![Image 3] | ![Image 4] |
at multiple eigenvalues, we get a non-singular linear system whose solution uniquely determines the coefficients of the remainder polynomial.

5.2.2 Stability of the eigenpairs and pseudo-spectrum

Assuming that $\mu := \rho(\lambda)$ is an eigenvalue of $\varphi(L)$ with multiplicity $m$ and rewriting the characteristic polynomial of $\varphi(L)$ as $P(s) = (s - \mu)^m Q(s)$, where $Q(\cdot)$ is a polynomial of degree $(n - m)$ and $Q(\mu) \neq 0$, we get that

$$(s - \mu)^m = \frac{P(s)}{Q(s)} \approx \frac{O(\delta)}{Q(s)} \delta \to 0,$$

as $P(s) \to 0$, $s \to \mu$, i.e., $s \approx \mu + O(\delta \mu)$. Indeed, modifying the spectrum (i.e., the underlying graph) in such a way that the filtered eigenvalues are perturbed by $\delta := 10^{-m}$ corresponds to a change of order $0.1$ in $\mu$ (i.e., $s \approx \mu + 0.1$) and this amplification becomes larger as the multiplicity of the eigenvalue increases.

Pseudo-spectrum To handle numerical instabilities associated with multiple or close eigenvalues (e.g., for symmetric graphs) of the Laplacian or spectral kernel matrix, we introduce the pseudo-spectrum, which defines the eigenvalues of a matrix with respect to a threshold. Let $A$ be a $n \times n$ symmetric matrix (or, more generally, a $D$-adjoint matrix as $\varphi(L)$) and $\lambda_r(A) := \{z \in \mathbb{R} : \|zI - A\|_2 \geq 1/\epsilon\}$ the pseudo-spectrum of $A$, i.e., the set of disks of radius $\epsilon$ centred around the eigenvalues of $A$. Indeed, $z$ is a pseudo-eigenvalues of $A$ if and only if $(zI - A)$ is sufficiently close to be singular (i.e., with respect to the threshold $\epsilon$). In fact,

$$\frac{1}{\epsilon} \leq \|zI - A\|_2 = |\lambda_{\text{max}}(zI - A)|^{-1} = |\lambda_{\text{min}}(zI - A)|^{-1},$$

i.e., $\lambda_{\text{min}}(zI - A) \leq \epsilon$. The pseudo-spectrum of $A$ generalises the spectrum (i.e., $\lambda_0(A) = \lambda(A)$) and $z \in \lambda_r(A)$ if and only if $z \in \lambda(A + E)$ and $\|E\|_2 \leq \epsilon$. If $\epsilon_1 \geq \epsilon_2$, then $\lambda_{\epsilon_1}(A) \subseteq \lambda_{\epsilon_2}(A)$.

5.2.3 Approximated spectral densities

Given a matrix $A$ with eigenvalues $(\lambda_i)_{i=1}^n$, its pointwise and smooth spectral densities \(^{[20]}\) in an interval $[a, b]$ are defined as the functions

$$\delta := \frac{1}{n} \int_a^b \sum_{i=1}^n \delta(s - \lambda_i) ds, \quad \delta(t) := \frac{1}{n \sigma} \sum_{i=1}^n h_\sigma(t - \lambda_i).$$

The pointwise spectral density measures the percentage of eigenvalues belonging to the selected interval, and the smooth density is defined by replacing the $\delta$-function with a smooth hat function, such as the Gaussian function $h_\sigma(t) := (2\pi\sigma^2)^{-1/2} \exp(-t^2/\sigma^2).$ Here, larger values of the width $\sigma$ provide smoother curves while reducing the interpolation of the eigenvalues. Assuming that the input matrix has been normalised in such a way that its eigenvalues belong to the interval $[-1, 1]$, we express the normalised spectral density $\hat{\delta}(t) := (1 - t^2)^{3/2} \delta(t)$ in terms of the Chebyshev polynomials as $\delta = \sum_{n=0}^{+\infty} \mu_n T_n$. Then,

$$\mu_k = \langle \delta, T_k \rangle_w, = \frac{1}{n} \sum_{i=1}^n \int_{-1}^1 \delta(s - \lambda_i) T_k(s) ds,$$

Indeed, each coefficient reduces to the evaluation of the trace of the matrix achieved by evaluating the corresponding Chebyshev polynomial on the input matrix. For more details on the approximation of the spectral densities of large matrices with polynomial methods and spectroscopic approaches, we refer the reader to \(^{[22]}\).
5.3 Numerical stability and computational cost

Rational spectrum-free computation According to Eq. (22), the vector \( R_{n+1}(\mathbf{L})f \) is recursively computed as

\[
f_{n+1} := R_{n+1}(\mathbf{L})f = 2(\mathbf{L} + \mathbf{I})^{-1}(\mathbf{L} - \mathbf{I})f_n - f_{n-1} = 2g_n - f_{n-1},
\]

where the vector \( g_n \) is the solution to the sparse, symmetric, and positive definite linear system \((\mathbf{L} + \mathbf{I})g_n = (\mathbf{L} - \mathbf{I})f_n\), or equivalently \((\mathbf{L} + \mathbf{D})g_n = (\mathbf{L} - \mathbf{D})f_n\). Since the coefficient matrix \((\mathbf{L} + \mathbf{D})\) is independent of the iteration and positive-definite, it can be pre-factorised and its pre-factorisation is used for the computation of \(a_n\), in line at each iteration. In an analogous way, we derive the discrete counterparts of Eqs. (19), (20). In Fig. 5, the diffusion kernel has been computed through the spectrum-free approximation with Chebyshev rational polynomials and the recursive relation in Eq. (23). The shape and distribution of the level-sets confirm the high accuracy of this approximation at small and large scales.

Conditioning of the spectral wavelet/kernel If \( \varphi \) is an increasing function (i.e., \( \varphi \) is a low pass filter), then the conditioning number of the spectral kernel is bounded as

\[
\kappa_2(K_\varphi) = \kappa_2(\varphi(\mathbf{L})) = \frac{\max_{i=1,...,n} \{ \varphi(\lambda_i) \}}{\min_{i=1,...,n} \{ \varphi(\lambda_i) \}} = \frac{\|\varphi\|_\infty}{\varphi(0)},
\]

and it is ill-conditioned when \( \varphi(0) \) is close to zero or \( \varphi \) is unbounded. If \( \rho \) is bounded and \( \varphi(0) \) is not too close to 0, then the spectral kernel is well-conditioned. If \( \varphi(0) \) is null, then we consider the smallest and not null filtered Laplacian eigenvalue at the denominator of the previous relation.

Computational cost The computational cost of the truncated spectral approximation of the kernel/wavelet depends on the sparsity degree of the Laplacian matrix and takes from \( O(kn \log n) \) to \( O(kn^2) \) time, where \( k \) is the number of selected eigenpairs. Choosing a rational approximation of the input filter of degree \((r, l)\), the evaluation of the spectral kernel/wavelet is reduced to \( r \) linear systems whose coefficient matrix is \( \mathbf{L} \). Through iterative solvers, the computational cost is \( O(\tau(n)) \), where \( \tau(n) \) is the cost for the solution of a sparse linear system, which varies from \( O(n) \) to \( O(n^2) \), according to the sparsity of the coefficient matrix, and it is \( O(n \log n) \) in the average case. Indeed, the polynomial and rational polynomial approximations have the same order of computational complexity.

Evaluating the spectral kernel/wavelet at one scale with the rational approximation is generally more efficient and accurate than the truncated spectral approximation, especially at small scales (Fig. 8). In the average case, the cost is \( O(rn \log n) \) versus \( O(kn \log n) \), with \( r << k \), e.g., \( r := 5, 7 \) and \( k = 50, 100 \). In case of multiple scales, the eigensystem is computed only once and applied for the evaluation of spectral wavelets at all scales in linear time, while the linear systems associated with the rational approximation is solved for each scale. Assuming \( s \) scales, the computational cost of the rational approximation is competitive with respect to the truncated spectral approximation if \( sr < k \), i.e., the number of scales is lower than the ration \( k/r \) between the number \( k \) of selected eigenpairs and the degree \( r \) of the rational polynomial.

6 Applications

Through Fourier-based and graph filters, we define the spectral kernels/functions \( \Phi(\delta_p) = K_\varphi(p, \cdot) \) centred at a point \( p \) as the action of the spectral operator on \( \delta_p \). In this setting, we discuss applications of spectral kernels to signal reconstruction and smoothing and shape correspondence.

Signal reconstruction and smoothing In Fig. 8, we report the \( \ell_\infty \) approximation error of the reconstruction of the geometry \((x, y, z \text{ coordinates})\) of different 3D shapes (e.g., the plane wings, the table legs, the legs and the harms, the tail and the legs).

Signal smoothing For spectral smoothing, we consider a noisy signal \( f : f + \delta \) (e.g., the \( x, y, z \) coordinates of the points on a surface), where \( \delta \) is a Gaussian noise, and a set \( B \) of diffusion functions centred at \( k := 100 \) samples, evaluated with the farthest point sampling from a seed point, and at \( s := 6 \) scales. Then, we compute a smoothed signal \( g \) as the least-squares projection of \( f \) on \( B \) and evaluate the corresponding approximation error as
consider a set of local and multi-scaled filtered spectral landmarks, which are used to initialise the functional, and Laplacian eigenfunctions and (ii) the diffusion functions of correspondences, the diffusion basis functions generally fusion functions has been compared with the one induced The functional map induced by the corresponding 48 differences correctly map local and global features on the source and articulated shapes (Fig. 11), the computed correspondences on partial 3D shapes with a similar (cut data set from SHREC’16 [32]) and irregular (FARM partial data set [33]) connectivity, and (ii) compared with state-of-the-art methods discussed in [32]. The results achieved with the spectral wavelet are comparable with the best current method [34] for partial functional maps (PFM) on 3D shape with similar connectivity (Fig. 12a); contrary to PFM, our results remain reliable when matching shapes with a highly different connectivity (Fig. 12a). For more details, we refer the reader to our recent work presented in [41].

\[ \| f - g \|_\infty / \| f \|_\infty \] The smoothness order and the approximation accuracy (Fig. 10 y-axis) increase as the number of functions at any selected scale. Best results are achieved by selecting diffusion wavelets at small scales in order to accurately recover the local and global details of the signal.

Shape correspondences We now apply specific classes of the spectral kernels to high-level tasks, such as shape correspondence through the functional map framework [35]. To this end, we select a low number \( k \) of ground-truth landmarks, which are used to initialise the functional, and consider a set of local and multi-scaled filtered spectral kernels, which generate the sub-space on which the shape descriptors will be projected to define the functional map. Firstly, we compare the quality of the correspondence map computed on the function space generated by (i) \( k = 60 \) Laplacian eigenfunctions and (ii) the diffusion functions (i.e., 15 seed points, 4 scales) 5 couples of shapes belonging to 5 classes of the SHREC’10 data set. Indeed, we consider the same number of functions but with different properties in terms of locality and information encoding. For the rigid and articulated shapes (Fig. 11), the computed correspondences correctly map local and global features on the source and target shape. To this end, we have selected 7 ground-truth landmarks, 5 uniformly sampled seed points and 4 scales on 15 shapes. The functional map induced by the corresponding 48 diffusion functions has been compared with the one induced by 60 Laplacian eigenfunctions. According to the variation of the mean correspondence error and the examples of correspondences, the diffusion basis functions generally provide a lower correspondence error before and after the optimisation based on the Iterative Closed Point (ICP, for short) [40]. Furthermore, the diffusion functions improve the quality of the correspondences with respect to the Laplacian eigenbasis, e.g., between the legs of the giraffe and the tail of the cow, the legs of the giraffe and the horns of the goat, the legs of the dog and the horns of the cow.

Partial shape correspondences For comparison with previous work, we focus on partial shape correspondence. As descriptor, we select the spectral kernel \( K_\varphi(p, \cdot) := L_\varphi \delta_p \), which is induced by spectral convolution operator \( \Psi_\varphi := \exp(-\Delta)\varphi \) and the filter \( \varphi(s) := \exp(-st)s \). Indeed, the spectral kernel \( K_\varphi(p, \cdot) \) is achieved by applying the diffusion operator \( \exp(-t\Delta) \) to the smooth approximation \( \Delta \delta_p \), which can be interpreted as a Mexican hat function. Then, the spectral operator is computed through the rational approximation \( \varphi(s) \approx (sP_r(s))/Q_r(s) \), which \( P_r/Q_r(r, r) \) Padé-Chebyshev approximation of thee exponential function. In Fig. 12 the spectral wavelet induced by the filter \( \varphi \) has been (i) applied to shape correspondence on partial 3D shapes with a similar (cut data set from SHREC’16 [32]) and irregular (FARM partial data set [33]) connectivity, and (ii) compared with state-of-the-art methods discussed in [32]. The results achieved with the spectral wavelet are comparable with the best current method [34] for partial functional maps (PFM) on 3D shape with similar connectivity (Fig. 12a); contrary to PFM, our results remain reliable when matching shapes with a highly different connectivity (Fig. 12a). For more details, we refer the reader to our recent work presented in [41].

7 CONCLUSIONS AND FUTURE WORK
This paper has discussed the definition of novel Fourier-based and rational spectral operators for graph processing, which generalise the notion of polynomial spectral filters and Fourier transform to non-Euclidean domains. As future work, we plan to investigate the usefulness of rational filters for the definition of a family of spectral bases for data analysis. In fact, several signal processing techniques generally represent the data in terms of a given basis in order to highlight a given class of underlying properties or features, e.g., localising content in both space and frequency through wavelet basis. Rational filters, and in particular rational Chebyshev polynomials, are particularly useful to enlarge the class of learning networks, as a generalisation of networks based on polynomial filters, such as PolyNet [23]. ChebNet [24], CayleyNet [25], in order to improve the discriminative capabilities of networks in the context of 3D geometric deep learning.

Acknowledgments: We thank the Reviewers for their thorough review and constructive comments, which helped us to improve the technical part and presentation of the paper. Graphs are courtesy of the “Gaimc: Graph Alg.” Library.

REFERENCES
[1] P. Berard, G. Besson, and S. Gallot, “Embedding Riemannian manifolds by their heat kernel,” Geometric and Functional Analysis, vol. 4, no. 4, pp. 373–398, 1994.
[2] M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data representation,” Neural Computations, vol. 15, no. 6, pp. 1373–1396, 2003.
[3] S. Ghosh, N. Das, T. Gonzalez, P. Quaresma, and M. Kundu, “The journey of graph kernels through two decades,” Computer Science Review, vol. 27, pp. 88 – 111, 2018.
[4] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs,” IEEE Trans. on Signal Processing, vol. 61, no. 7, pp. 1644–1656, 2013.
[5] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs: frequency analysis,” *IEEE Trans. on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[6] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, “Discrete signal processing on graphs: Sampling theory,” *IEEE Trans. on Signal Processing*, vol. 63, no. 24, pp. 6510–6523, 2015.

[7] A. Heimowitz and Y. C. Eldar, “A Markov variation approach to smooth graph signal interpolation,” https://arxiv.org/abs/1806.03174, 2018.

[8] S. Mahadevan and M. Maggioni, “Value function approximation with diffusion wavelets and laplacian eigenshapes,” in *Conf. on Neural Information Processing Systems*, 2005, pp. 843–850.

[9] S. Lafon, Y. Keller, and R. R. Coifman, “From graph to manifold Laplacian: the convergence rate,” *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 128 – 134, 2006.

[10] D. K. Hammond, P. Vandergheynst, and R. Gribonval, “Wavelets on graphs via spectral graph theory,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.

[11] H. Bahonar, A. Mirzaei, S. Sadri, and R. Wilson, “Graph embedding using frequency filtering,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. In press, 2019.

[12] N. Perraudin, B. Ricaud, D. J. Shuman, and P. Vandergheynst, “Global and local uncertainty principles for signals on graphs,” *APSIPA Trans. on Signal and Information Processing*, vol. 7, Ééd, 2018.

[13] K. Ramani and A. Sinha, “Multiscale kernels using random walks,” *Computer Graphics Forum*, vol. 33, no. 1, pp. 164–177, 2013.

[14] X. Zhu, Z. Ghahramani, and J. Lafferty, “Semi-supervised learning with gaussian fields and harmonic functions,” in *Conf. on Machine Learning*, 2003, pp. 912–919.

[15] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: analysis and an algorithm,” in *Proc. of Conf. on Neural Information Processing Systems: Natural and Synthetic*, 2001, pp. 849–856.

[16] O. Rippel, J. Snoek, and R. P. Adams, “Spectral representations for convolutional neural networks,” in *Conf. on Neural Information Processing Systems*, 2015, pp. 2449–2457.

[17] M. Bronstein and A. Bronstein, “Convolutional neural networks on graphs with fast localized spectral filtering,” in *Conf. on Neural Information Processing Systems*, 2016, pp. 3844–3852.

[18] M. Bronstein and A. Bruna, “Shape recognition with spectral distances,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1065 –1071, 2011.

[19] R. C. Wilson, E. R. Hancock, and Bin Loo, “Pattern vectors from algebraic graph theory,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 7, pp. 1112–1124, 2005.

[20] G. Golub and G. VanLoan, *Matrix Computations*. John Hopkins University Press, 2nd Edition, 1989.

[21] G. L. Litvinov, “Approximate construction of rational approximations and the effect of error autocorrection. applications,” *Russian Journal of Mathematical Physics*, vol. 1:3, pp. 313–352, 1993.

[22] X. Zhang, Z. Li, C. C. Loy, and D. Lin, “PolyNet: A pursuit of structural diversity in very deep networks,” in *Conf. on Computer Vision and Pattern Recognition*, 2017, pp. 3900–3908.

[23] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” https://arxiv.org/abs/1609.02907, 2016.

[24] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, “Cayleynets: Graph convolutional neural networks with complex rational spectral filters,” *IEEE Trans. on Signal Processing*, vol. 67, no. 1, pp. 97–109, Jan 2019.

[25] G. A. Baker and P. Graves-Morris, *Padé Approximants*, 2nd ed., ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996.

[26] C. Brezinski and M. R. Zaglia, *Extrapolation Methods: Theory and Practice*, ser. Elsevier Science. Cambridge University Press, 1991.

[27] R. Lehoucq and D. C. Sorensen, “Deflation techniques for an implicitly re-started Arnoldi iteration,” *SIAM Journal of Matrix Analysis and Applications*, vol. 17, no. 4, pp. 789–821, 1996.

[28] D. C. Sorensen, “Implicit application of polynomial filters in a k-step arnoldi method,” *SIAM Journal of Matrix Analysis and Applications*, vol. 13, no. 1, pp. 357-385, 1992.

[29] L. Lin, Y. Saad, and C. Yang, “Approximating spectral densities of large matrices,” *SIAM Review*, vol. 58, no. 1, pp. 34–65, 2016.

[30] A. Bronstein, M. M. Bronstein, U. Castellani, B. Falcidieno, A. Fusiello, A. Godil, L. Guibas, I. Kokkinos, Z. Lian, M. Ovsjanikov, G. Patane, M. Spagnuolo, and R. Toldo, “SHREC 2010: robust large-scale shape retrieval benchmark,” *Workshop on 3D Object Retrieval*, 2010.

[31] L. Cosmo, E. Rodolà, M. Bronstein, A. Torsello, D. Cremers, and Y. Sahillioğlu, “Partial matching of deformable shapes,” in *Workshop on 3D Object Retrieval*, 2016, pp. 61–67.

[32] S. Melzi, R. Marin, E. Rodolà, U. Castellani, J. Ren, A. Poulenard, P. Wonka, and M. Ovsjanikov, “SHREC 2019: Matching Humans with Different Connectivity,” in *Eurographics Workshop on 3D Object Retrieval*, 2019.

[33] E. Rodolà, L. Cosmo, M. Bronstein, A. Torsello, and D. Cremers, “Partial functional correspondence,” *Computer Graphics Forum*, vol. 36, no. 1, pp. 222–236, 2017.

[34] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. J. Guibas, “Functional maps: a flexible representation of maps between shapes,” *ACM Trans. on Graphics*, vol. 31, no. 4, p. 30, 2012.

[35] E. Rodolà, S. Bulò, T. Windheuser, M. Vestner, and D. Cremers, “Dense non-rigid shape correspondence using random forests,” in *Int. Conf. on Computer Vision and Pattern Recognition*, 2014, pp. 4177–4184.

[36] J. Sun, M. Ovsjanikov, and L. J. Guibas, “A concise and provably informative multi-scale signature based on heat diffusion,” *Computer Graphics Forum*, vol. 28, no. 5, pp. 1383–1392, 2009.

[37] E. Rodolà, A. Torsello, T. Harada, Y. Kuniyoshi, and D. Cremers, “Elastic net constraints for shape matching,” in *Int. Conf. on Computer Vision*, 2013, pp. 1169–1176.

[38] E. Rodolà, A. M. Bronstein, A. Albarelli, F. Bergamasco, and A. Torsello, “A game-theoretic approach to deformable shape matching,” in *Conf. on Computer Vision and Pattern Recognition*, 2012, pp. 182–189.

[39] D. Nogneng and M. Ovsjanikov, “Informative descriptor preservation via commutativity for shape matching,” *Computer Graphics Forum*, vol. 36, no. 2, pp. 259–267, 2017.

[40] M. Kirgo, S. Melzi, G. Patanë, E. Rodolà, and M. Ovsjanikov, “Wavelet-based heat kernel derivatives: Towards informative localized shape analysis,” https://arxiv.org/abs/2007.11632, 2020.