torch-projectors: A High-Performance Differentiable Projection Library for PyTorch

Dimitry Tegunov
tegunovd@gene.com

August 12, 2025

Abstract

1 Introduction

2 Methods

2.1 Data Conventions

`torch-projectors` operates exclusively in Fourier space using PyTorch's RFFT format, following FFTW conventions. For rotations to work correctly, reconstruction and projection data must be shifted (fftshift) in real space to place their center at the 0th tensor element before FFT. Results must be inverse-shifted (ifftshift) in real space to restore conventional centering.

All spatial dimensions must be square (2D) or cubic (3D) with even sizes. 2D reconstructions use shape $[B, N, N/2+1]$ where B is the batch dimension. 3D reconstructions use shape $[B, D, H, W/2+1]$ where the spatial dimensions are equal and even. Projections follow the same RFFT convention with shape $[B, P, N, N/2+1]$, where B is the reconstruction batch dimension, and P is the number of poses.

Rotations are specified as matrices: $[B, P, 2, 2]$ for 2D and $[B, P, 3, 3]$ for 3D. Translation shifts use shape $[B, P, 2]$ and are applied via phase modulation. A positive shift value moves the image contents along the positive direction of the respective axis. In back-projection, the shift is applied in the opposite direction to ensure it is the inverse of the forward projection with the same parameters. Batch broadcasting is supported where B can be 1 or match the reconstruction batch size B; or the P dimension in either rotations or shifts (but not both) can be 1 to use the same value over all poses. This enables efficient processing of multiple reconstructions with shared or individual poses.

2.2 Interpolation

`torch-projectors` supports linear and cubic interpolation methods for sampling Fourier-space data at non-integer coordinates. Linear interpolation uses standard multilinear kernels: bilinear for 2D operations (4-point support) and trilinear for 3D operations (8-point support).

Cubic interpolation employs separable Catmull-Rom kernels with parameter $a = -0.5$, providing C^1 continuity and exact interpolation through control points. The kernel function is defined as:

$$w(s) = \begin{cases} (a+2)|s|^3 - (a+3)|s|^2 + 1 & \text{if } |s| \leq 1 \\ a|s|^3 - 5a|s|^2 + 8a|s| - 4a & \text{if } 1 < |s| \leq 2 \\ 0 & \text{if } |s| > 2 \end{cases} \tag{1}$$

For bicubic interpolation, a 4×4 neighborhood is sampled around each coordinate, while tricubic interpolation samples a $4 \times 4 \times 4$ neighborhood. The final interpolated value is computed as the separable product of 1D kernel evaluations along each dimension.

All interpolation operations support oversampling, where coordinates are scaled by a factor > 1 to sample from reconstructions that were previously zero-padded in real space to increase the Fourier space sampling rate. This improves interpolation accuracy at constant computational load, but at the expense of increased memory usage. The sampling grid is sparse, matching the box size of the unpadded reconstruction. Because of this, no additional real-space cropping is required for the final projections.

A hard low-pass filter can be applied during the projection operation.

Shifts can be applied as part of the forward and backward projection operations. This is achieved by applying a phase modulation to the complex Fourier space components according to the Fourier shift theorem: $F\{f(\mathbf{x}-\mathbf{s})\} = F\{f(\mathbf{x})\} \cdot e^{-i2\pi\mathbf{k}\cdot\mathbf{s}}$, where $\mathbf{k}$ are the Fourier coordinates and $\mathbf{s}$ is the shift vector.

2.3 Friedel Symmetry

The Fourier transform of a real-valued function exhibits Friedel symmetry, characterized by the relationship:

$$F(-\mathbf{k}) = F^*(\mathbf{k}) \tag{2}$$

where $F^*(\mathbf{k})$ denotes the complex conjugate at frequency $\mathbf{k}$. This symmetry property allows PyTorch's RFFT format to store only the non-negative frequency components along the last dimension, yielding the characteristic $[\dots, N/2+1]$ shape.

However, the zero-frequency axis (where $k_c = 0$) requires careful handling during projection operations to maintain mathematical consistency and prevent double-counting:

Forward Pass: Standard Friedel symmetry lookup is applied when sampling negative k_c coordinates. For $k_c < 0$, the implementation accesses $F(-k_c, -k_r)$ and returns its complex conjugate $F^*(-k_c, -k_r)$ according to equation 2.

Backward Pass: To avoid over-representation in gradient accumulation, the implementation skips processing the negative half of the zero-frequency axis (specifically, components where $k_c = 0$ and $k_r < 0$), since these are present twice in the RFFT-formatted gradient.

Accumulation on Zero-Frequency Axis: When projection operations result in values that must be accumulated at positions on the zero-frequency axis, the implementation maintains Friedel symmetry by inserting contributions at both the target location $(k_r, 0)$ and its symmetric counterpart $(-k_r, 0)$.

2.4 Implementation Architecture

`torch-projectors` follows PyTorch's hybrid C++/Python extension architecture using the `TORCH_LIBRARY` registration system. The library implements a multi-backend design with separate optimized kernels for CPU, Apple Silicon (MPS), and CUDA devices.

Each backend maintains its own implementation in dedicated directories (`csrc/cpu/`, `csrc/mps/`, `csrc/cuda/`), with common utilities factored into `csrc/cpu/common/`. All projection operators are registered in a unified `torch_projectors` namespace using `TORCH_LIBRARY` declarations, enabling automatic device dispatch through PyTorch's operator registration system.

The Python interface wraps C++ operators with custom `torch.autograd.Function` classes that handle gradient computation. Library initialization loads the compiled C++ extension and registers Python autograd functions using `torch.library.register_autograd`. This architecture provides seamless integration with PyTorch's automatic differentiation while maintaining high performance through backend-specific optimizations.

2.5 2D $\rightarrow$ 2D Forward Projection

The 2D forward projection operator generates rotated 2D projections from 2D Fourier-space reconstructions. Given a reconstruction tensor $[B, N, N/2+1]$ and rotation matrices $[B_{rot}, P, 2, 2]$, the operator produces projections $[B, P, N_{out}, N_{out}/2+1]$ by sampling the reconstruction at transformed Fourier coordinates.

2.5.1 Forward Pass

For each output Fourier coordinate $\mathbf{k}_{out}$ in the projection, the operator computes the corresponding coordinate in the reconstruction as:

$$\mathbf{k}_{rec} = \mathbf{R}^{-1}\mathbf{k}_{out} \tag{3}$$

where $\mathbf{R}$ is the 2D rotation matrix. The projection value is obtained by interpolating the reconstruction at $\mathbf{k}_{rec}$, then applying phase modulation for translation:

$$P(\mathbf{k}_{out}) = F_{rec}(\mathbf{k}_{rec}) \cdot e^{-i2\pi\mathbf{k}_{out}\cdot\mathbf{s}} \tag{4}$$

where $\mathbf{s}$ is the shift vector.

2.5.2 Backward Pass

During backpropagation, the incoming projection gradients are first phase-modulated using the conjugate shift:

$$\nabla P' = \nabla P \cdot e^{i2\pi\mathbf{k}_{out}\cdot\mathbf{s}} \tag{5}$$

Reconstruction gradients are then accumulated via transposed interpolation operations at the transformed coordinates $\mathbf{k}_{rec}$.

Analytical rotation matrix gradients require careful application of the chain rule: Since $\mathbf{k}_{rec} = \mathbf{R}^{-1}\mathbf{k}_{out}$, the gradient with respect to rotation matrix element R_{ij} is:

$$\frac{\partial P}{\partial R_{ij}} = \frac{\partial F_{rec}(\mathbf{k}_{rec})}{\partial \mathbf{k}_{rec}} \cdot \frac{\partial \mathbf{k}_{rec}}{\partial R_{ij}} \tag{6}$$

For bilinear interpolation, spatial derivatives are computed from the 2×2 sample grid using separable linear kernel derivatives:

$$\frac{\partial F}{\partial r} = \sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij} \cdot l'(r_f - i) \cdot l(c_f - j) \tag{7}$$

$$\frac{\partial F}{\partial c} = \sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij} \cdot l(r_f - i) \cdot l'(c_f - j) \tag{8}$$

where $l(s) = 1 - |s|$ for $|s| \leq 1$ and $l'(s) = \text{sign}(s)$ for $|s| < 1$.

For bicubic interpolation, derivatives follow from separable Catmull-Rom kernel derivatives over the 4×4 neighborhood:

$$\frac{\partial F}{\partial r} = \sum_{i=-1}^{2}\sum_{j=-1}^{2} p_{ij} \cdot w'(r_f - i) \cdot w(c_f - j) \tag{9}$$

$$\frac{\partial F}{\partial c} = \sum_{i=-1}^{2}\sum_{j=-1}^{2} p_{ij} \cdot w(r_f - i) \cdot w'(c_f - j) \tag{10}$$

where $w'(s)$ is the cubic kernel derivative. The coordinate transformation derivatives $\frac{\partial \mathbf{k}_{rec}}{\partial R_{ij}}$ are computed from the matrix inverse relationships.

Shift gradients follow from the phase modulation derivative:

$$\frac{\partial P}{\partial \mathbf{s}} = -i2\pi\mathbf{k}_{out}P(\mathbf{k}_{out}) \tag{11}$$

2.6 2D $\rightarrow$ 2D Backward Projection

The 2D backward projection operator is the mathematical adjoint of the forward projection operation. It accumulates 2D Fourier-space projections $[B, P, N, N/2+1]$ into 2D reconstructions $[B, N_{rec}, N_{rec}/2+1]$, where the reconstruction size accounts for oversampling: $N_{rec} = N \cdot \text{oversampling}$.

2.6.1 Forward Pass

For each projection coordinate $\mathbf{k}_{proj} = (k_r, k_c)$, the corresponding coordinate in the 2D reconstruction is computed using equation 3 with the same 2×2 rotation matrix $\mathbf{R}$. The projection value is then accumulated into the reconstruction at the transformed coordinate. When shifts are present, the projection data is first conjugate phase-modulated:

$$P'(\mathbf{k}_{proj}) = P(\mathbf{k}_{proj}) \cdot e^{i2\pi \mathbf{k}_{proj} \cdot \mathbf{s}} \tag{12}$$

where the conjugate phase ensures the mathematical adjoint relationship with forward projection.

The accumulation process uses 2D interpolation kernels in transpose mode. For bilinear interpolation, contributions are distributed to the 2×2 neighborhood around each fractional coordinate. For bicubic interpolation, contributions are spread across the 4×4 neighborhood according to the Catmull-Rom weights defined in equations 7–8 and 9.

Optional weight accumulation supports applications requiring per-Fourier component weights, such as CTF correction in cryo-EM. Weights are accumulated using the absolute values of the 2D interpolation kernel weights, maintaining a reference for downstream normalization (e.g. for Wiener-like filters).

2.6.2 Backward Pass

During backpropagation, projection gradients are computed using forward projection of reconstruction gradients, exploiting the adjoint relationship. Rotation matrix gradients follow equation 6 with spatial derivatives computed using equations 7–8 for bilinear interpolation or equation 9 for bicubic interpolation. Shift gradients are computed using equation 11, applied to the accumulated reconstruction data.

2.7 3D $\rightarrow$ 2D Forward Projection

The 3D→2D forward projection operator implements the Central Slice Theorem, generating 2D projections from 3D Fourier-space reconstructions. Given a 3D reconstruction tensor $[B, D, H, W/2+1]$, 3×3 rotation matrices $[B, P, 3, 3]$, and optional 2D shifts $[B, P, 2]$, the operator produces projections $[B, P, H_{out}, W_{out}/2+1]$ by sampling central slices through the 3D volume.

2.7.1 Forward Pass

For each output projection coordinate $\mathbf{k}_{proj} = (k_r, k_c)$, the operator extends it to 3D by setting the third coordinate to zero: $\mathbf{k}_{3D} = (k_c, k_r, 0)$. This implements the central slice through the origin required by the Central Slice Theorem.

The corresponding coordinate in the 3D reconstruction is computed using the 3×3 rotation matrix:

$$\mathbf{k}_{rec} = \mathbf{R}^{-1}\mathbf{k}_{3D} \tag{13}$$

where $\mathbf{R}$ is the 3×3 rotation matrix. The projection value is obtained by trilinear or tricubic interpolation in the 3D volume at $\mathbf{k}_{rec}$, then applying phase modulation for translation:

$$P(\mathbf{k}_{proj}) = F_{rec}(\mathbf{k}_{rec}) \cdot e^{-i2\pi \mathbf{k}_{proj} \cdot \mathbf{s}} \tag{14}$$

where $\mathbf{s}$ is the 2D shift vector applied to the projection coordinates.

2.7.2 Backward Pass

During backpropagation, the incoming projection gradients are first phase-modulated using the conjugate shift applied to the 2D projection coordinates (using equation 5).

Reconstruction gradients are then accumulated via 3D transposed interpolation operations at the transformed coordinates $\mathbf{k}_{rec}$.

Analytical 3×3 rotation matrix gradients require extension of the chain rule to three dimensions: Since $\mathbf{k}_{rec} = \mathbf{R}^{-1}\mathbf{k}_{3D}$, the gradient with respect to rotation matrix element R_{ij} is:

$$\frac{\partial P}{\partial R_{ij}} = \frac{\partial F_{rec}(\mathbf{k}_{rec})}{\partial \mathbf{k}_{rec}} \cdot \frac{\partial \mathbf{k}_{rec}}{\partial R_{ij}} \tag{15}$$

For trilinear interpolation, spatial derivatives are computed from the 2×2×2 sample grid using separable linear kernel derivatives:

$$\frac{\partial F}{\partial d} = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} p_{ijk} \cdot l'(d_f - i) \cdot l(r_f - j) \cdot l(c_f - k) \tag{16}$$

$$\frac{\partial F}{\partial r} = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} p_{ijk} \cdot l(d_f - i) \cdot l'(r_f - j) \cdot l(c_f - k) \tag{17}$$

$$\frac{\partial F}{\partial c} = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} p_{ijk} \cdot l(d_f - i) \cdot l(r_f - j) \cdot l'(c_f - k) \tag{18}$$

where $l(s) = 1 - |s|$ for $|s| \leq 1$ and $l'(s) = \text{sign}(s)$ for $|s| < 1$.

For tricubic interpolation, derivatives follow from separable Catmull-Rom kernel derivatives over the 4×4×4 neighborhood:

$$\frac{\partial F}{\partial d} = \sum_{i=-1}^{2}\sum_{j=-1}^{2}\sum_{k=-1}^{2} p_{ijk} \cdot w'(d_f - i) \cdot w(r_f - j) \cdot w(c_f - k) \tag{19}$$

$$\frac{\partial F}{\partial r} = \sum_{i=-1}^{2}\sum_{j=-1}^{2}\sum_{k=-1}^{2} p_{ijk} \cdot w(d_f - i) \cdot w'(r_f - j) \cdot w(c_f - k) \tag{20}$$

$$\frac{\partial F}{\partial c} = \sum_{i=-1}^{2}\sum_{j=-1}^{2}\sum_{k=-1}^{2} p_{ijk} \cdot w(d_f - i) \cdot w(r_f - j) \cdot w'(c_f - k) \tag{21}$$

where $w'(s)$ is the cubic kernel derivative. The coordinate transformation derivatives $\frac{\partial \mathbf{k}_{rec}}{\partial R_{ij}}$ are computed from the matrix inverse relationships.

Shift gradients follow from the phase modulation derivative applied to the 2D projection coordinates:

$$\frac{\partial P}{\partial \mathbf{s}} = -i2\pi \mathbf{k}_{proj} P(\mathbf{k}_{proj}) \tag{22}$$

2.8 2D → 3D Backward Projection

The 2D→3D backward projection operator is the mathematical adjoint of the 3D→2D forward projection operation. It accumulates 2D Fourier-space projections $[B, P, H, W/2+1]$ into 3D reconstructions $[B, D, H_{rec}, W_{rec}/2+1]$, where the reconstruction dimensions account for oversampling and form a cubic volume: $D = H_{rec} = W_{rec} = H \cdot \text{oversampling}$.

2.8.1 Forward Pass

For each projection coordinate $\mathbf{k}_{proj} = (k_r, k_c)$, the operator extends it to 3D by setting the third coordinate to zero, implementing the central slice through the origin required by the Central Slice Theorem:

$$\mathbf{k}_{3D} = (k_c, k_r, 0) \tag{23}$$

The 3D sampling coordinate in the reconstruction is computed using equation 13 with the same 3×3 rotation matrix **R**. The projection value is then accumulated into the reconstruction at the transformed coordinate. When shifts are present, the projection data is first conjugate phase-modulated using equation 12 to ensure the mathematical adjoint relationship.

The accumulation process uses 3D interpolation kernels in transpose mode. For trilinear interpolation, contributions are distributed to the 2×2×2 neighborhood around each fractional coordinate. For tricubic interpolation, contributions are spread across the 4×4×4 neighborhood according to the Catmull-Rom weights defined in equations 16–18 and 19.

Optional weight accumulation supports applications requiring per-Fourier component weights, such as CTF correction in cryo-EM. Weights are accumulated using the absolute values of the 3D interpolation kernel weights, with proper handling of 3D Friedel symmetry for real-valued reconstructions.

2.8.2 Backward Pass

During backpropagation, projection gradients are computed using 3D→2D forward projection of reconstruction gradients, exploiting the adjoint relationship. The 3×3 rotation matrix gradients follow equation 15 with spatial derivatives computed using equations 16–18 for trilinear interpolation or equation 19 for tricubic interpolation. Shift gradients are computed using equation 22, applied to the 2D projection coordinates since shifts affect only the 2D projection plane.

3 Results

3.1 Performance Benchmarks

3.2 Accuracy Validation

4 Discussion

5 Availability

The `torch-projectors` library is available as an open-source Python package at `https://github.com/warpem/torch-projectors` and can be installed via pip. For CUDA-enabled packages, please refer to the repository's README.

References