

Neural Field Convolutions by Repeated Differentiation

NTUMBA ELIE NSAMPI, MPI Informatik, Germany
ADARSH DJEACOMAR, MPI Informatik, Germany
HANS-PETER SEIDEL, MPI Informatik, Germany
TOBIAS RITSCHEL, University College London, United Kingdom
THOMAS LEIMKÜHLER, MPI Informatik, Germany

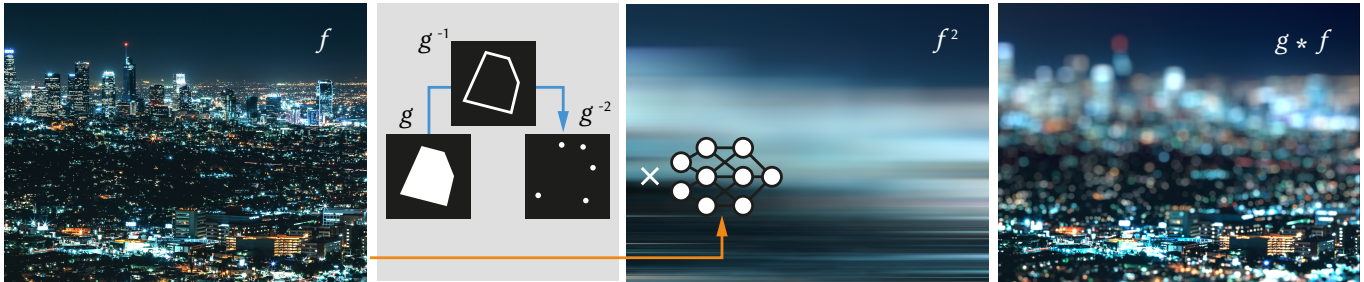


Fig. 1. We introduce an algorithm to perform efficient continuous convolution of neural fields f by piecewise polynomial kernels g . The key idea is to convolve the sparse repeated derivative of the kernel (g^{-n}) with the repeated antiderivative of the signal (f^n).

Neural fields are evolving towards a general-purpose continuous representation for visual computing. Yet, despite their numerous appealing properties, they are hardly amenable to signal processing. As a remedy, we present a method to perform general continuous convolutions with general continuous signals such as neural fields. Observing that piecewise polynomial kernels reduce to a sparse set of Dirac deltas after repeated differentiation, we leverage convolution identities and train a repeated integral field to efficiently execute large-scale convolutions. We demonstrate our approach on a variety of data modalities and spatially-varying kernels.

CCS Concepts: • **Computing methodologies** → **Neural networks**; **Computer graphics**.

Additional Key Words and Phrases: Convolution, Geometry Processing, Image Processing, Neural Fields, Signal Processing, Sparsity

ACM Reference Format:

Ntumba Elie Nsambi, Adarsh Djeacomar, Hans-Peter Seidel, Tobias Ritschel, and Thomas Leimkühler. 2023. Neural Field Convolutions by Repeated Differentiation. *ACM Trans. Graph.* 42, 6, Article 206 (December 2023), 12 pages. <https://doi.org/10.1145/3618340>

1 INTRODUCTION

Neural fields have recently emerged as a powerful way of representing signals and have witnessed widespread adoption in particular for visual data [Tewari et al. 2022; Xie et al. 2022]. Also referred to as implicit or coordinate-based neural representations, neural fields typically use a multi-layer perceptron (MLP) to encode a mapping

from coordinates to values. This representation is universal and allows to capture a multitude of modalities, such as mapping from 2D location to color for images [Stanley 2007], from 3D location to the signed distance to a surface for geometry [Park et al. 2019], from 5D light field coordinates to emitted radiance of an entire scene [Mildenhall et al. 2020], and many more.

The appealing properties of neural fields are three-fold: First, they represent signals in a continuous way, which is a good fit for the mostly continuous visual structure of our world. Second, they are compact, since they encode complex signals into a relatively small number of MLP weights [Dupont et al. 2021], while adapting well to local signal complexities. Third, they are easy to optimize by construction. Taking all of these properties together, it comes at no surprise that neural fields are rapidly evolving towards a general-purpose data representation [Dupont et al. 2022a]. However, to be a true alternative to established specialized representations such as pixel arrays, meshes, point clouds, etc., neural fields are still lacking in a fundamental aspect: They are hardly amenable to *signal processing* [Xu et al. 2022; Yang et al. 2021]. As a remedy, in this work, we propose a general framework to apply a core signal processing technique to neural fields: convolutions.

The versatility and expressivity of neural representations have evolved significantly over the last couple of years, mostly due to advances in architectures and training methodologies [Hertz et al. 2021; Müller et al. 2022; Sitzmann et al. 2020; Tancik et al. 2020]. However, at their core, neural fields only support *point samples*. This is sufficient for point operations, such as the remapping of input coordinates, e.g., for the purpose of deformations [Kopanas et al. 2022; Park et al. 2021; Treitschke et al. 2021; Yuan et al. 2022], or the remapping of output values [Vicini et al. 2022]. In contrast, a convolution requires the *continuous integration* of values over coordinates weighted by a continuous kernel.

Authors' addresses: Ntumba Elie Nsambi, MPI Informatik, Germany, nnsampi@mpi-inf.mpg.de; Adarsh Djeacomar, MPI Informatik, Germany, adjeacou@mpi-inf.mpg.de; Hans-Peter Seidel, MPI Informatik, Germany, hpseidel@mpi-sb.mpg.de; Tobias Ritschel, University College London, United Kingdom, t.ritschel@ucl.ac.uk; Thomas Leimkühler, MPI Informatik, Germany, thomas.leimkuehler@mpi-inf.mpg.de.

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM Transactions on Graphics*, <https://doi.org/10.1145/3618340>.

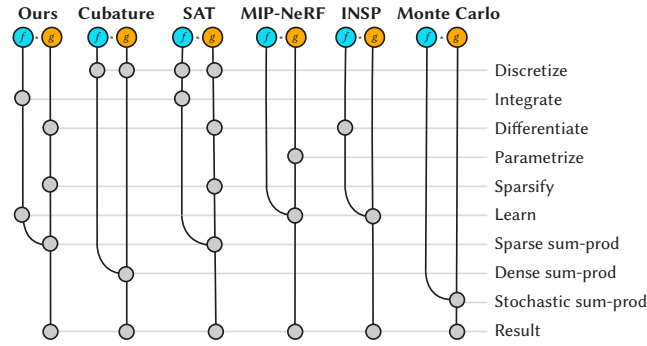


Fig. 2. The landscape of convolution methods as combinations of different operations applied to a signal and kernel (*top*), leading to a result (*bottom*).

Aggregation in neural fields can be approximated using either discretization followed by cubature, or Monte Carlo sampling, resulting in excessive memory requirements and noise, respectively. Another solution is to consider a narrow parametric family of kernels and train the field using supervision on filtered versions of the signal [Barron et al. 2021]. Representation and learned convolution operation can be explicitly disentangled using higher-order derivatives [Xu et al. 2022], but this comes at the cost of only supporting small spatially invariant kernels. AutoInt [Lindell et al. 2021] performs analytic integration using automatic differentiation, but only considers unweighted integrals. We advance the state of the art by presenting a method to *efficiently* perform *general, large-scale, spatially-varying* convolutions *natively* in neural fields.

In our approach, we consider neural fields to be convolved with piecewise polynomial kernels, which reduce to a sparse set of Dirac deltas after repeated differentiation [Heckbert 1986]. Combining this insight with convolution identities on differentiation and integration, our approach requires only a small number of point samples from a neural integral field to perform an exact continuous convolution, independent of kernel size. This integral field needs to be trained in a specific way, supervised via continuous higher-order finite differences, corresponding to a minimal kernel of a certain polynomial degree. Once trained, our neural fields are ready to be convolved with *any* piecewise polynomial kernel of that degree.

We showcase the generality and versatility of our approach using different modalities, such as images, videos, geometry, character animations, and audio, all natively processed in a neural field representation. Further, we demonstrate (spatially-varying) convolutions with a variety of kernels, such as smoothing and edge detection of different shapes and sizes. In summary, our contributions are:

- A principled and versatile framework for performing convolutions in neural fields.
- Two novel enabling ingredients: An efficient method to train a repeated integral field, and the optimization of continuous kernels that are sparse after repeated differentiation.
- The evaluation of our framework on a range of modalities and kernels.

2 RELATED WORK

The convolution of a signal, eventually in some higher dimension, with a kernel is a central operation in modern signal processing [Zhang 2022]. In this work, we consider a slightly generalized form of convolution, where a kernel *varies spatially* across the signal. Refer to appendix A in the supplemental document for a tabulated overview of the different solutions discussed next.

Discrete. For discrete signals and kernels (Cubature and SAT in Fig. 2), most convolutions are based on cubature, i.e., a dense sum-product operation across all dimensions. This, unfortunately, does not scale to larger kernels or higher dimensions but allows spatially-varying kernels. A common acceleration is the *Fourier transform* [Brigham 1988], which also requires time and space to perform and store the transformed signal. Most of all, it requires the kernel to be spatially invariant. As a remedy, certain spatially-varying convolutions can be realized using spatially-varying combinations or transformations of stationary filters [Fournier and Fiume 1988; Freeman et al. 1991; Mitchel et al. 2020]. For a large class of filters, *pyramidal* schemes [Farbman et al. 2011; Williams 1983] can be a solution, but require additional memory. The key idea is that intermediate pyramid values store a partial aggregate of the signal. Techniques that store integrals without reducing the resolution are called *summed-area tables* (SAT) or *integral images* [Crow 1984; Viola and Jones 2001]. Notably, SATs and their variants allow efficient spatially-varying convolution by considering differentiated kernels [Heckbert 1986; Leimkühler et al. 2018; Simard et al. 1998]. This efficiency comes from the fact that the differentiated kernel is sparse (“Sparsify” for SAT in Fig. 2) and the SAT only needs to be evaluated at very few locations. Our approach will take this idea to the continuous neural domain.

Continuous. Convolution becomes more challenging, if the signal, the kernel, or both are continuous. *Monte Carlo* methods, that straightforwardly sample signal and kernel randomly and sum the result (Monte Carlo in Fig. 2) can handle this case. These scale very well to high dimensions, but at the expense of noise that only vanishes with many samples, even when specialized blue noise [Singh et al. 2019] or low-discrepancy [Niederreiter 1992; Sobol 1967] samplers are used. Similar to our approach, the use of derivatives has been shown to be beneficial [Kettunen et al. 2015]. Practical unstructured convolution [Hermosilla et al. 2018; Shocher et al. 2020; Vasconcelos et al. 2023; Wang et al. 2018] does away with cubature and evaluates the product of kernel and signal only at specific sparse positions such as the points in a point cloud. Our approach does not rely on random sampling but works directly on a continuous signal.

Neural. It has recently been proposed to replace discrete representations with continuous neural networks, so-called *neural fields* [Tewari et al. 2022; Xie et al. 2022]. These have applications in geometry representation [Park et al. 2019], novel-view synthesis [Mildenhall et al. 2020; Sitzmann et al. 2019], dynamic scene reconstruction [Park et al. 2021; Treitschke et al. 2021; Yuan et al. 2022], etc. Replacing a discrete grid with complex continuous functions requires developing the same operations available to grids [Dupont et al. 2022a], including convolutions, as we set out to do in this work. Early work has been conducted to explore the manifold of all

natural neural fields [Du et al. 2021] and to build a generative model of neural fields [Dupont et al. 2022b; Erkoç et al. 2023]. Specialized network architectures allow the decomposition of signals into a discrete set of frequency bands [Fathony et al. 2020; Lindell et al. 2022; Yang et al. 2022]. Further, limited forms of geometry processing have been considered in this representation [Yang et al. 2021]. However, none of them is looking into general, efficient, large-scale, and/or spatially-varying convolutions.

A very specific form of convolution occurs as anti-aliasing or depth-of-field in image-based rendering. To account for these effects, neural fields can be learned that are conditioned on the parameters of the convolution kernel, such as its bandwidth [Barron et al. 2021, 2022; Isaac-Medina et al. 2023; Wang et al. 2022]. The network can then be evaluated to directly produce the filtered result, i.e., signal representation and convolution operation are intertwined. This Mip-NeRF-style convolution is in principle applicable to other filters, as long as they can be parametrized to become conditions to input into the network (MIP-NeRF in Fig. 2). Unfortunately, this limits the kernels that can be applied to a parametric family that needs to be known in advance. Further, it significantly increases training time, since kernel parameters act as additional input dimensions to the network. We use inductive knowledge of integration and differentiation to arrive at a more efficient formulation that generalizes across kernels.

The key to efficient convolutions is a combination of sparsity, differentiation, and integration. Fortunately, tools to perform integration and differentiation on neural fields are available. AutoInt [Lindell et al. 2021] proposes to learn a neural network that, when automatically differentiated, fits a signal. By evaluating the original network without differentiation, the antiderivative can be evaluated conveniently. Unfortunately, this approach does not scale well to the higher-order antiderivatives needed for efficient convolutions, as the size of the derivative graphs grows quickly. In contrast, our approach leverages higher-order finite differences to train a repeated integral field, which scales with the number of integrations required.

Recently, Xu et al. [2022] have proposed a method with the same aim as ours (INSP in Fig. 2). Given a trained neural field, they learn to combine higher-order derivatives of the field to approximate a convolution. Similar to a Taylor expansion, this requires high-order derivatives to reason about larger neighborhoods and, unfortunately, hence only allows for very small, spatially invariant kernels.

Finally, aggregation in neural fields in the form of range queries has been studied by Sharp and Jacobson [2022]. Their approach allows to retrieve a conservative estimate of the field’s extrema within a query volume, which unfortunately does not provide enough information to perform accurate continuous convolutions.

3 BACKGROUND

We consider the convolution of arbitrary continuous signals $f \in \mathbb{R}^{d_{\text{in}}} \rightarrow \mathbb{R}^{d_{\text{out}}}$ with arbitrary continuous kernels $g \in \mathbb{R}^d \rightarrow \mathbb{R}$. Both inputs and outputs of f are low- to medium-dimensional. The signal can be any continuous function, including but not limited to a neural network. We assume the kernel has compact but potentially large support. The kernel does not necessarily extend across *all* input dimensions of f , i.e., $d \leq d_{\text{in}}$. To simplify our exposition,

without loss of generality, we assume that the first d dimensions of f correspond to the filter dimensions of g . Further, we allow g to vary for different locations in the input space. An example of this setup is a space-time signal f encoding an RGB ($d_{\text{out}} = 3$) video with two spatial and a temporal dimension ($d_{\text{in}} = 3$), to be convolved with a kernel g that applies a foveated blur to each time slice of the video ($d = 2$). In the following derivations, we assume a spatially invariant kernel for ease of notation. Sec. 4.1.3 explains how our method can be easily extended to the spatially-varying case.

Formally, we seek to carry out the continuous convolutions

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^d} f(\mathbf{x} - \boldsymbol{\tau})g(\boldsymbol{\tau})d\boldsymbol{\tau}. \quad (1)$$

This integral operation does not have a closed-form solution for all but the most constrained sets of signals and/or kernels. In particular, it is unclear how this continuous operation can be applied to generic neural fields, which naturally only support point samples. The typical solution for these integrals is numerical approximation: For low-dimensional integration domains, quadrature rules are feasible, while the scalable gold standard in higher dimensions is Monte Carlo integration. The latter proceeds by sampling the integration domain and approximating the integral by a weighted sum of integrand evaluations:

$$(f * g)(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\tau}} [f(\mathbf{x} - \boldsymbol{\tau})g(\boldsymbol{\tau})] \approx \frac{1}{N} \sum_{\boldsymbol{\tau} \sim p} \frac{f(\mathbf{x} - \boldsymbol{\tau})g(\boldsymbol{\tau})}{p(\boldsymbol{\tau})}, \quad (2)$$

where \mathbb{E} is the expectation and $\boldsymbol{\tau}$ are now random samples drawn from the probability density function p . Unfortunately, a high number N of samples is required for large kernels g and/or high-frequency signals f , rendering this approach inefficient.

In the following, we develop a method that performs continuous convolutions in the form of Eq. 1, while only requiring a very low number of network evaluations, independent of the kernel size.

4 METHOD

We efficiently convolve a continuous signal f with a continuous kernel g by approximating g with a piecewise polynomial function, which becomes sparse after repeated differentiation. Our approach requires the evaluation of the repeated integral of f at a sparse set of sample positions dictated by the differentiated kernel (Sec. 4.1), leading to a substantial speed-up of the convolution operation. This general approach has first been studied by Heckbert [1986] in the context of discrete representations. Our method lifts the idea to the continuous neural setting by optimizing for a sparse differential representation of g (Sec. 4.2), and obtaining the repeated continuous integral of f by supervising on a minimal kernel using higher-order finite differences (Sec. 4.3). Once trained, any sparsity-optimized convolution kernel can be applied to f efficiently, without requiring additional input parameters. Our method supports spatially-varying convolutions in the form of continuously transformed kernels, leveraging the continuous nature of the representation. Fig. 3 gives an overview of our approach.

4.1 Convolution by Repeated Differentiation

Our approach requires two conceptual ingredients: First, the convolution operation in Eq. 1 reduces to a discrete sum if g consists

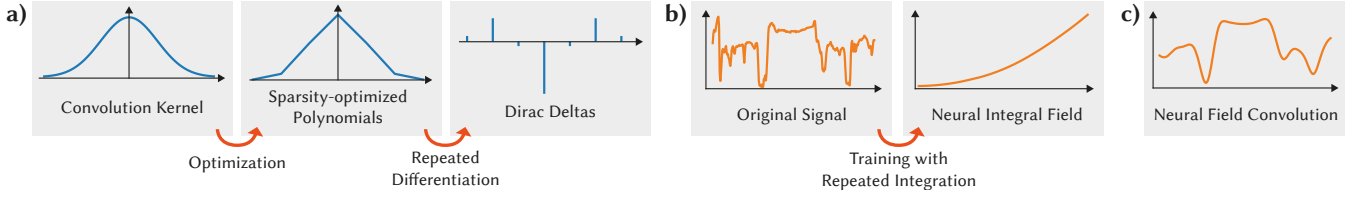


Fig. 3. Overview of our approach. *a)* Given an arbitrary convolution kernel, we optimize for its piecewise polynomial approximation, which under repeated differentiation yields a sparse set of Dirac deltas. *b)* Given an original signal, we train a neural field that captures the repeated integral of the signal. *c)* The continuous convolution of the original signal and the convolution kernel is obtained by a discrete convolution of the sparse Dirac deltas from *a)* and corresponding sparse samples of the neural integral field from *b)*.

of only Dirac deltas (Sec. 4.1.1). Second, the right-hand side of Eq. 1 can be transformed using repeated differentiation and integration (Sec. 4.1.2). Putting both ingredients together leads to an efficient discrete formulation of the continuous convolution operation (Sec. 4.1.3) involving a repeated integral field and a sparse differential kernel consisting of Dirac deltas [Heckbert 1986].

4.1.1 Dirac Kernels. As our first ingredient, consider a kernel g that is non-zero only at a small set of m locations in \mathbb{R}^d , i.e., g is

$$g(\mathbf{x}) = \sum_{i=1}^m \delta(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)}, \quad (3)$$

a sum of Dirac deltas δ , where $\mathbf{x}^{(i)} \in \mathbb{R}^d$ denotes the location and $w^{(i)} \in \mathbb{R}$ the magnitude¹ of the i 'th impulse. Then, by the *sifting* property of Dirac deltas, Eq. 1 simplifies to

$$(f * g)(\mathbf{x}) = \sum_{i=1}^m f(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)}, \quad (4)$$

i.e., we have reduced the computation from a continuous integral to a discrete sum – an efficient operation if m is small.

4.1.2 Convolutions with Differentiation and Integration. Our second ingredient is the following identity:

$$f * g = \left(\int f d\mathbf{x}_i \right) * \left(\frac{\partial}{\partial \mathbf{x}_i} g \right),$$

i.e., in order to convolve f with g we might as well convolve the antiderivative of f with the corresponding derivative of g . Applying this principle repeatedly yields [Heckbert 1986; Perlin 1984]

$$f * g = \underbrace{\left(\int \dots \int f d\mathbf{x}_1^n \dots d\mathbf{x}_d^n \right)}_{f^n} * \underbrace{\left(\frac{\partial^n}{\partial \mathbf{x}_1^n \dots \partial \mathbf{x}_d^n} g \right)}_{g^{-n}}. \quad (5)$$

Here, we sequentially differentiate g n times with respect to *each* of its dimensions. We denote this multidimensional repeated derivative as g^{-n} . For equality in Eq. 5 to hold, this pattern is mirrored for f , replacing differentiations with antiderivatives, where superscripts n denote repeated integrations along the individual dimensions. We denote the repeated multidimensional antiderivative of f as f^n . We refer to Heckbert [1986] for a proof of Eq. 5. Notice that input and

¹Technically, $\delta(0) = \infty$. But since a Dirac delta integrates to one, in the context of continuous convolutions, we refer to $w^{(i)}$ as “magnitudes” nevertheless.

output dimensions of f and g do not change after integration and differentiation.

4.1.3 Efficient Neural Field Convolutions. The central idea of our approach is to combine both ingredients presented above for the case of piecewise polynomial kernels \hat{g} . Concretely, we observe that piecewise polynomial functions turn into a sparse set of Dirac deltas after repeated differentiation [Heckbert 1986] (Fig. 4), i.e., \hat{g}^{-n} reduces to the form of Eq. 3. This implies that Eq. 4 can be used to perform a convolution with this kernel. Combining Eq. 4 and Eq. 5, our final convolution operation reads

$$(f * \hat{g})(\mathbf{x}) = \sum_{i=1}^m f^n(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)}. \quad (6)$$

Notice that this formulation requires only m evaluations of the repeated integral of f at locations dictated by the Dirac deltas of the differentiated kernel to yield the same result as the equivalent continuous convolution in Eq. 1.

The number of integrations and differentiations n directly depends on the desired order of the kernel polynomials, as detailed in Sec. 4.2. A disk-shaped kernel simulating thin-lens depth of field in an image can be approximated well using a piecewise *constant* function (corresponds to $n = 1$), while a Gaussian might require a piecewise *quadratic* approximation (corresponds to $n = 3$) to yield high-quality results with a low number of Dirac deltas.

Notice that our approach allows us to realize *spatially-varying* convolutions as well: The evaluation of Eq. 6 is independent for different evaluation locations \mathbf{x} . Therefore, we can make the choice of the convolution kernel \hat{g} a function of \mathbf{x} itself. In Sec. 4.2.2 we give details on how to obtain continuous parametric kernel families.

In summary, our method requires two components: *(i)* A piecewise polynomial kernel that results in a sparse set of Dirac deltas after repeated differentiation, and *(ii)* an efficient way to obtain and evaluate the repeated multidimensional integral of a continuous signal. These are detailed in Sec. 4.2 and Sec. 4.3, respectively.

4.2 Sparse Differential Kernels

Our approach requires a kernel g which, after repeated differentiation, results in a sparse set of Dirac deltas with positions $\mathbf{x}^{(i)}$ and magnitudes $w^{(i)}$:

$$g^{-n}(\mathbf{x}) = \sum_{i=1}^m \delta(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)}. \quad (7)$$

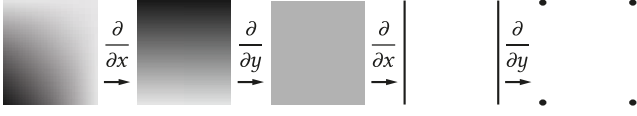


Fig. 4. Repeated differentiation of a bilinear patch ($d = n = 2$). After the first differentiation, the patch exhibits linear variation only along the vertical dimension. After subsequent differentiations, we obtain a constant patch, two vertical lines, and, finally, four Dirac deltas.

This property is satisfied for piecewise polynomial kernels of degree $n - 1$, which reduce to Dirac deltas positioned at the junctions between the segments after n differentiations per dimension [Heckbert 1986] (Fig. 4). Thus, given a kernel g , we seek to find its optimal piecewise polynomial approximation \hat{g} adhering to a user-specified budget of m Dirac deltas (Fig. 5).

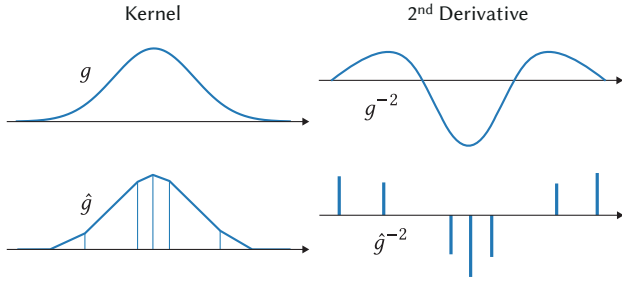


Fig. 5. Kernel representation in 1D for the case $n = 2$, i.e., a piecewise linear function. *Top row*: The original continuous kernel g has a continuous second derivative. *Bottom row*: We approximate g with a piecewise linear function \hat{g} , which reduces to a sparse set of Dirac deltas in its second derivative.

To parameterize \hat{g} , we utilize the linear structure of Eq. 7 and the linearity of differentiation: We consider the d -dimensional n -fold repeated antiderivative of the Dirac delta function

$$\delta^n(\mathbf{x}) = \begin{cases} \frac{\prod_{i=1}^d x_i^{n-1}}{(n-1)!} & \min_i x_i \geq 0 \\ \text{else} & \end{cases}$$

which is referred to as the n 'th-order *ramp* (Fig. 6). We now write our polynomial kernel \hat{g} as a linear combination of shifted ramps:

$$\hat{g}(\mathbf{x}) = \sum_{i=1}^m \delta^n(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)}. \quad (8)$$

Please note that Eq. 7 is the n 'th derivative of Eq. 8 by construction. Thus, we have established a parameterization of a C^{n-2} -continuous piecewise polynomial kernel \hat{g} , from which we can directly read off Dirac delta positions and magnitudes.

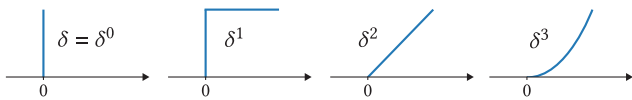


Fig. 6. 1D ramps of different orders (redrawn from Heckbert [1986]). Each ramp is the antiderivative of its predecessor.

We now optimize the following objective:

$$\min_{\mathbf{x}^{(i)}, w^{(i)}} \left[\mathbb{E}_{\mathbf{x} \in \mathbb{R}^d} [\|g(\mathbf{x}) - \hat{g}(\mathbf{x})\|_2^2] + \lambda \left| \sum_{i=1}^m w^{(i)} \right| \right]. \quad (9)$$

The first term encourages the solution to be close to the reference kernel on the entire domain. The second term steers the optimization to prefer solutions where the Dirac magnitudes sum to zero, which effectively enforces the kernel to be compact.

4.2.1 Optimization. We initialize the ramp positions $\mathbf{x}^{(i)}$ on a regular grid and their magnitudes $w^{(i)}$ to zero. We use the Adam [Kingma and Ba 2015] optimizer with standard parameters and set $\lambda = 0.1$ in all our experiments. In each iteration, we uniformly sample the continuous kernel domain. We observe that when the provided ramp budget m is too high, the magnitudes of individual ramps will approach zero, further increasing sparsity. We capitalize on this fact by monitoring ramp magnitudes in regular intervals. If an absolute magnitude falls below a small threshold, we remove the ramp from the mixture and continue optimizing. Separable kernels can be obtained by optimizing the respective 1D filters and combining them with an outer product. We provide timings of the optimization for different kernels in Supplemental appendix C.

We note the strong connection to spline-based approximations of functions [Ahlberg et al. 2016]. In the low-order regime under the continuity requirements we operate on, we find that our practical stochastic gradient descent-based approach yields high-quality results without the need for more elaborate techniques.

4.2.2 Kernel Transformations. Many applications of convolutions require kernels of different sizes and shapes, in particular in the case of spatially-varying convolutions. For example, foveated imagery or the simulation of depth-of-field require continuous and fine-grained control over the size of a blur kernel. Our approach supports on-the-fly kernel transformations without the need to sample and optimize entire parametric families of kernels by leveraging the continuous nature of the kernel and the signal.

Concretely, to continuously shift and (anisotropically) scale an optimized kernel \hat{g} using a matrix T , we simply apply T to the Dirac delta positions, i.e., $\mathbf{x}_T^{(i)} = T\mathbf{x}^{(i)}$. The updated Dirac delta magnitudes are given by $w_T^{(i)} = \frac{w^{(i)}}{\det(T)^n}$. Thus, we need to run the optimization for a kernel type only once in a canonical position and size, and obtain continuously transformed kernel instances at virtually no computational cost. Notice that, as a useful consequence, our approach enables continuous scale-space analysis [Lindeberg 2013; Witkin 1987], as illustrated in supplemental.

4.3 Neural Repeated Integral Field

To compute Eq. 6, we need to evaluate f^n , the n -th antiderivative of f as the second ingredient. We choose to implement f^n as a neural field \hat{f}^n . Ideally, it would hold that $\hat{f}^n = f^n$. This might be difficult to achieve without knowing an analytic form of the antiderivative. We could try Monte Carlo-estimating the antiderivative from the signal, leading to a loss like

$$\mathbb{E}_{\mathbf{x} \in \mathbb{R}^d} \left[\left\| \hat{f}^n(\mathbf{x}) - \mathbb{E}_{\tau \geq 0} [f(\mathbf{x} - \tau)] \right\| \right]. \quad (10)$$

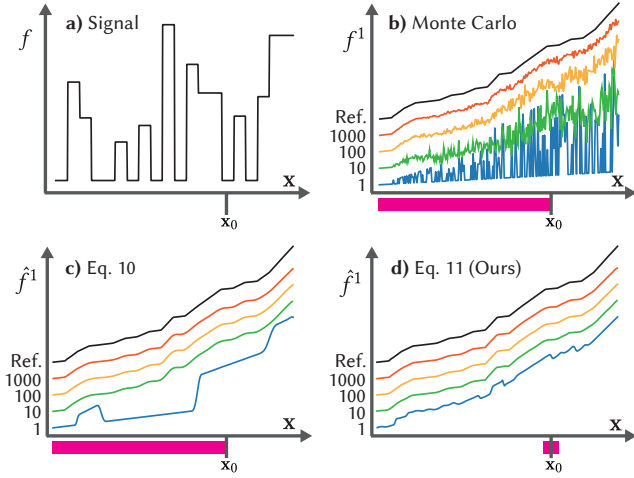


Fig. 7. (a): Given a signal f , we are interested in finding its antiderivative f^1 . (b): A scalable way to obtain the antiderivative is Monte Carlo estimation $\mathbb{E}_{\tau \geq 0} [f(x - \tau)]$. Increasing the number of samples (different colors) gets us closer to the true solution (black). (c): Using the estimates from (b) to supervise the training of a network \hat{f}^1 as per Eq. 10 requires many samples, while the regressed antiderivative remains blurry. (d): In contrast, our approach only requires a convolution with a small kernel (Eq. 11) to yield high-quality results, including sharp features, with only a low number of Monte Carlo samples. For each method, a pink bar marks the region that needs to be considered for estimating/training the antiderivative value at location x_0 . The different graphs in (b)-d) use different constants of integration for improved visualization.

The inner expectation would sum over the entire half-domain $\tau \geq 0$, leading to a high variance and a low-quality \hat{f}^n . An example of this is shown in Fig. 7, where the input is a 1D HDR signal (Fig. 7a). To estimate the antiderivative at x_0 , we have to sample the entire pink area in Fig. 7b, resulting in significant variance, even if the sample count increases. When using these estimates to train \hat{f}^n , this variance leads to a low-quality, blurry regression (Fig. 7c), as no loss function is known that properly captures the statistical properties of Monte Carlo noise [Lehtinen et al. 2018].

For our purpose, convolution, what really needs to hold, is that $\hat{f}^n * h_n^{-n} = f * h_n$, for any piecewise polynomial kernel h_n of degree $n - 1$. The resulting loss to achieve this is

$$\mathbb{E}_{\mathbf{x} \in \mathbb{R}^{d_{in}}} \left\| \sum_{i=1}^m \hat{f}^n(\mathbf{x} - \mathbf{x}^{(i)}) w^{(i)} - \mathbb{E}_{\tau \in \text{supp}(h_n)} [f(\mathbf{x} - \tau) h_n(\tau)] \right\|, \quad (11)$$

where $\mathbf{x}^{(i)}$ and $w^{(i)}$ are the Dirac delta positions and magnitudes of h_n^{-n} . The first term is due to Eq. 6 and the second term is a Monte Carlo estimate of the convolved signal.

For efficient training and a high-quality antiderivative, it is crucial for h_n to be very compact: First, the right part of Eq. 11 becomes a tame interval, significantly reducing variance and thus enabling training with a low number of Monte Carlo samples (Fig. 7d). Second, it prevents \hat{f}^n from “cheating” by not learning the antiderivative of the signal but the antiderivative of a convolved signal. However,

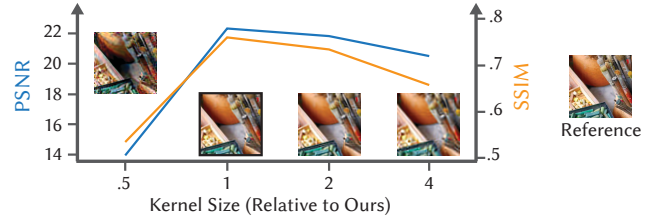


Fig. 8. Antiderivative quality as a function of kernel size used in Eq. 11. We measure and display quality by repeated automatic differentiation of the learned antiderivatives. We see that our *minimal kernel* is optimal in terms of quality. Smaller kernels lead to instabilities and larger ones to blur.

we cannot reduce the support of h_n arbitrarily, as the left part of Eq. 11 tends to result in instabilities, as the distances between the $\mathbf{x}^{(i)}$ shrink. Fig. 8 illustrates the inherent trade-off of this situation: There exists a sweet spot for the kernel size, producing the highest-quality antiderivatives. Smaller kernels lead to training instabilities, larger kernels to blur. We refer to the optimal solution as the *minimal kernel*.

While we could use any filter shape as minimal kernel h_n , we choose the n -fold convolution of a box with itself (Fig. 9). This has the advantage that the left part of the loss Eq. 11 becomes a sum over only $m = (n + 1)^d$ elements that is efficient to compute, corresponding to higher-order finite differences.

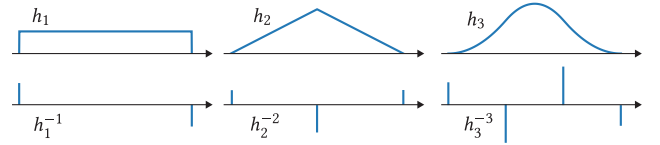


Fig. 9. Minimal piecewise polynomial kernels of different degrees (top row) and their corresponding Dirac deltas (bottom row).

4.4 Implementation Details

All source code is accessible via <https://neural-fields-conv.mpi-inf.mpg.de>. We have implemented our prototype within the PyTorch [Paszke et al. 2017] environment. Our integral fields are realized using MLPs, where exact architectures vary slightly depending on the modality to be represented, as detailed in appendix B of the supplemental document.

For training our repeated integral field, we again use the Adam [Kingma and Ba 2015] optimizer with standard parameters. We handle boundaries by mirror-padding the training signal. Considering a unit domain, we train with a size of 0.025 for the minimal kernel h_n until convergence, followed by a fine-tuning on a kernel of size 0.0125. Empirically, we found this size to produce the highest-quality antiderivatives (Fig. 8). As the kernel size cannot be further decreased, our repeated integral field \hat{f}^n is a slightly low-pass filtered version of f^n , resulting in a lower limit of filter sizes our convolutions can faithfully compute. Fortunately, these small filters are highly amenable to efficient Monte Carlo estimation. Thus, at test time, whenever a kernel size falls below the threshold, we Monte-Carlo-estimate the convolution.

Table 1. Settings for all applications.

Modality	Input			Kernel					m^b
	d_{in}	d_{out}	Format	Shape	SV ^a	d	Order		
Images	2	3	Grid	Gauss	×	2	Linear	169	
	2	3	Grid	DoG	×	1	Linear	13	
	2	3	Grid	Circle	✓	2	Const.	141	
Bilat. Images	3	3	Grid	Gauss	✓	3	Const.	343	
Video	3	3	Grid	Tent	×	1	Linear	3	
Geometry	3	1	SDF	Box	✓	3	Const.	8	
Animation	1	69	Paths	Gauss	×	1	Linear	13	
Audio	1	1	Wave	Box	×	1	Const.	2	

^a Experiments with spatially-varying kernels.

^b The number of Diracs automatically adapts to the kernel as described in Sec. 4.2.1.

Convolution per Eq. 6 can be efficiently implemented as an augmented neural field based on \hat{f}^n , by prepending positional offsets $\mathbf{x}^{(i)}$ and appending a linear layer containing $w^{(i)}$.

5 APPLICATIONS

To demonstrate the generality and efficiency of our approach, we consider five signal modalities: images, videos, geometry, character animations, and audio. An overview of settings for all modalities is given in Tab. 1. References are computed using Monte Carlo estimation as per Eq. 2 until convergence. Further, we consider the following baselines:

INSP [Xu et al. 2022]. Similar to our approach, INSP relies on higher-order derivatives but uses them in a point-wise fashion to reason about local neighborhoods, reminiscent of a Taylor expansion. This method has to be trained for each convolution kernel separately, while our integral fields can be used with any kernel of a fixed polynomial degree. We follow their original implementation and provide all second-order partial derivatives. We found that providing more derivatives did not markedly improve their results.

BACON [Lindell et al. 2022] and *PNF* [Yang et al. 2022]. While our method supports arbitrary filters at test time, both BACON and PNF are limited to a discrete cascade of k specific and fixed intermediate network outputs with different frequency contents, i.e., a set of pre-filtered versions of the signal. To compare ours to BACON and PNF, we approximate the convolution with an arbitrary filter as a linear combination of their intermediate outputs: Given the reference result, we optimize for a set of weights that when multiplied with corresponding intermediate outputs is closest to the target. Notice that this procedure requires a reference, while ours does not. For PNF, we observed that finer-scale intermediate outputs are not zero-mean. We compensate for this by subtracting the mean from all intermediate outputs and adding the sum of these means to the coarsest output. This way, finer levels only add higher-frequency details to the solution, but no global color shifts.

5.1 Images

In this application, we consider (high dynamic range) RGB images $f \in \mathbb{R}^2 \rightarrow \mathbb{R}^3$ as signals.



Fig. 10. Gaussian 2D image blur of the input signal, with increasing bandwidth. It can be seen how our approach works also for large kernels.

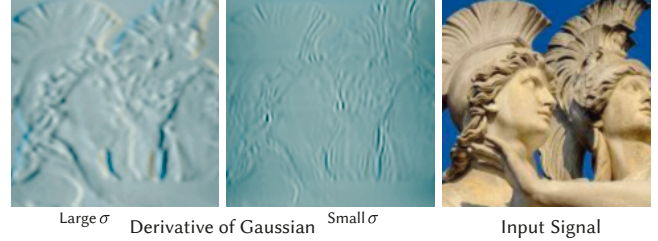


Fig. 11. Derivative-of-Gaussian filtering 2D result. Note, that our approach supports such non-convex filters, producing signed results.

Table 2. Image quality comparison for Gaussian kernels.

	$\sigma = 0.04$			$\sigma = 0.05$			$\sigma = 0.07$		
	PSNR	LPIPS	SSIM	PSNR	LPIPS	SSIM	PSNR	LPIPS	SSIM
INSP	27.2	0.305	0.774	25.8	0.398	0.713	23.9	0.535	0.622
BACON	35.2	0.079	0.946	33.6	0.091	0.935	30.6	0.139	0.904
PNF	34.5	0.072	0.938	33.9	0.086	0.935	32.0	0.132	0.922
Ours	28.6	0.076	0.934	31.0	0.042	0.963	34.1	0.032	0.98

Linear filtering. We show results for a Gaussian image filter in Fig. 10 and a derivative-of-Gaussian filter in Fig. 11. More results can be found in supplemental. Table 2 and Fig. 12 evaluate image quality on a set of 50 images. Since no other method can produce competitive results for large kernels, to facilitate an insightful analysis, we limit our numerical evaluation to the small-kernel regime. We see that for very small kernels, BACON and PNF tend to produce higher-quality results, while we significantly outperform all methods as the kernel size increases.

Non-linear filtering. Our approach can also be used for non-linear filtering, such as bilateral filtering [Tomasi and Manduchi 1998]. We implement this by optimizing for the repeated integral field of a bilateral grid [Chen et al. 2007], which augments a 2D image with an additional signal dimension, reducing filtering to a linear operation in this extended space. We see in Fig. 13 that our approach is able to faithfully produce edge-aware filtering results.

Spatially-varying filtering. is demonstrated in Fig. 1 and Fig. 14. In these cases, the strength of a circular blur filter is determined by a spatially-varying auxiliary signal in the form of a depth buffer.



Fig. 12. A qualitative comparison between INSP [Xu et al. 2022], BACON [Lindell et al. 2022], PNF [Yang et al. 2022], and our approach using a Gaussian kernel. INSP approach does hardly perform any filtering, when the kernel is larger, such as our approach supports, but suffers from noise. BACON fares better, but shows ringing at high-contrast edges such as the house against the sky, while PNF also struggles to produce low-pass filtering results.

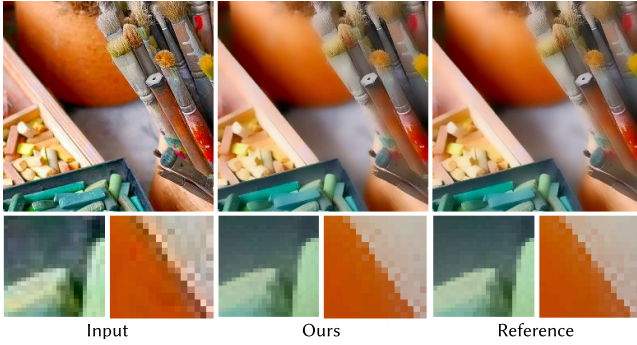


Fig. 13. Non-linear (bilateral) filtering of an input image (left) by our method (middle) and the reference (right).



Fig. 14. Depth-of-field applied to a synthetic image with a depth map.

5.2 Videos

Videos are fields $f \in \mathbb{R}^3 \rightarrow \mathbb{R}^3$, mapping 2D location and time to RGB color. In Fig. 15, we apply smoothing with a tent filter along the time dimension to create appealing non-linear motion blur. We refer to our supplemental material for a more extensive evaluation.

5.3 3D Geometry

Surfaces can be modeled using signed distance functions (SDFs) $f \in \mathbb{R}^3 \rightarrow \mathbb{R}$. We apply 3D box filters with different side lengths σ to an SDF, resulting in a progressively smoothed surface. We again compare our approach to INSP and BACON (no code is available to apply PNF to the 3D case) in Fig. 16 and Table 3, where numerical evaluations are averaged across the three 3D objects studied in INSP. As metrics, we consider the mean squared error (MSE) of the SDF, the intersection over union (IoU), as well as the Chamfer distance of the reconstructed surface. We observe that we outperform INSP and

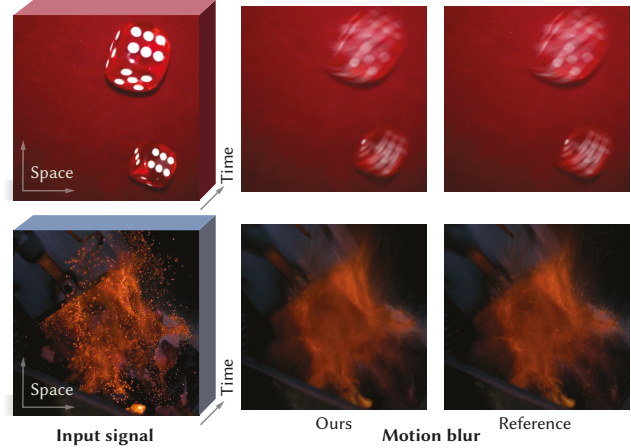


Fig. 15. We apply our approach to a 3D space-time HDR field (video, seen left) to filter along the time axis with a tent kernel. The resulting motion blur (middle) compares favorably to the reference.

Table 3. Quality comparisons of filtered SDFs using 3D box kernels.

	$\sigma = 0.05$			$\sigma = 0.15$		
	MSE	Cham.	IoU	MSE	Cham.	IoU
INSP	292.50000	171.77	0.90	281.30000	388.26	0.73
BACON	0.40145	360.90	0.82	1.85791	932.84	0.61
Ours	0.00109	20.97	0.99	0.00026	13.24	0.99

BACON across all kernel sizes and metrics. More qualitative results can be found in supplemental. Additionally, Fig. 17 demonstrates a spatially-varying SDF filtering result.

5.4 Animation

We consider the task of filtering a neural field representation of a motion-capture sequence, which contains a significant amount of noise. Our test sequence consists of 23 3D joint position paths over time, resulting in a field $f \in \mathbb{R} \rightarrow \mathbb{R}^69$. In Fig. 18 and the supplemental, we show the result of applying a Gaussian filter to the noisy animation data, resulting in smooth motion trajectories.

5.5 Audio

Finally, we apply our framework to the task of filtering an audio signal $f \in \mathbb{R} \rightarrow \mathbb{R}$, available for listening in the supplemental.



Fig. 16. Two geometric shapes represented by an SDF (left) are filtered with a box kernel ($\sigma = 0.05$). While INSP, in the second column, suffers from noise, and BACON, in the third column, cannot reproduce larger-scale filtering, our result is close to the MC reference.



Fig. 17. Spatially-varying blur from two views computed using our method.

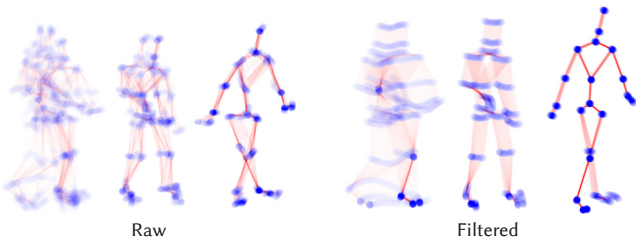


Fig. 18. A noisy motion-capture sequence (left) is filtered using our approach to yield smooth motion trajectories (right).

6 ANALYSIS

Here, we analyze further individual aspects of our approach.

Repeated Integral Fields. We seek to gain more insights into our learned integral fields. To this end, we consider the 2D image case (Sec. 5.1) both for single and double integrals per dimension, as required for convolutions with piecewise constant and linear kernels, respectively. In Tab. 4 we compare our antiderivatives against AutoInt [Lindell et al. 2021]. As the original implementation does not support integration with respect to more than one variable, we re-implemented this baseline using the `functorch` library for repeated differentiation. We compute the mean squared error (MSE) between the original signal and the obtained integral fields after repeated automatic differentiation, averaged over three images. We do not compare integrals directly, as their absolute values are dominated by higher-order constants of integration. Further, we measure the time required to train the fields, as well as the size of the network during training in terms of the number of nodes. As there is no straightforward procedure to count the number of nodes of the

Table 4. Integral field evaluation.

Int. Op.	Method	MSE ($\times 10^{-3}$)	Time	Graph Size
\iint	AutoInt	6.83	1.3h	339
	Ours	6.48	1.1h	11
$\int^2 \int^2$	AutoInt	7.71	14.7h	15,407
	Ours	7.28	1.2h	11

repeated-derivative graphs, we use the official AutoInt implementation for this particular calculation and differentiate two and four times with respect to *one* input variable, to get a good approximation of the two cases studied. Corresponding graphs are visualized in Fig. 19.

We see that our approach produces higher-quality antiderivatives than AutoInt while taking significantly less time to train, in particular for higher-order integrals. Further, our network size during training is independent of the integration order, while the computational graphs of AutoInt grow quickly due to the symbolic differentiation required.

We are further interested in how the quality of the learned antiderivative affects convolution quality. For this analysis, we consider antiderivative MSE as above (lower is better), and convolution quality in terms of PSNR (higher is better). We find the two measures highly correlated: Pearson's $R = -0.98$.

In appendix C of the supplemental document, we study the accuracy of our optimized kernels.

Comparison to equal-effort MC. While we use *converged* Monte Carlo estimates of convolutions (Eq. 2) as references in our experiments, we are also interested in the quality of such an estimate when reducing the number of samples to the number of Dirac deltas we use in our method, providing an equal-effort comparison. In Fig. 20 we show an illustration of this analysis. We observe that MC suffers from extensive noise, while our solution is smooth.

7 DISCUSSION AND CONCLUSION

We have presented a novel approach to perform general, spatially-varying convolutions in continuous signals. Capitalizing on the fact that piecewise polynomial kernels become sparse after repeated

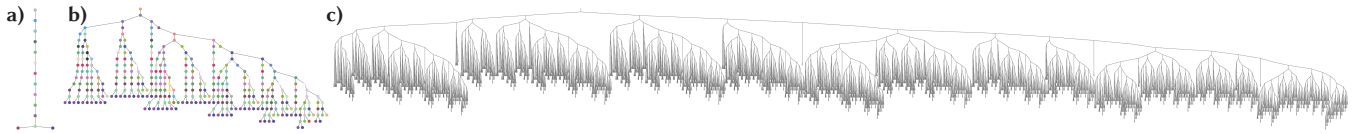


Fig. 19. Comparison of different graph sizes. *a*) Our method retains a small graph independent of integration order. *b*) The AutoInt graph after two differentiations (first Int. Op. in Tab. 4). *c*) The AutoInt graph after four differentiations (second Int. Op. in Tab. 4). Same colors represent same operations.

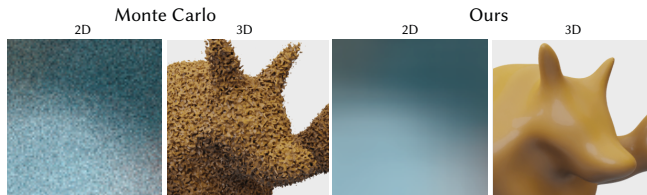


Fig. 20. Ours vs. an equal-effort Monte Carlo estimate of a convolution.

differentiation, we only require a small number of integral-network evaluations to perform large-scale continuous convolutions.

Since our work is one of the first steps in this direction, there is ample opportunity for future work. Currently, one of our biggest limitations is that we need access to the entire signal to train our repeated integral field. This prevents the treatment of signals that are only partially observed through a differentiable forward map, e.g., as prominently is the case for neural radiance fields [Mildenhall et al. 2020].

Our integral fields are trained using generalized finite differences, which we found to become unstable for small kernels (Sec. 4.4). This naturally imposes an upper frequency limit on the learned antiderivatives. Fortunately, this becomes noticeable only for small kernels, in which case we resort to Monte Carlo sampling of the convolution, which is efficient in this condition.

Our assumption is that kernels to be used for spatially-varying filtering are (anisotropically) scaled versions of a reference kernel, which, arguably, covers a broad range of applications. We do not have a scalable solution for situations that violate this assumption – except for special problems such as bilateral filtering. One solution could be blending between the Dirac deltas of a set of pre-computed reference kernels. Further, kernel transformations are limited to axis-aligned operations. We envision that re-parameterizations leveraging the continuous nature of the representation might be able to lift this restriction.

In this work, we do not claim superiority over operating on a grid, if enough space is available to represent the signal that way. But if one needs to use a neural field, e.g. for extreme compression, continuous queries, or other advantages (Sec. 1), an approach such as ours is needed. We hope to inspire future work on signal processing in continuous neural representations to help them reach their full potential.

REFERENCES

J Harold Ahlberg, Edwin Norman Nilson, and Joseph Leonard Walsh. 2016. *The Theory of Splines and Their Applications*. Vol. 38. Elsevier.
Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. 2021. Mip-NeRF: A Multiscale Representation for

Anti-Aliasing Neural Radiance Fields. *ICCV* (2021).
Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. 2022. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. *CVPR* (2022).
E Oran Brigham. 1988. *The fast Fourier transform and its applications*. Prentice-Hall, Inc.
Jiawen Chen, Sylvain Paris, and Frédo Durand. 2007. Real-time edge-aware image processing with the bilateral grid. *ACM Trans. Graph.* 26, 3 (2007), 1–9.
Franklin C Crow. 1984. Summed-area tables for texture mapping. In *SIGGRAPH*. 207–212.
Yilun Du, M. Katherine Collins, B. Joshua Tenenbaum, and Vincent Sitzmann. 2021. Learning Signal-Agnostic Manifolds of Neural Fields. In *NeurIPS*.
Emilien Dupont, Adam Goliński, Milad Alizadeh, Yee Whye Teh, and Arnaud Doucet. 2021. Coin: Compression with implicit neural representations. In *ICLR (Neural Compression Workshop)*.
Emilien Dupont, Hyunjik Kim, S. M. Ali Eslami, Danilo Jimenez Rezende, and Dan Rosenbaum. 2022a. From data to functa: Your data point is a function and you can treat it like one, Vol. 162. PMLR, 5694–5725.
Emilien Dupont, Yee Whye Teh, and Arnaud Doucet. 2022b. Generative Models as Distributions of Functions, Vol. 151. PMLR, 2989–3015.
Ziya Erkoç, Fangchang Ma, Qi Shan, Matthias Nießner, and Angela Dai. 2023. Hyperdiffusion: Generating implicit neural fields with weight-space diffusion. In *ICCV*.
Zeev Farberman, Raanan Fattal, and Dani Lischinski. 2011. Convolution pyramids. *ACM Trans. Graph.* 30, 6 (2011), 1–8.
Rizal Fathony, Anit Kumar Sahu, Devin Willmott, and J Zico Kolter. 2020. Multiplicative filter networks. In *ICLR*.
Alain Fournier and Eugene Fiume. 1988. Constant-time filtering with space-variant kernels. *ACM Trans. Graph.* 22, 4 (1988), 229–238.
William T Freeman, Edward H Adelson, et al. 1991. The design and use of steerable filters. *IEEE TPAMI* 13, 9 (1991), 891–906.
Paul S Heckbert. 1986. Filtering by repeated integration. *ACM Trans. Graph.* 20, 4 (1986), 315–321.
Pedro Hermosilla, Tobias Ritschel, Pere-Pau Vázquez, Alvar Vinacua, and Timo Ropinski. 2018. Monte carlo convolution for learning on non-uniformly sampled point clouds. *ACM Trans. Graph.* 37, 6 (2018), 1–12.
Amir Hertz, Or Perel, Raja Giryes, Olga Sorkine-Hornung, and Daniel Cohen-Or. 2021. Sape: Spatially-adaptive progressive encoding for neural optimization. *NeurIPS* 34 (2021), 8820–8832.
Brian K. S. Isaac-Medina, Chris G. Willcocks, and Toby P. Breckon. 2023. Exact-NeRF: An Exploration of a Precise Volumetric Parameterization for Neural Radiance Fields. *CVPR* (2023).
Markus Kettunen, Marco Manzi, Miika Aittala, Jaakko Lehtinen, Frédo Durand, and Matthias Zwicker. 2015. Gradient-domain path tracing. *ACM Trans. Graph.* 34, 4 (2015), 1–13.
Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR*.
Georgios Kopanas, Thomas Leimkühler, Gilles Rainer, Clément Jambon, and George Drettakis. 2022. Neural Point Catacaustics for Novel-View Synthesis of Reflections. *ACM Trans. Graph.* 41, 6 (2022), 1–15.
Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. 2018. Noise2Noise: Learning Image Restoration without Clean Data. In *ICML*. 2965–2974.
Thomas Leimkühler, Hans-Peter Seidel, and Tobias Ritschel. 2018. Laplacian Kernel Splatting for Efficient Depth-of-field and Motion Blur Synthesis or Reconstruction. *ACM Trans. Graph.* 37, 4 (2018), 1–11.
Tony Lindeberg. 2013. *Scale-space theory in computer vision*. Vol. 256. Springer Science & Business Media.
David B Lindell, Julien NP Martel, and Gordon Wetzstein. 2021. AutoInt: Automatic integration for fast neural volume rendering. In *CVPR*. 14556–14565.
David B Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. 2022. Bacon: Band-limited coordinate networks for multiscale scene representation. In *CVPR*. 16252–16262.
Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2020. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*. 405–421.

- Thomas W. Mitchel, Benedict Brown, David Koller, Tim Weyrich, Szymon Rusinkiewicz, and Michael Kazhdan. 2020. *Efficient Spatially Adaptive Convolution and Correlation*. Technical Report 2006.13188. arXiv preprint.
- Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. *ACM Trans. Graph.* 41, 4 (2022), 1–15.
- Harald Niederreiter. 1992. Low-discrepancy point sets obtained by digital constructions over finite fields. *Czechoslovak Mathematical Journal* 42, 1 (1992), 143–166.
- Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. 2019. DeepSDF: Learning continuous signed distance functions for shape representation. In *CVPR*, 165–174.
- Keunhong Park, Utkarsh Sinha, Jonathan T Barron, Sofien Bouaziz, Dan B Goldman, Steven M Seitz, and Ricardo Martin-Brualla. 2021. Nerfies: Deformable neural radiance fields. In *ICCV*. 5865–5874.
- Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch. (2017).
- Kenneth Perlin. 1984. Personal communication with Paul Heckbert, mentioned in Heckbert [1986].
- Prajit Ramachandran, Barret Zoph, and Quoc V Le. 2017. Searching for activation functions. *arXiv preprint arXiv:1710.05941* (2017).
- Nicholas Sharp and Alec Jacobson. 2022. Spelunking the Deep: Guaranteed Queries on General Neural Implicit Surfaces via Range Analysis. *ACM Trans. Graph.* 41, 4 (2022), 1–16.
- Assaf Shocher, Ben Feinsein, Niv Haim, and Michal Irani. 2020. From discrete to continuous convolution layers. *arXiv preprint arXiv:2006.11120* (2020).
- Patrice Simard, Léon Bottou, Patrick Haffner, and Yann LeCun. 1998. Boxlets: a fast convolution algorithm for signal processing and neural networks. *NeurIPS* 11 (1998).
- Gurprit Singh, Cengiz Öztireli, Abdalla GM Ahmed, David Coeurjolly, Kartic Subr, Oliver Deussen, Victor Ostromoukhov, Ravi Ramamoorthi, and Wojciech Jarosz. 2019. Analysis of sample correlations for Monte Carlo rendering. In *Comp. Graph. Forum*, Vol. 38. Wiley Online Library, 473–491.
- Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. 2020. Implicit neural representations with periodic activation functions. *NeurIPS* 33 (2020), 7462–7473.
- Vincent Sitzmann, Michael Zollhöfer, and Gordon Wetzstein. 2019. Scene Representation Networks: Continuous 3D-Structure-Aware Neural Scene Representations. In *NeurIPS*.
- Ilya Meerovich Sobol. 1967. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7, 4 (1967), 784–802.
- Kenneth O Stanley. 2007. Compositional pattern producing networks: A novel abstraction of development. *Genetic programming and evolvable machines* 8 (2007), 131–162.
- Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. 2020. Fourier features let networks learn high frequency functions in low dimensional domains. *NeurIPS* 33 (2020), 7537–7547.
- Ayush Tewari, Justus Thies, Ben Mildenhall, Pratul Srinivasan, Edgar Tretschk, W Yifan, Christoph Lassner, Vincent Sitzmann, Ricardo Martin-Brualla, Stephen Lombardi, et al. 2022. Advances in neural rendering. *Comp. Graph. Forum* 41, 2 (2022), 703–735.
- Carlo Tomasi and Roberto Manduchi. 1998. Bilateral filtering for gray and color images. In *ICCV*. 839–846.
- Edgar Tretschk, Ayush Tewari, Vladislav Golyanik, Michael Zollhöfer, Christoph Lassner, and Christian Theobalt. 2021. Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In *ICCV*. 12959–12970.
- Cristina Vasconcelos, Kevin Swersky, Mark Matthews, Milad Hashemi, Cengiz Öztireli, and Andrea Tagliasacchi. 2023. CUF: Continuous Upsampling Filters. In *CVPR*.
- Delio Vicini, Sébastien Speierer, and Wenzel Jakob. 2022. Differentiable signed distance function rendering. *ACM Trans. Graph.* 41, 4 (2022), 1–18.
- Paul Viola and Michael Jones. 2001. Rapid object detection using a boosted cascade of simple features. In *CVPR*, Vol. 1. I–I.
- Shenlong Wang, Simon Suo, Wei-Chiu Ma, Andrei Pokrovsky, and Raquel Urtasun. 2018. Deep parametric continuous convolutional neural networks. In *CVPR*. 2589–2597.
- Yinhuai Wang, Shuzhou Yang, Yujie Hu, and Jian Zhang. 2022. NeRFocus: Neural Radiance Field for 3D Synthetic Defocus. *arXiv preprint arXiv:2203.05189* (2022).
- Lance Williams. 1983. Pyramidal parametrics. In *SIGGRAPH*, Vol. 17. 1–11.
- Andrew P Witkin. 1987. Scale-space filtering. In *Readings in Computer Vision*. 329–332.
- Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. 2022. Neural fields in visual computing and beyond. *Comp. Graph. Forum* 41, 2 (2022), 641–676.
- Dejia Xu, Peihao Wang, Yifan Jiang, Zhiwen Fan, and Zhangyang Wang. 2022. Signal Processing for Implicit Neural Representations. In *NeurIPS*.
- Guandao Yang, Serge Belongie, Bharath Hariharan, and Vladlen Koltun. 2021. Geometry processing with neural fields. *NeurIPS* 34 (2021), 22483–22497.
- Guandao Yang, Sagie Benaim, Varun Jampani, Kyle Genova, Jonathan Barron, Thomas Funkhouser, Bharath Hariharan, and Serge Belongie. 2022. Polynomial neural fields for subband decomposition and manipulation. *NeurIPS* 35 (2022), 4401–4415.
- Yu-Jie Yuan, Yang-Tian Sun, Yu-Kun Lai, Yuewen Ma, Rongfei Jia, and Lin Gao. 2022. NeRF-editing: geometry editing of neural radiance fields. In *ICCV*. 18353–18364.
- Xian-Da Zhang. 2022. Modern signal processing. In *Modern Signal Processing*. De Gruyter.

Table 5. Different convolution methods. n is the size of the filter, m the size of the signal (samples or weights to represent it), and d the signal dimension.

	Time	Spat. vary	Noisy	Cont.
Classic	$O(m \times n^d)$	✓	×	×
Fourier	$O(m \times \log(m) \times d)$	×	×	×
Monte Carlo	$O(m \times n)$	✓	✓	✓
SAT	$O(m \times d)$	✓	×	×
Mip-NeRF	$O(m)$	✓	×	✓
INSP	$O(m)$	×	×	✓
Ours	$O(m)$	✓	×	✓

Table 6. Architecture details of our integral fields.

Application	#Layers	#Features	#Trainable Params.
Images ¹	5	256	270,851
Images ²	5	512	1,065,987
Videos	9	256	534,019
Geometry	5	256	270,851
Audio	5	256	270,851
Animation	5	256	270,851

¹ Low-resolution (256x256) images used for large-scale comparisons.² High-resolution (3000x3000) images used for displaying results.

Table 7. Accuracy of kernels and time to optimize them.

m	3		7		13		24	
	1	2	1	2	1	2	1	2
MSE	1.6e-1	1.5e-2	1.5e-2	7.9e-4	4.0e-3	7.6e-5	1.1e-3	2.3e-5
Time (s)	3	25	4	60	6	75	9	132

A CONVOLUTION METHODS

In Tab. 5 we compare different solutions to the convolution problem.

B INTEGRAL FIELD MODEL DETAILS

In Tab. 6 we give details of network architectures for the repeated integral fields we use per application. In all cases we use a multi-layer perceptron (MLP). Similar to Lindell et al. [2021], we observed best results with Swish [Ramachandran et al. 2017] activation functions. We report the number of hidden layers, the number of features per layer, and the resulting number of trainable parameters.

C KERNELS

In Tab. 7 we provide details on our optimized kernels. Here, we consider a 1D Gaussian kernel, represented by different numbers m of Diracs, using different orders of differentiation n . We give the reconstruction error in terms of the mean squared error (MSE) and the time (in seconds) our unoptimized implementation takes to obtain a converged result.