Combining Scattering Transform paths cuts off the exponential growth while retaining key features

Collating Transform

David Weber, Naoki Saito
UC Davis

Scattering Transform

A Scattering layer consists of a collection of frames $\Psi^\ell = \{\phi\} \bigcup \{\psi_{\lambda^\ell}\}_{\lambda^\ell \in \Lambda^\ell}$, a nonlinearity σ such as ReLU, and a subsampler R_r with rate r:

$$U[\lambda^\ell](f) = R_r\sigma(\psi_{\lambda^\ell} \star f)$$

A *scattering path* $p = (\lambda^1, \dots, \lambda^L)$ consists of repeatedly applying these operations and finally averaging:

$$U^\ell[p](f) = U[\lambda^\ell] \cdots U[\lambda^1](f)$$
$$S[p](f) = \phi \star U[p](f)$$

Collating Transform

Given a set of signals $f_1, \dots, f_c$, a collating layer interposes a sum across frames using a collating matrix $A \in \mathbb{R}^{K \times c}$:

$$C[\lambda^\ell, k](f_1, \dots, f_c) = R_r\sigma\left(\sum_{i=1}^{c} A_{ki}\psi_{\lambda^\ell} \star f_i\right)$$
$$S[\lambda^\ell, k](f) = \phi \star C[\lambda^\ell, k](f)$$

Theoretical properties of both

The following carry over simply from the scattering transform to the collating transform:

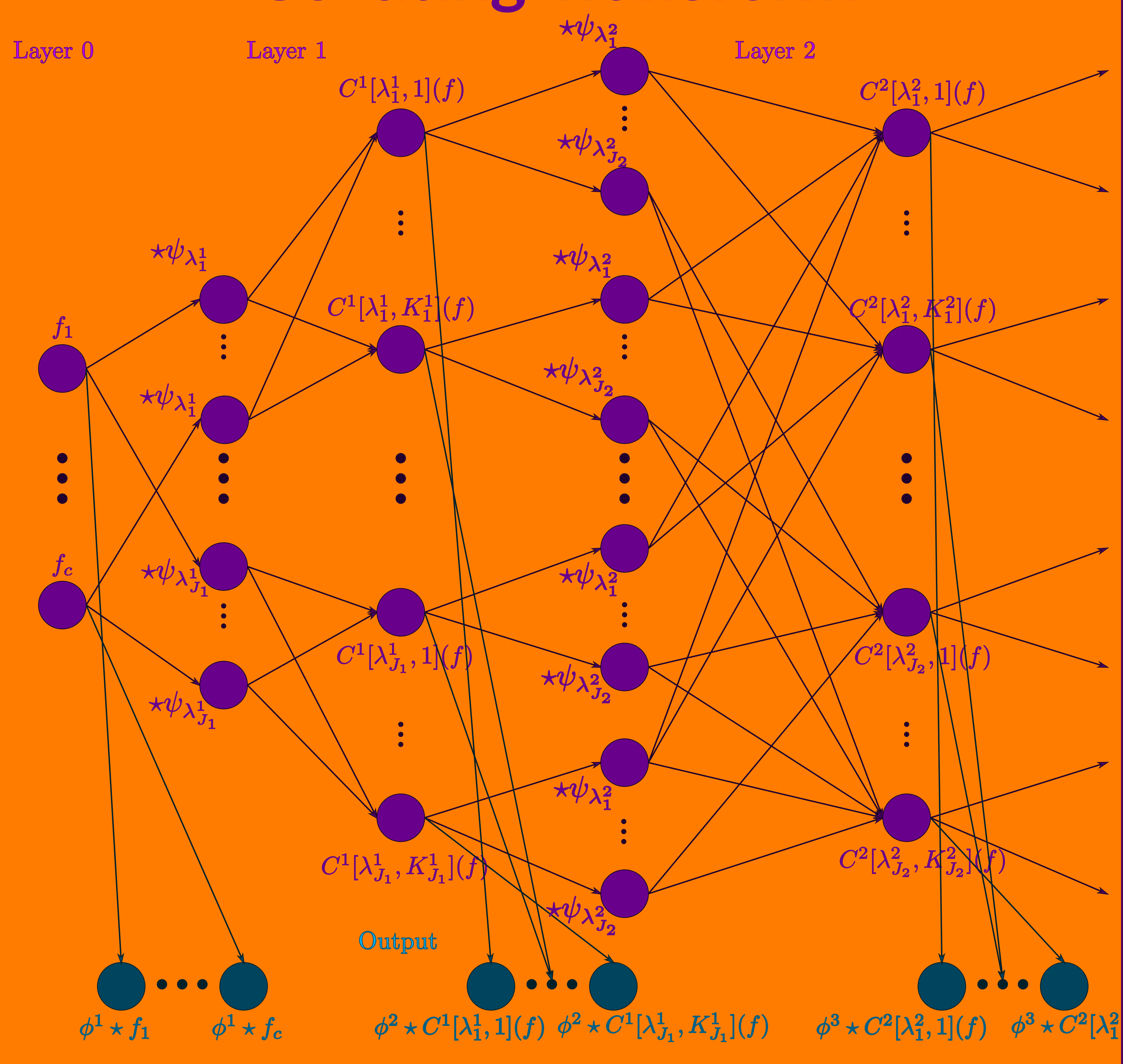
Theorem 1. *Translation decays exponentially with depth:*

$$S[\lambda^\ell, k](T_c f) = T_{c/r_1 \dots r_\ell} S[\lambda^\ell, k](f)$$

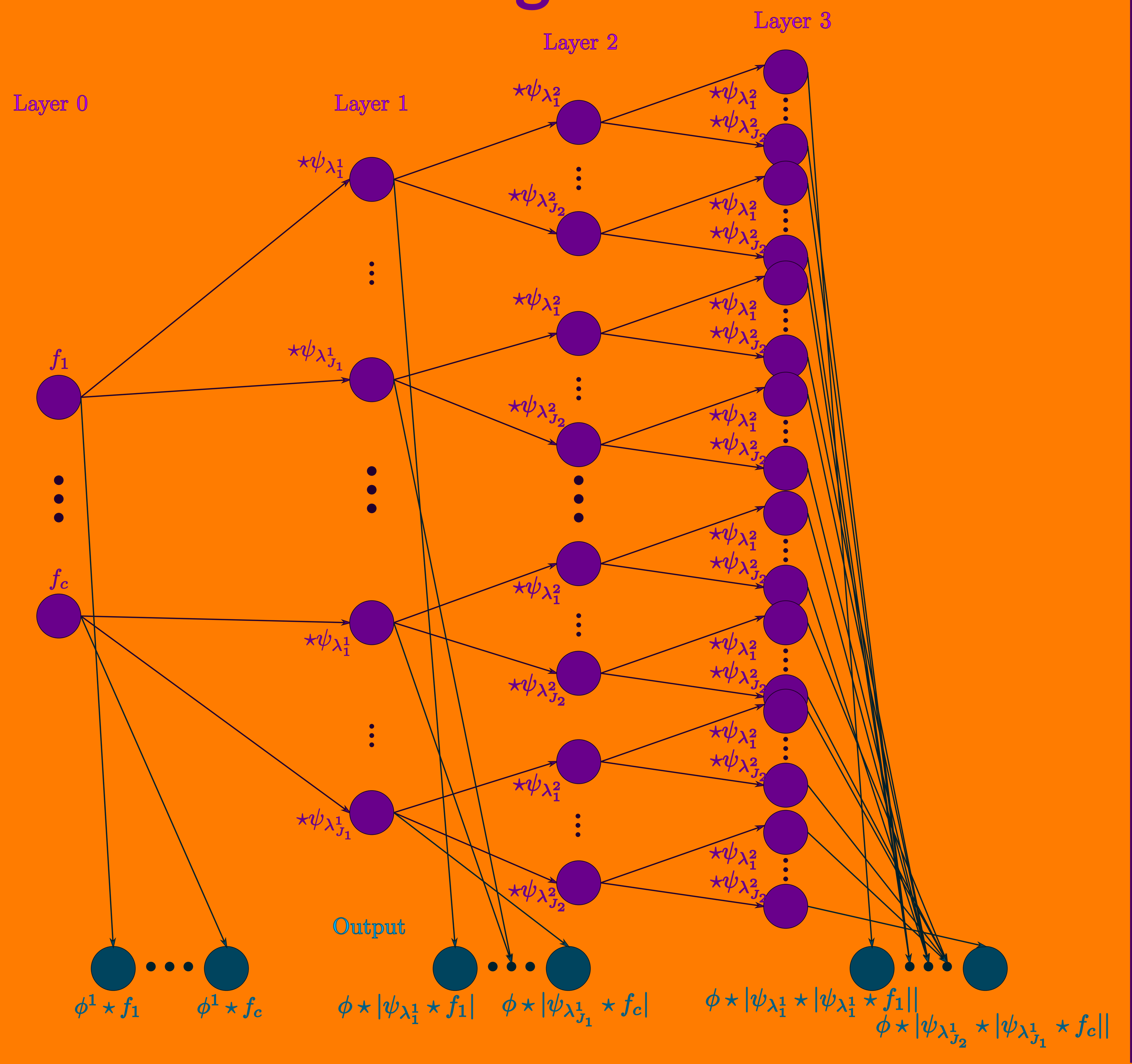
Theorem 2. *Given upper frame bound B_ℓ, lipschitz bound γ_ℓ on σ, and c signals of dimension d, if we have $\max\{B_\ell, B_\ell r_{\ell+1}^{-d}\gamma_{\ell+1}(\max_{ji}|A_{ji}^\ell|^2)\} \leq 1$, then there is a universal C so for R band-limited functions $f \in \mathscr{L}^2(\mathbb{R}^d, \mathbb{R}^c)$, $\omega \in \mathscr{C}(\mathbb{R}^d, \mathbb{R}^c)$, $\tau \in \mathscr{C}^1(\mathbb{R}^d, \mathbb{R}^d)$,*

$$\left(F_{\tau,\omega} f\right)_i := e^{i\omega_i(x)} f_i(x + \tau(x))$$
$$\left\| S\left[F_{\tau,\omega} f\right] - S[f] \right\| \leq C\left(R\|\tau\|_\infty + \max_j \|\omega_j\|_\infty\right) \|f\|_2$$

Collating Transform



Scattering Transform



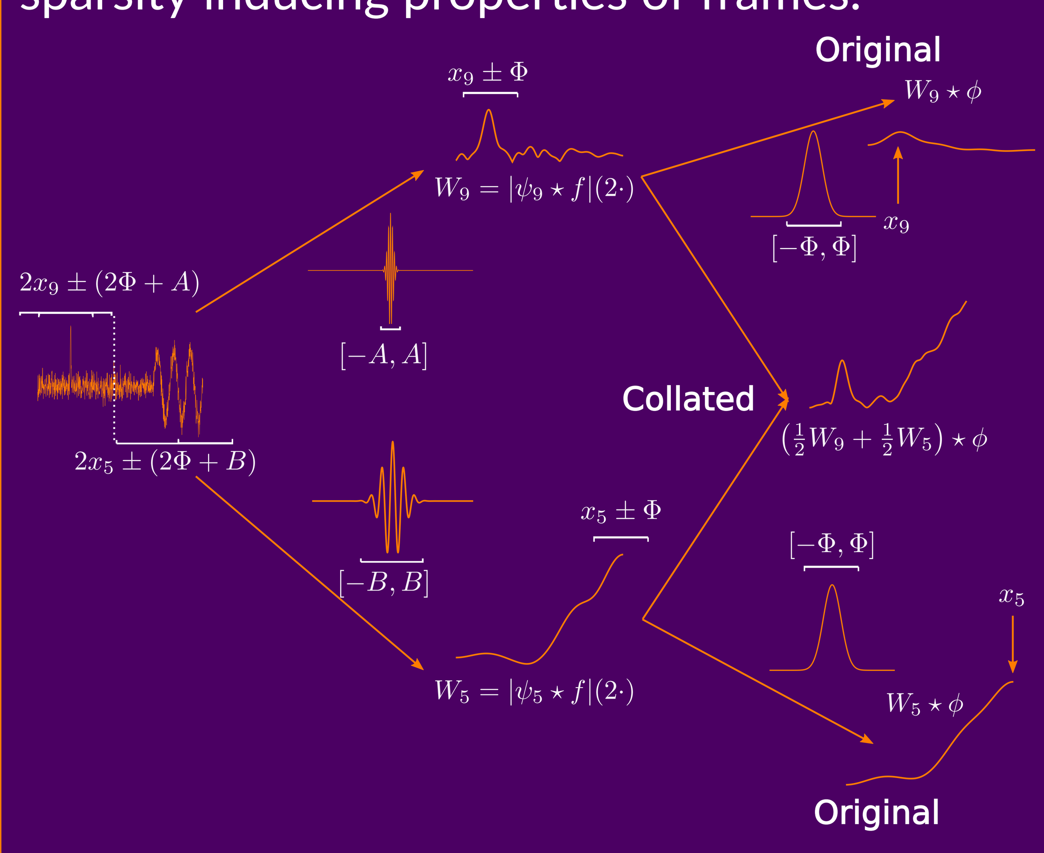
Comparison

Learned CNNs match the structure of the collating transform, but learn the entire frame collator $A_{ki}\psi_{\lambda^\ell}$ product tensor, rather than just the collation.

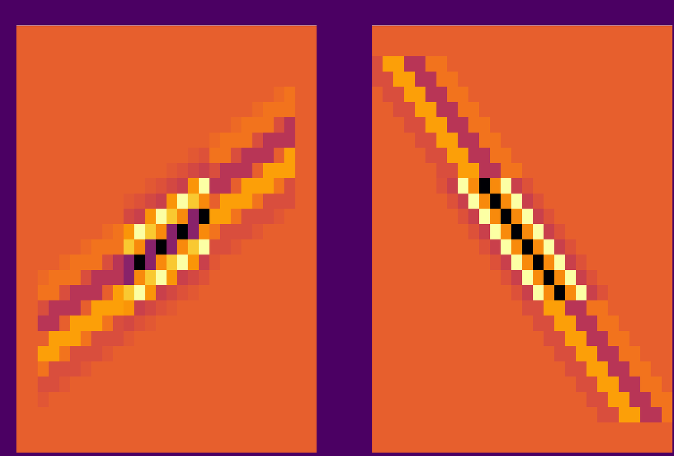
Given F frames in each layer, c input channels, and n^d coefficients in each channel, there are cF^L scattering paths at layer L, and $\sum_{\ell=1}^{L} cn^d\left(\frac{F}{r^d}\right)^\ell$ output coefficients. If there are K_ℓ collators at layer ℓ, there are FK_ℓ nodes, and a total of $\sum_{\ell=0}^{L} \frac{n^d}{r^{d\ell}} FK_\ell$ coefficients.

Combining Paths

If path a contributes values at x_a, and path b at x_b, then if $\{x_a \pm (2\Phi + A)\}$ and $\{x_B \pm (2\Phi + B)\}$ don't overlap, adding them keeps the same value at those locations, using the sparsity inducing properties of frames.



For 2D transforms, there is also a spatial element to overlap



Extends to groups of paths that use the same frame at a given layer, decomposition can happen at different levels.