Equivariant Graph Neural Networks Improve Optical Spectra Prediction for Materials Screening

Kasper Helverskov Petersen [1] **François Raymond J Cornet** [1] **Martin Ovesen** [2] **Mikkel Jordahn** [1] **Kristian S. Thygesen** [2] **Mikkel N. Schmidt** [1]

Abstract

Scalable prediction of optical spectra is a critical component of high-throughput materials screening for optoelectronic applications such as solar cells. Existing surrogate models are trained on spectra computed from lower levels of theory or rely on rotation-invariant scalar features, limiting their geometric expressiveness. We explore the use of equivariant graph neural networks for optical spectra prediction, adapting GotenNet to this task and evaluating it on multiple datasets including a recently published collection of 10,533 structures with spectra computed at the level of the random phase approximation (RPA). The proposed model outperforms the current state of the art, with the largest gains in the 0–8 eV range and on predicting the static real permittivity, both of particular relevance for thin-film optics.

1. Introduction

AI is increasingly applied across scientific disciplines, and materials discovery, sitting at the confluence of physics, chemistry, and engineering, is a natural target. The combinatorial vastness of chemical and structural space makes AI-driven exploration particularly compelling: generative models propose candidate materials; surrogate models screen them for target properties; and first-principles calculations and experimental characterization evaluate the most promising candidates. Together, these form a pipeline bridging machine learning, computational physics, and materials synthesis and characterization. For discovering improved materials for solar cells, a critical component is the scalable prediction of optical spectra. Experimental measurement and ab initio methods such as density functional theory (DFT) (Hohenberg & Kohn, 1964; Kohn & Sham, 1965) and many-body perturbation theory (MBPT) (Onida et al., 2002) are resource-intensive and computationally prohibitive for large-scale screening. Machine learning surrogates trained on first-principles databases offer a compelling alternative, but must be accurate enough to reliably distinguish promising candidates from poor ones, making architectural improvements a key priority.

Existing approaches (Hung et al., 2024; Ibrahim & Ataca, 2024; Ginter et al., 2025) are trained on limited data at lower levels of theory, such as the independent-particle approximation (IPA) or tight-binding (TB). Meanwhile, Grunert et al. (2026c) recently introduced a dataset of 10,533 optical spectra at the level of the random phase approximation (RPA) using an IPA-pretrained model and achieving state-of-the-art results. Their architecture, OptiMate3B, encodes geometry only through rotation-invariant scalar features. Furthermore, it explicitly constructs a line-graph over edges to capture angular features, which necessitates a Voronoi graph construction to limit connectivity. Equivariant architectures avoid both limitations: they encode directional information by construction, providing a geometric inductive bias that is especially valuable given limited training data, and capture angular correlations implicitly, allowing distance-based graph construction that more naturally reflects the physical interaction ranges of atoms.

In this work, we adapt GotenNet (Aykent & Xia, 2025), a state-of-the-art equivariant graph neural network (GNN), to optical spectra prediction, and evaluate it on both the IPA and RPA OptiMate3B datasets. The resulting model, $\text{GotenNet}_{\text{Opt}}$, achieves consistent improvements with the most pronounced gains in the 0–8 eV range and on predicting the static real permittivity, both of particular relevance for materials screening in thin-film optics. We further benchmark against GNNOpt (Hung et al., 2024) on their smaller dataset of 944 IPA-computed spectra, demonstrating improvements in a low-data setting.

[1]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark [2]Department of Physics, Technical University of Denmark, Kongens Lyngby, Denmark. Correspondence to: Kasper Helverskov Petersen <khepe@dtu.dk>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

2. Background

2.1. The dielectric function

The predicted optical spectra consist of the complex dielectric function, which describes the response of a solid to an applied electric field $\mathbf{E}$ of angular frequency ω. The dielectric function is actually a second-rank tensor, $\varepsilon_{ij}(\omega)$, relating the electric polarization density along the Cartesian direction i induced by an electric field polarized along the direction j. The tensorial nature reflects that the polarization response can differ along different crystallographic directions. Mathematically, this can be expressed as the linear relationship between the components of the electric displacement field $\mathbf{D}$ and the components of the applied electric field $\mathbf{E}$

$$D_i(\omega) = \varepsilon_0 \sum_j \varepsilon_{ij}(\omega) E_j(\omega), \tag{1}$$

where ε_0 is the permittivity of free space and the second-rank tensor is unitless. Alternatively, the second-rank tensor can be formulated as the electric susceptibility, χ_{ij}, describing the linear relationship between the polarization density $\mathbf{P}$ and the applied electric field

$$P_i(\omega) = \varepsilon_0 \sum_j \chi_{ij}(\omega) E_j(\omega), \tag{2}$$

where the relationship between ε_{ij} and χ_{ij} is simply

$$\varepsilon_{ij}(\omega) = \delta_{ij} + \chi_{ij}(\omega). \tag{3}$$

Typically, determining all 9 tensor elements is unnecessary due to the crystal structure being invariant w.r.t. certain sets of symmetry operations. For example, for materials with cubic symmetry or isotropic structure, the off-diagonal elements of the tensor vanish while the diagonal elements reduce to a scalar, i.e. $\varepsilon_{ij}(\omega) = \varepsilon(\omega)\delta_{ij}$. More generally, the isotropic part of the dielectric function tensor can be defined as

$$\bar{\varepsilon}(\omega) = \mathrm{Tr}(\varepsilon(\omega))/3, \tag{4}$$

which captures the mean linear response across all crystallographic directions. It consists of a real and imaginary part which are related to the refractive index $n(\omega)$ and the extinction coefficient $\kappa(\omega)$, describing the amount of light attenuation, through

$$n(\omega) + i\kappa(\omega) = \sqrt{\bar{\varepsilon}(\omega)}, \tag{5}$$

which can also be written as

$$n = \sqrt{\frac{\sqrt{\mathrm{Re}(\bar{\varepsilon})^2 + \mathrm{Im}(\bar{\varepsilon})^2} + \mathrm{Re}(\bar{\varepsilon})}{2}}, \tag{6}$$

and

$$\kappa = \sqrt{\frac{\sqrt{\mathrm{Re}(\bar{\varepsilon})^2 + \mathrm{Im}(\bar{\varepsilon})^2} - \mathrm{Re}(\bar{\varepsilon})}{2}}. \tag{7}$$

While the real and imaginary parts of $\bar{\varepsilon}$ (or equivalently, $\bar{\chi}$) can be calculated independently of each other, a consequence of causality is that they must satisfy the criterion that they are related to each other by Hilbert transformations (Toll, 1956)

$$\mathrm{Re}\left(\bar{\chi}(\omega)\right) = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \frac{\mathrm{Im}\left(\bar{\chi}\left(\omega'\right)\right)}{\omega' - \omega} \mathrm{d}\omega', \tag{8}$$

and

$$\mathrm{Im}\left(\bar{\chi}(\omega)\right) = -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \frac{\mathrm{Re}\left(\bar{\chi}\left(\omega'\right)\right)}{\omega' - \omega} \mathrm{d}\omega', \tag{9}$$

where $\mathcal{P}$ denotes that the singularity at $\omega' = \omega$ is resolved as the Cauchy principal value. The equations (8) and (9) are known as the Kramers-Kronig relations.

For the OptiMate3B dataset, we predict the concatenated real and imaginary parts of $\bar{\varepsilon}(\omega)$, sampled from 0 to 20 eV in 10 meV steps, yielding

$$\mathrm{out} = \mathrm{Re}[\mathrm{Tr}(\varepsilon(\omega))/3] \,\|\, \mathrm{Im}[\mathrm{Tr}(\varepsilon(\omega))/3], \tag{10}$$

which is a 4002-dimensional output vector. Whereas, for the GNNOpt dataset, a separate model is trained and evaluated for the real and imaginary components, with each predicting a single component interpolated onto a uniform grid of 251 points over $0 \leq \hbar\omega \leq 50$ eV, where $\hbar$ is the reduced Planck constant:

$$\mathrm{out}_{\mathrm{Re}} = \mathrm{Re}[\mathrm{Tr}(\varepsilon(\omega))/3], \qquad \mathrm{out}_{\mathrm{Im}} = \mathrm{Im}[\mathrm{Tr}(\varepsilon(\omega))/3], \tag{11}$$

each yielding a 251-dimensional output vector. A brief background section on DFT is provided in Section A.

2.2. 3D graph representation of crystal structures

A crystal is an infinite periodic solid described by a unit cell: a parallelepiped spanned by lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$ and a finite set of atoms at positions $\{\mathbf{x}_i\}$, given in Cartesian or fractional coordinates $\{\mathbf{s}_i\} \subset [0,1)^3$ via $\mathbf{x}_i = \sum_k s_{i,k}\mathbf{a}_k$. Under periodic boundary conditions (PBC), each atom at $\mathbf{x}_i$ has images at $\mathbf{x}_i + \sum_k n_k \mathbf{a}_k$ for all $n_k \in \mathbb{Z}$.

We represent crystals as a graph $G = (V, E)$, where nodes V are the unit cell atoms and edges E encode atomic interactions, with the neighborhood of atom i defined as $\mathcal{N}(i) = \{j \in V \mid (i,j) \in E\}$. PBC means the neighborhood search includes periodic images, so multiple edges may exist between the same pair of atoms, each with a distinct edge vector $\mathbf{r}_{ij}$. We use a distance-based cutoff of $r_c = 6.0$ Å, and additionally evaluate a Voronoi-based construction, which yields fewer edges and is the method used by OptiMate3B.

Each atom i is featurized by concatenating one-hot encodings of its periodic table group (18 classes) and period (7

classes), giving $\mathbf{z}_i \in \mathbb{R}^{25}$. Each edge carries the unit vector $\hat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/\|\mathbf{r}_{ij}\|$ and the interatomic distance $\|\mathbf{r}_{ij}\|$.

2.3. Equivariant graph neural networks

GNN architectures for atomic systems are commonly formalized within the message passing neural network (MPNN) framework (Gilmer et al., 2017), where atom embeddings are iteratively updated by aggregating information from neighbors. At each round $t \in \{1, \ldots, T\}$, messages are computed and aggregated, the embeddings updated, and after T rounds a readout layer produces the final output:

$$\mathbf{m}_i^{(t+1)} = \bigoplus_{j \in \mathcal{N}(i)} M_t\left(\mathbf{h}_i^{(t)}, \mathbf{h}_j^{(t)}\right), \tag{12}$$

$$\mathbf{h}_i^{(t+1)} = U_t\left(\mathbf{h}_i^{(t)}, \mathbf{m}_i^{(t+1)}\right), \tag{13}$$

$$\hat{y} = R\left(\left\{\mathbf{h}_i^{(T)} \mid i \in V\right\}\right), \tag{14}$$

where M_t, U_t, and R are the message, update, and readout functions, and $\bigoplus$ is a permutation-invariant aggregation.

Early architectures relied on invariant scalar features derived from pairwise distances, angles, and torsions (Gilmer et al., 2017; Schütt et al., 2017; Lu et al., 2019; Klicpera et al., 2020; Liu et al., 2021). Equivariant architectures instead propagate vector or higher-order tensor features that transform consistently under symmetry operations (Brandstetter et al., 2021; Satorras et al., 2021; Schütt et al., 2021; Le et al., 2022; Thölke & Fabritiis, 2022; Batatia et al., 2023; Simeon & de Fabritiis, 2023; Liao & Smidt, 2023; Aykent & Xia, 2025). Formally, a function f is equivariant to the action of a group $\mathcal{G}$ if

$$f(g \cdot \mathbf{x}) = g \cdot f(\mathbf{x}) \quad \forall\, g \in \mathcal{G}. \tag{15}$$

By design, this property guarantees orientation-consistent predictions, eliminating the need for the model to learn these geometric symmetries directly from the training data.

3. Method

3.1. GotenNet$_{\text{Opt}}$ Architecture

Figure 1 outlines the proposed architecture, which builds upon GotenNet (Aykent & Xia, 2025). For clarity, we reformulate the model below and highlight our specific modifications.

3.1.1. Tensor Features

For each edge (i, j), we encode tensor edge features: $\tilde{\mathbf{r}}_{ij} = \{\tilde{\mathbf{r}}_{ij}^{(l)}\}_{l=0}^{L_{\max}}$, where $\tilde{\mathbf{r}}_{ij}^{(0)} = \|\mathbf{r}_{ij}\|$ is the scalar distance, $\tilde{\mathbf{r}}_{ij}^{(1)} = \hat{\mathbf{r}}_{ij}$ is the normalized direction vector, and $\tilde{\mathbf{r}}_{ij}^{(l)} = Y^{(l)}(\hat{\mathbf{r}}_{ij}) \in \mathbb{R}^{2l+1}$ for $l > 1$ are degree-l real spherical harmonics in Cartesian coordinates, normalized to unit norm. Each node i similarly carries steerable features $\{\tilde{\mathbf{X}}_i^{(l)}\}_{l=1}^{L_{\max}}$, with $\tilde{\mathbf{X}}_i^{(l)} \in \mathbb{R}^{(2l+1)\times F}$ transforming as a rank-l irreducible representation of SO(3). Both the edge and node tensor features thus capture increasingly higher-order geometric information: degree $l = 1$ encodes directional information, while higher degrees encode more complex angular patterns such as planar orientations ($l = 2$). The edge tensor features are fixed geometric inputs, while the steerable node features are initialized to zero and iteratively updated across message passing layers.

3.1.2. Embedding Layer

Scalar node feature initialization. Scalar node features $\mathbf{h}_i \in \mathbb{R}^F$ are initialized by aggregating over neighbors $j \in \mathcal{N}(i)$ using element-wise products of atom, covalent radius, and distance embeddings. Inspired by GNNOpt (Hung et al., 2024), which demonstrated that incorporating covalent radius as an explicit physical descriptor improves prediction of optical spectra, we augment the original GotenNet embedding with covalent radius gating. Each atom's covalent radius $c_i \in \mathbb{R}$ is expanded into a fixed-width representation via a Gaussian RBF basis $\boldsymbol{\psi}(c_i)$, providing a continuous encoding of atomic size.

Neighbor messages are then formed by gating the projected species embedding with the neighbor's covalent radius embedding:

$$\mathbf{m}_i = \sum_{j \in \mathcal{N}(i)} \left(\sigma(\mathbf{W}_{\text{nbr}}\mathbf{z}_j) \circ \mathbf{W}_{\text{cbr}}\boldsymbol{\psi}(c_j)\right) \circ \mathbf{d}_{ij}, \tag{16}$$

where

$$\mathbf{d}_{ij} = f_{\text{cut}}(\|\mathbf{r}_{ij}\|) \cdot \mathbf{W}_{\text{ndp}}\, \boldsymbol{\varphi}(\|\mathbf{r}_{ij}\|) \tag{17}$$

projects a exponential RBF expansion $\boldsymbol{\varphi}(\|\mathbf{r}_{ij}\|)$, cosine-cutoff weighted to suppress distant neighbors. The source atom embedding is similarly gated by its own covalent radius:

$$\mathbf{a}_i = \sigma(\mathbf{W}_{\text{na}}\mathbf{z}_i) \circ \mathbf{W}_{\text{cna}}\boldsymbol{\psi}(c_i). \tag{18}$$

This multiplicative gating allows atomic size to modulate the contribution of species identity, encouraging the model to distinguish atoms that share a period or group but differ in effective radius. Node features are then initialized by a two-layer MLP over the concatenation of $\mathbf{a}_i$ and $\mathbf{m}_i$:

$$\mathbf{h}_{i,\text{init}} = \mathbf{W}_{\text{nru}}\, \sigma(\text{LN}(\mathbf{W}_{\text{nrd}}[\mathbf{a}_i, \mathbf{m}_i])) + \mathbf{b}_{\text{nru}}, \tag{19}$$

where σ denotes SiLU and LN denotes layer normalization.

Scalar edge feature initialization. The scalar edge features $\mathbf{t}_{ij} \in \mathbb{R}^F$ are initialized as the element-wise product of the summed initial node features and a linear projection of the RBF-expanded edge distances:

$$\mathbf{t}_{ij,\text{init}} = (\mathbf{h}_{i,\text{init}} + \mathbf{h}_{j,\text{init}}) \circ (\mathbf{W}_{\text{erp}}\varphi(||\mathbf{r}_{ij}||) + \mathbf{b}_{\text{erp}}) \tag{20}$$

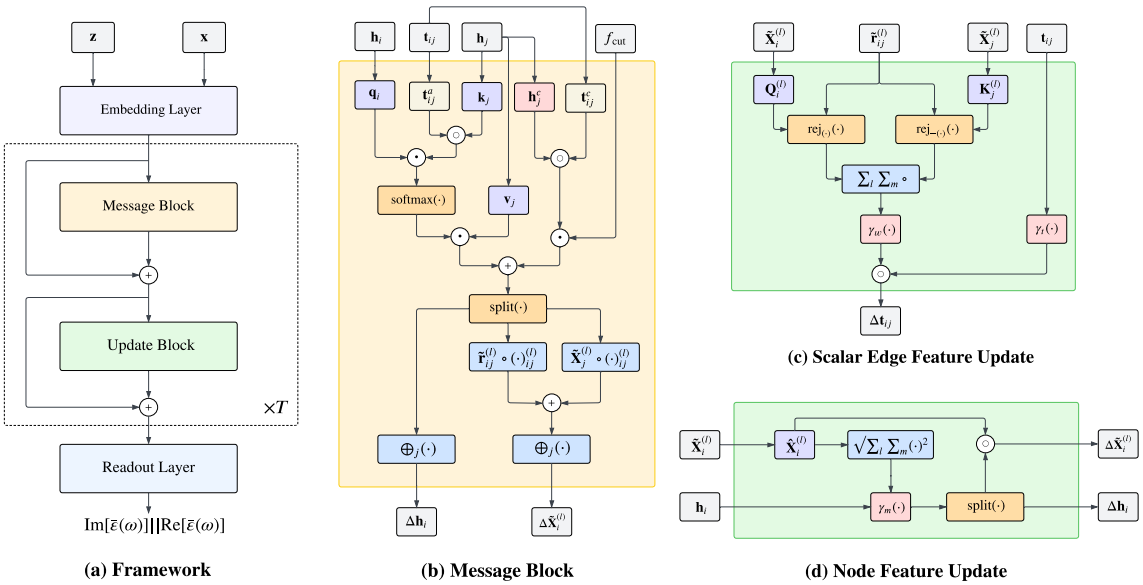

Figure 1. Architecture of GotenNet$_{\text{Opt}}$. (a) The model follows the MPNN framework: an embedding layer initializes atomic and bond features from elemental information and 3D coordinates, T message-passing rounds update the representations, and a readout layer predicts the optical spectra. (b) The message block. (c) Scalar edge feature update. (d) Node feature update. The summation $\sum_l \sum_m$ is over $l = 1, \ldots, L_{\text{max}}$ and $m = 2l + 1$. Symbols: $\cdot$ dot product, $\circ$ element-wise product. Figure adapted from (Aykent & Xia, 2025).

3.1.3. Message Block

The message block updates node features via the Geometry-Aware Tensor Attention (GATA) module from GotenNet, adopted without modification except for removing the layer normalization on scalar node features.

First, queries, keys, and a projected edge embedding are computed:

$$\mathbf{q}_i = \mathbf{W}_q \mathbf{h}_i + \mathbf{b}_q, \tag{21}$$

$$\mathbf{k}_j = \mathbf{W}_k \mathbf{h}_j + \mathbf{b}_k, \tag{22}$$

$$\mathbf{t}^a_{ij} = \sigma(\mathbf{W}_{\text{re}}\, \mathbf{t}_{ij} + \mathbf{b}_{\text{re}})\,. \tag{23}$$

Splitting into H heads of dimension $F_{\text{head}} = F/H$, the attention score for head h is computed and normalized as:

$$\alpha^{(h)}_{ij} = \big(\mathbf{q}^{(h)}_i\big)^\top \big(\mathbf{k}^{(h)}_j \circ \mathbf{t}^{a,(h)}_{ij}\big), \tag{24}$$

$$\alpha^{(h)}_{ij} \leftarrow \frac{\exp\big(\alpha^{(h)}_{ij}\big)}{\sum_{k\in\mathcal{N}(i)} \exp\big(\alpha^{(h)}_{ik}\big)} \cdot \frac{\sqrt{|\mathcal{N}(i)|}}{\sqrt{F}}, \tag{25}$$

where $|\mathcal{N}(i)|$ denotes the degree of node i. The value vector $\mathbf{v}_j = \gamma_v(\mathbf{h}_j) \in \mathbb{R}^{S\cdot F}$ is obtained via a two-layer MLP $\gamma_v \colon \mathbb{R}^F \to \mathbb{R}^{S\cdot F}$, with $S = 2L_{\text{max}} + 1$. Reshaping into H heads of dimension $F_{\text{ch}} = S \cdot F/H$, the scaled edge attention (SEA) for head h is:

$$\mathbf{sea}^{(h)}_{ij} = \alpha^{(h)}_{ij}\, \mathbf{v}^{(h)}_j \in \mathbb{R}^{F_{\text{ch}}}. \tag{26}$$

To compute the per-head coefficient vectors, we additionally introduce a second edge projection $\mathbf{t}^c_{ij} = \mathbf{W}_{\text{rs}}\, \mathbf{t}_{ij} + \mathbf{b}_{\text{rs}} \in \mathbb{R}^{S\cdot F}$ and an MLP-transformed neighbor node feature $\mathbf{h}^c_j = \gamma_s(\mathbf{h}_j) \in \mathbb{R}^{S\cdot F}$, both reshaped into H heads so that $\mathbf{t}^{c,(h)}_{ij}, \mathbf{h}^{c,(h)}_j \in \mathbb{R}^{F_{\text{ch}}}$:

$$\mathbf{c}^{(h)}_{ij} = \mathbf{sea}^{(h)}_{ij} + f_{\text{cut}}(\|\mathbf{r}_{ij}\|) \cdot \big(\mathbf{t}^{c,(h)}_{ij} \circ \mathbf{h}^{c,(h)}_j\big) \in \mathbb{R}^{F_{\text{ch}}}. \tag{27}$$

Concatenating across all heads and reshaping gives $\mathbf{c}_{ij} \in \mathbb{R}^{S\times F}$, which we decompose into S components of dimension F:

$$\mathbf{c}_{ij} = \Big[\mathbf{c}^s_{ij},\ \big\{\mathbf{c}^{d,(l)}_{ij}\big\}_{l=1}^{L_{\text{max}}},\ \big\{\mathbf{c}^{t,(l)}_{ij}\big\}_{l=1}^{L_{\text{max}}}\Big]. \tag{28}$$

These coefficients define the messages to the scalar and steerable node features:

$$\Delta\mathbf{h}_i = \bigoplus_{j\in\mathcal{N}(i)} \mathbf{c}^s_{ij}, \tag{29}$$

$$\Delta\widetilde{\mathbf{X}}^{(l)}_i = \bigoplus_{j\in\mathcal{N}(i)} \Big(\mathbf{c}^{d,(l)}_{ij} \circ \widetilde{\mathbf{r}}^{(l)}_{ij} + \mathbf{c}^{t,(l)}_{ij} \circ \widetilde{\mathbf{X}}^{(l)}_j\Big), \tag{30}$$

which are added to the respective features:

$$\mathbf{h}_i \leftarrow \mathbf{h}_i + \Delta\mathbf{h}_i, \qquad \widetilde{\mathbf{X}}^{(l)}_i \leftarrow \widetilde{\mathbf{X}}^{(l)}_i + \Delta\widetilde{\mathbf{X}}^{(l)}_i. \tag{31}$$

Note that the high-degree steerable features are initialized as $\mathbf{X}^{(l)} \equiv \mathbf{0}$, which makes the $\mathbf{c}^{t,(l)}_{ij} \circ \tilde{\mathbf{X}}^{(l)}_j$ term equal zero in the first message block, resulting in the initial features:

$$\tilde{\mathbf{X}}^{(l)}_{i,\text{init}} = \bigoplus_{j\in\mathcal{N}(j)} (\mathbf{c}^{d,(l)}_{ij} \circ \tilde{\mathbf{r}}^{(l)}_{ij}) \tag{32}$$

3.1.4. Update Block

The update block updates the scalar edge features through the Hierarchical Tensor Refinement (HTR) module and the scalar and steerable node features through the Equivariant Feed-Forward (EQFF) network proposed in GotenNet.

Scalar edge feature update. For each degree $l = 1, \ldots, L_{\text{max}}$, we compute transformed steerable features:

$$\mathbf{Q}_i^{(l)} = \tilde{\mathbf{X}}_i^{(l)} \mathbf{W}_{vq}, \quad \mathbf{K}_j^{(l)} = \tilde{\mathbf{X}}_j^{(l)} \mathbf{W}_{vk}^{(l)}, \tag{33}$$

where $\mathbf{W}_{vq}$ is shared across degrees and $\mathbf{W}_{vk}^{(l)}$ is degree-specific. We then obtain $\hat{\mathbf{Q}}_i^{(l)}$ and $\hat{\mathbf{K}}_j^{(l)}$ by applying vector rejection of $\mathbf{Q}_i^{(l)}$ onto $\tilde{\mathbf{r}}_{ij}^{(l)}$ and of $\mathbf{K}_j^{(l)}$ onto $-\tilde{\mathbf{r}}_{ij}^{(l)}$, respectively. This ensures the outputs retain only what the network has learned beyond the edge geometry.

Next an interaction vector $\mathbf{w}_{ij}$ is obtained by summing element-wise products across all degrees and components:

$$\mathbf{w}_{ij} = \sum_{l=1}^{L_{\text{max}}} \sum_{m=1}^{2l+1} \hat{\mathbf{Q}}_{i,m}^{(l)} \circ \hat{\mathbf{K}}_{j,m}^{(l)}, \tag{34}$$

capturing learned atomic interactions not explained by edge geometry alone. Finally, the residual update to the scalar edge features is:

$$\Delta \mathbf{t}_{ij} = \gamma_w(\mathbf{w}_{ij}) \circ \gamma_t(\mathbf{t}_{ij}), \tag{35}$$

where γ_w and γ_t are two-layer MLPs, giving $\mathbf{t}_{ij} \leftarrow \mathbf{t}_{ij} + \Delta \mathbf{t}_{ij}$.

Node feature updates. For each $l = 1, \ldots, L_{\text{max}}$ the steerable feature go through a linear projection $\hat{\mathbf{X}}_i^{(l)} = \tilde{\mathbf{X}}_i^{(l)} \mathbf{W}_{\text{vu}}$. Second, we compute a channel-wise norm over all projected features:

$$\mathbf{n}_i = \sqrt{\sum_{l=1}^{L_{\text{max}}} \sum_{m=1}^{2l+1} (\hat{\mathbf{X}}_{i,m}^{(l)})^2 + \epsilon} \tag{36}$$

Next, the scalar node features and this norm are concatenated and passed through a two-layer MLP, which is split into two parts:

$$[\mathbf{m}_{i,1}, \mathbf{m}_{i,2}] = \gamma_m((\mathbf{n}_i, \mathbf{h}_i)) \tag{37}$$

After which, the scalar and high-degree steerable node features are updated as:

$$\mathbf{h}_i \leftarrow \mathbf{m}_{1,i} + \mathbf{h}_i \tag{38}$$

$$\tilde{\mathbf{X}}_i^{(l)} \leftarrow (\mathbf{m}_{i,2} \circ \hat{\mathbf{X}}_i^{(l)}) + \tilde{\mathbf{X}}_i^{(l)} \tag{39}$$

3.1.5. Readout Layer

For the readout layer, we use the same design as OptiMate3B. It consists of an attention-weighted global pooling to produce a graph-level embedding, followed by an output MLP to predict the spectrum. First, a linear layer with SiLU activation computes a per-node attention vector $\mathbf{a}_i = \sigma(\mathbf{W}_{\text{pool}} \mathbf{h}_i + \mathbf{b}_{\text{pool}}) \in \mathbb{R}^F$ from the node features, which is then normalised across all nodes in the graph:

$$\boldsymbol{\alpha}_i = \frac{\exp(\mathbf{a}_i)}{\sum_{j \in V} \exp(\mathbf{a}_j)} \in \mathbb{R}^F, \tag{40}$$

where exp and division are applied elementwise, and the normalisation is taken over all nodes j belonging to the same graph. The graph-level embedding is then the weighted sum:

$$\mathbf{h}_G = \sum_{i \in V} \boldsymbol{\alpha}_i \circ \mathbf{h}_i \in \mathbb{R}^F. \tag{41}$$

Second, the graph embedding is passed through a final MLP, with D hidden layers of dimension F_s, each followed by a SiLU activation (replacing the ReLU used in OptiMate3B), and a plain linear output layer.

For the OptiMate3B dataset, the MLP produces $\hat{\mathbf{y}} = \gamma_{\text{out}}(\mathbf{h}_G) \in \mathbb{R}^{4002}$, where the first 2001 elements correspond to the imaginary part and the remaining elements to the real part. For the GNNOpt dataset, the output is instead $\hat{\mathbf{y}} = \gamma_{\text{out}}(\mathbf{h}_G) \in \mathbb{R}^{251}$, predicting either the real or imaginary part.

4. Results

4.1. OptiMate3B Dataset

We follow the same procedure as OptiMate3B, training first on IPA spectra and finetuning on the RPA dataset. Training details and dataset descriptions are provided in Sections B.1 and C.1. Code to reproduce results is available at https://github.com/khelverskovp/GotenNetOpt. We evaluate using three metrics over the frequency range $[0, \omega_{max}]$: the similarity coefficient (SC), mean squared error (MSE), and mean absolute error (MAE), defined as:

$$\text{SC}[X(\omega); Y(\omega)] = 1 - \frac{\int |X(\omega) - Y(\omega)| \, d\omega}{\int |Y(\omega)| \, d\omega}, \tag{42}$$

$$\text{MSE}[X(\omega); Y(\omega)] = \frac{1}{\omega_{\max}} \int |X(\omega) - Y(\omega)|^2 \, d\omega, \tag{43}$$

$$\text{MAE}[X(\omega); Y(\omega)] = \frac{1}{\omega_{\max}} \int |X(\omega) - Y(\omega)| \, d\omega. \tag{44}$$

where $X(\omega)$ is the target spectra and $Y(\omega)$ is the predicted spectra for $\text{Im}(\bar{\varepsilon})$ or $\text{Re}(\bar{\varepsilon})$.

Table 1. IPA and RPA test set results for the imaginary (Im) and real (Re) parts of the dielectric function, measured by mean absolute error (MAE), mean squared error (MSE), and similarity coefficient (SC). Each cell shows the mean with median in parentheses below. Best results are shown in bold.

	$\text{Im}(\bar{\varepsilon}_{\text{IPA}})$			$\text{Re}(\bar{\varepsilon}_{\text{IPA}})$			$\text{Im}(\bar{\varepsilon}_{\text{RPA}})$			$\text{Re}(\bar{\varepsilon}_{\text{RPA}})$		
MODEL	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑
OPTIMATE3B	0.265 (0.193)	0.744 (0.108)	0.883 (0.903)	0.276 (0.201)	0.677 (0.108)	0.882 (0.904)	0.195 (0.115)	0.539 (0.035)	0.906 (0.930)	0.199 (0.120)	0.393 (0.035)	0.910 (0.933)
$\text{GOTENNET}_{\text{OPT}}$	**0.238** **(0.173)**	**0.643** **(0.089)**	**0.894** **(0.913)**	**0.247** **(0.178)**	**0.601** **(0.089)**	**0.894** **(0.914)**	**0.164** **(0.093)**	**0.453** **(0.024)**	**0.919** **(0.942)**	**0.168** **(0.099)**	**0.318** **(0.024)**	**0.923** **(0.946)**

Table 1 reports mean and median values for each metric on both the IPA and RPA datasets, separated by imaginary and real parts. Additionally, we report results on using Voronoi graph construction and a scaled down version of $\text{GotenNet}_{\text{Opt}}$ in Section D. Median values are considerably lower than means reflecting a tail of outlier samples. RPA predictions are substantially better than IPA for both models due to the fine-tuning procedure and the greater physical accuracy of RPA spectra. The distribution of each evaluation metric on the RPA test set are shown in Figure 2.

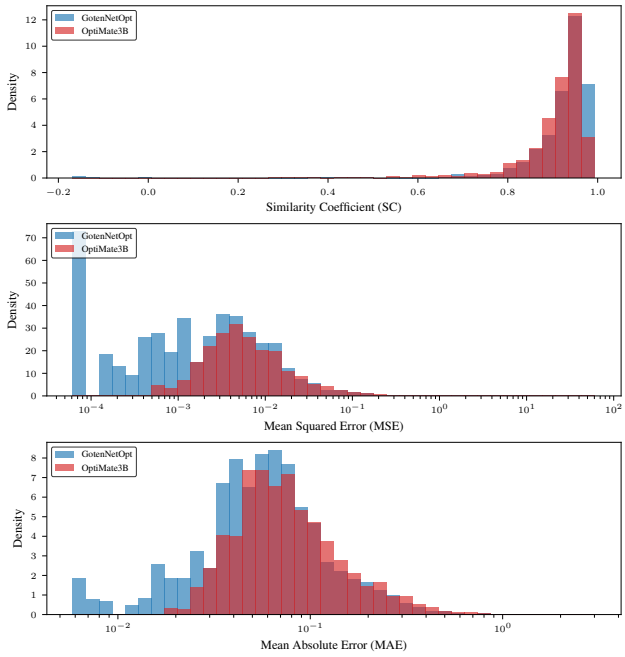

Figure 2. Density histograms of SC (top), MSE (middle), and MAE (bottom) across the RPA test samples for the two models. MSE and MAE are shown on logarithmic x-axes to accommodate their heavy-tailed distributions.

The difference in distributions is most pronounced for MAE and MSE, particularly the latter, as evidenced by the performance summary in Table 2. We report percentage improvements for both the full spectrum (0–20 eV) and the low-energy sub-spectrum (0–8 eV), which is most relevant for thin-film optics. The relative performance improvements are largest on the RPA dataset and in the 0–8 eV sub-spectrum.

Table 2. Percentage improvement of $\text{GotenNet}_{\text{Opt}}$ over OptiMate3B in median MAE, MSE, and SC across the test set, reported for the full 0–20 eV range and the low-energy 0–8 eV sub-range.

	METRIC	Δ% (0–20 eV)	Δ% (0–8 eV)
$\text{Im}(\bar{\varepsilon}_{\text{IPA}})$	MAE ↓	10.1%	8.9%
	MSE ↓	18.0%	16.0%
	SC ↑	1.1%	1.6%
$\text{Re}(\bar{\varepsilon}_{\text{IPA}})$	MAE ↓	11.0%	11.4%
	MSE ↓	17.9%	18.2%
	SC ↑	1.1%	1.1%
$\text{Im}(\bar{\varepsilon}_{\text{RPA}})$	MAE ↓	18.4%	19.2%
	MSE ↓	31.2%	31.3%
	SC ↑	1.2%	1.4%
$\text{Re}(\bar{\varepsilon}_{\text{RPA}})$	MAE ↓	17.9%	21.5%
	MSE ↓	31.6%	33.2%
	SC ↑	1.4%	1.3%
AVERAGE	MAE/MSE ↓	**19.6%**	**20.0%**
	SC ↑	**1.2%**	**1.4%**

Table 3 compares both models on predicting the static real permittivity by evaluating the median MAE and MSE on $\text{Re}(\bar{\varepsilon}(0))$ from the predicted spectra for both the IPA and RPA test sets.

Table 3. Test set MAE and MSE (median) on the static real permittivity, and the percentage improvement of $\text{GotenNet}_{\text{Opt}}$.

MODEL	$\text{Re}(\bar{\varepsilon}(0))_{\text{IPA}}$		$\text{Re}(\bar{\varepsilon}(0))_{\text{RPA}}$	
	MAE ↓	MSE ↓	MAE ↓	MSE ↓
OPTIMATE3B	0.131	0.017	0.121	0.015
$\text{GOTENNET}_{\text{OPT}}$	0.080	0.006	0.074	0.006
Δ%	**39.0%**	**62.8%**	**38.5%**	**62.1%**

Figure 3 shows predicted versus target spectra at the 10th, 40th, 60th and 90th SC percentiles of the RPA test set. For additional spectra plots, Figures 7–10 in Section D.4 show representative spectra for the cases where $\text{GotenNet}_{\text{Opt}}$ outperforms OptiMate3B, where OptiMate3B outperforms $\text{GotenNet}_{\text{Opt}}$, and where both models perform worst and best simultaneously. The imaginary and real parts are predicted with comparable accuracy across percentiles, consistent with the Kramers-Kronig relations.

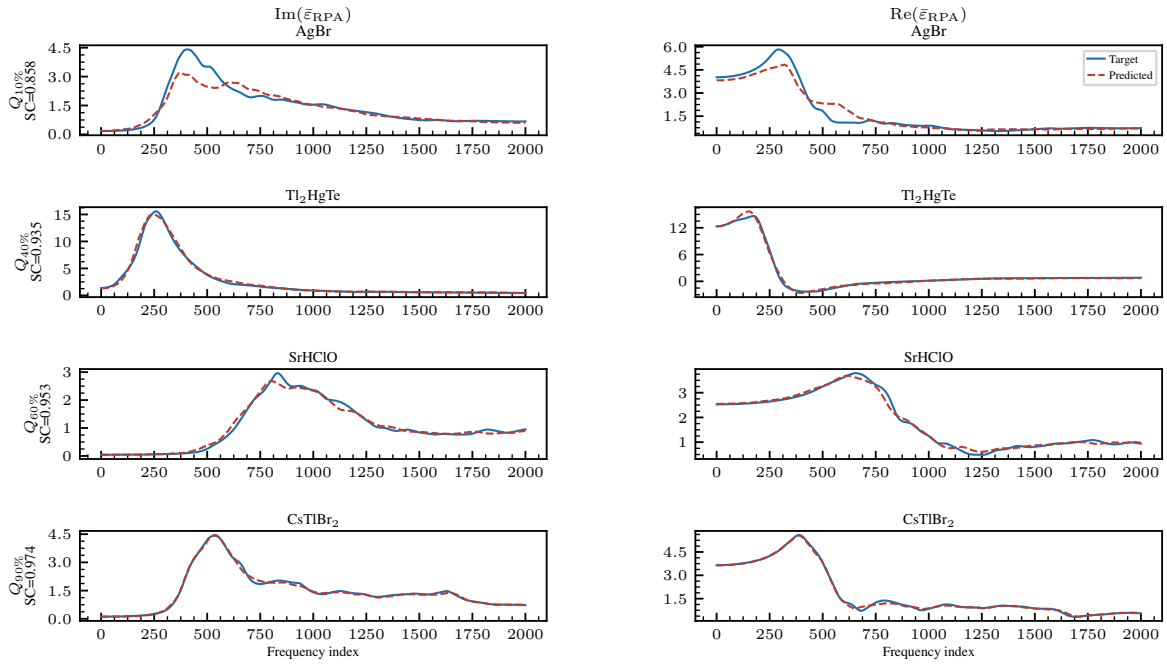

Figure 3. Predicted (dashed, red) vs. target (solid, blue) dielectric spectra at the 10th, 40th, 60th, and 90th percentiles of the RPA test set ranked by SC. Each row shows $\mathrm{Im}(\bar{\varepsilon}_{\mathrm{RPA}})$ and $\mathrm{Re}(\bar{\varepsilon}_{\mathrm{RPA}})$ components for a single structure (chemical formula shown above).

Figure 4 confirms this explicitly, having a strong-per sample correlation between SC on $\mathrm{Re}(\bar{\varepsilon})$ and $\mathrm{Im}(\bar{\varepsilon})$.

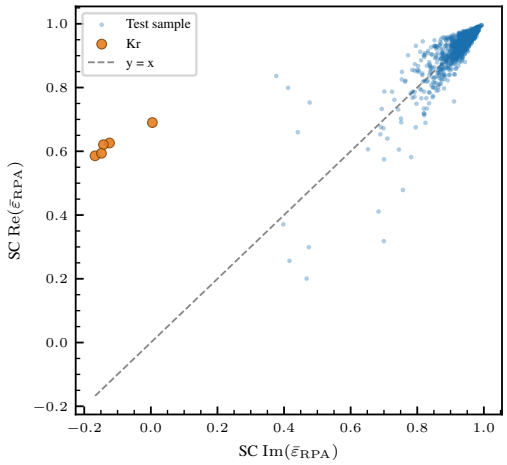

Figure 4. Kramers-Kronig consistency check. Per-sample SC for $\mathrm{Re}(\bar{\varepsilon})$ against SC for $\mathrm{Im}(\bar{\varepsilon})$.

Five outliers are Kr-only crystals. Kr appears in only 2 training samples with no other group-18 elements (Figure 5), leaving its atom embedding poorly informed; Kr also forms no covalent bonds under ordinary conditions, meaning the covalent radius feature is physically meaningless. These outliers are an extreme case of a broader log-linear relationship between prediction accuracy and per-element training count (Figure 6), consistent with findings reported for OptiMate3B (Grunert et al., 2026c).

4.2. GNNOpt Dataset

Table 4 presents the performance of GotenNet$_{\mathrm{Opt}}$ on the GNNOpt dataset; training details and dataset descriptions are provided in Sections B.2 and C.2. Following Hung et al. (2024), we report MSE and R^2 on the test set. R^2 is evaluated on the energy-weighted mean $\bar{X}_i = \sum_k \omega_k X_i(\omega_k) / \sum_k \omega_k$ of each predicted and target spectrum,

$$R^2[X;Y] = 1 - \frac{\sum_{i=1}^{N} \left(\bar{X}_i - \bar{Y}_i\right)^2}{\sum_{i=1}^{N} \left(\bar{X}_i - \overline{X}\right)^2}, \tag{45}$$

where $\bar{Y}_i$ is defined analogously and $\overline{X}$ is the mean of $\bar{X}_i$ over the test set. MSE is computed over all $K = 251$ frequency points, consistent with the definition in Equation (43).

Table 4. Test set results on the GNNOpt dataset measured by MSE and coefficient of determination (R^2) along with percentage improvement compared to GNNOpt.

	METRIC	GNNOPT	GOTENNET$_{\mathrm{OPT}}$	Δ%
IM($\bar{\varepsilon}$)	MSE ↓	1.765	1.371	22.3%
	R^2 ↑	0.737	0.864	17.2%
RE($\bar{\varepsilon}$)	MSE ↓	2.363	1.907	19.3%
	R^2 ↑	0.715	0.806	12.7%

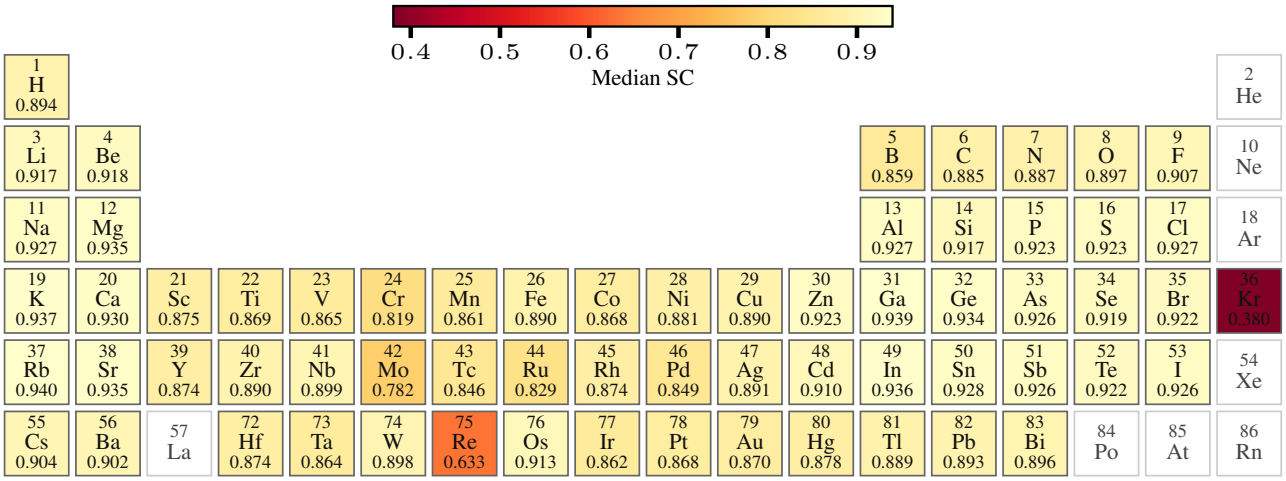

Figure 5. Median SC for each compound in the IPA test set that contains a given element. Elements that are not present in any compounds are colored white.

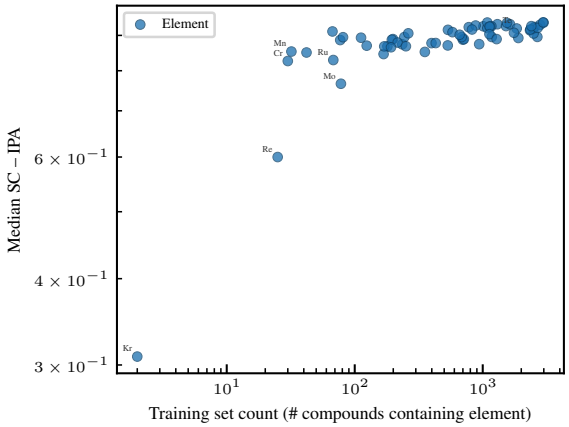

Figure 6. Log-log plot of median SC for all compounds in the test set that contain a given element as a function of how often that element occurs in the training set for the IPA data.

5. Discussion

This work suggests that equivariant GNNs can meaningfully improve optical spectra prediction, with the most pronounced gains in the 0–8 eV range and for prediction of the static real permittivity. Gains are also proportionally larger on the RPA dataset than on the IPA dataset, which may reflect that physically accurate spectra encoding richer many-body interactions benefit more from expressive geometric representations

The adaptation of GotenNet involved two main changes: incorporating covalent radius embeddings into the scalar node feature initialization, and replacing the readout layer with the OptiMate3B readout design. Covalent radius is a physically motivated feature, though it is not universally informative, and more generally, the most useful atomic features may depend on the chemical subspace being screened. Future work could explore other physical properties as node features, tailored to the target application. The readout architecture is similarly not necessarily optimal, and other designs remain to be explored. More broadly, other equivariant GNN architectures have not yet been tested on this task.

Despite performance improvements, the dominant failure mode is still distributional: prediction accuracy correlates with per-element training frequency. Discarding these outliers, the similarity coefficient metric suggests the model is approaching highly accurate predictions on the current data. Hence, for elements well represented in the training data, the model is viable as a large-scale screening tool, but for structures containing underrepresented elements, predictions should be treated with caution, and expanding the training set would be necessary before using the model on poorly covered chemical spaces.

6. Conclusion

We have presented GotenNet$_{\text{Opt}}$, an adaptation of the equivariant GNN GotenNet for optical spectra prediction. The results show that using an equivariant architecture leads to consistent performance increases, indicating that geometric equivariance is a beneficial inductive bias for this task. Gains are most pronounced in the low-energy range and in the static real permittivity, quantities of direct relevance for high-throughput screening of thin-film optical materials, and are proportionally larger on the RPA dataset, suggesting that geometric expressiveness matters most when the training data is physically accurate.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Aykent, S. and Xia, T. Gotennet: Rethinking efficient 3d equivariant graph neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=5wxCQDtbMo.

Batatia, I., Kovács, D. P., Simm, G. N. C., Ortner, C., and Csányi, G. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields, 2023. URL https://arxiv.org/abs/2206.07697.

Brandstetter, J., Hesselink, R., van der Pol, E., Bekkers, E. J., and Welling, M. Geometric and physical quantities improve E(3) equivariant message passing. *CoRR*, abs/2110.02905, 2021. URL https://arxiv.org/abs/2110.02905.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. *CoRR*, abs/1704.01212, 2017. URL http://arxiv.org/abs/1704.01212.

Ginter, C., Choudhary, K., and Mandal, S. Accelerated prediction of dielectric functions in solar cell materials with graph neural networks, 2025. URL https://arxiv.org/abs/2510.08738.

Grunert, M., Großmann, M., and Runge, E. Data for evaluating OptiMate3B. 1 2026a. doi: 10.6084/m9.figshare.30257551.v1. URL https://figshare.com/articles/dataset/Data_for_evaluating_OptiMate3B/30257551.

Grunert, M., Großmann, M., and Runge, E. Dielectric Function of Semiconductors and Insulators under the Independent Particle Approximation and the Random Phase Approximation. 1 2026b. doi: 10.6084/m9.figshare.30257689.v1. URL https://figshare.com/articles/dataset/Dielectric_Function_of_Semiconductors_and_Insulators_under_the_Independent_Particle_Approximation_and_the_Random_Phase_Approximation/30257689.

Grunert, M., Großmann, M., and Runge, E. Optical spectra prediction using three-body information. *Applied Physics Letters*, 128(5):053302, 02 2026c. ISSN 0003-6951. doi: 10.1063/5.0305586. URL https://doi.org/10.1063/5.0305586.

Hohenberg, P. and Kohn, W. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964. doi: 10.1103/PhysRev.136.B864. URL https://link.aps.org/doi/10.1103/PhysRev.136.B864.

Hung, N. T., Okabe, R., Chotrattanapituk, A., and Li, M. Ensemble-embedding graph neural network for direct prediction of optical spectra from crystal structure, 2024. URL https://arxiv.org/abs/2406.16654.

Ibrahim, A. and Ataca, C. Prediction of frequency-dependent optical spectrum for solid materials: A multi-output & multi-fidelity machine learning approach, 2024. URL https://arxiv.org/abs/2407.07736.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., and Persson, K. A. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 07 2013. ISSN 2166-532X. doi: 10.1063/1.4812323. URL https://doi.org/10.1063/1.4812323.

Klicpera, J., Giri, S., Margraf, J. T., and Günnemann, S. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *CoRR*, abs/2011.14115, 2020. URL https://arxiv.org/abs/2011.14115.

Kohn, W. and Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

Le, T., Noé, F., and Clevert, D.-A. Equivariant graph attention networks for molecular property prediction, 2022. URL https://arxiv.org/abs/2202.09891.

Liao, Y.-L. and Smidt, T. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs, 2023. URL https://arxiv.org/abs/2206.11990.

Liu, Y., Wang, L., Liu, M., Zhang, X., Oztekin, B., and Ji, S. Spherical message passing for 3d graph networks. *CoRR*, abs/2102.05013, 2021. URL https://arxiv.org/abs/2102.05013.

Lu, C., Liu, Q., Wang, C., Huang, Z., Lin, P., and He, L. Molecular property prediction: A multilevel quantum interactions modeling perspective, 2019. URL https://arxiv.org/abs/1906.11081.

Onida, G., Reining, L., and Rubio, A. Electronic excitations: density-functional versus many-body green's-function approaches. *Rev. Mod. Phys.*, 74:601–659, Jun 2002. doi: 10.1103/RevModPhys.74.601. URL https://link.aps.org/doi/10.1103/RevModPhys.74.601.

Satorras, V. G., Hoogeboom, E., and Welling, M. E(n) equivariant graph neural networks. *CoRR*, abs/2102.09844, 2021. URL https://arxiv.org/abs/2102.09844.

Schütt, K. T., Unke, O. T., and Gastegger, M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. *CoRR*, abs/2102.03150, 2021. URL https://arxiv.org/abs/2102.03150.

Schütt, K. T., Kindermans, P.-J., Sauceda, H. E., Chmiela, S., Tkatchenko, A., and Müller, K.-R. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions, 2017. URL https://arxiv.org/abs/1706.08566.

Simeon, G. and de Fabritiis, G. Tensornet: Cartesian tensor representations for efficient learning of molecular potentials, 2023. URL https://arxiv.org/abs/2306.06482.

Thölke, P. and Fabritiis, G. D. Torchmd-net: Equivariant transformers for neural network based molecular potentials. *CoRR*, abs/2202.02541, 2022. URL https://arxiv.org/abs/2202.02541.

Toll, J. S. Causality and the dispersion relation: logical foundations. *Physical review*, 104(6):1760, 1956.

Yang, R. X., Horton, M. K., Munro, J., and Persson, K. A. High-throughput optical absorption spectra for inorganic semiconductors, 2022. URL https://arxiv.org/abs/2209.02918.

A. Density functional theory

Optical response tensors of materials are calculated through quantum mechanical perturbation theory which necessitates an accurate description of the electronic ground state of the material to be perturbed first. For many-body systems, that is, systems containing many electrons, the ground state of the system is typically obtained through DFT. Strictly speaking, determining the ground state of a quantum system with N electrons means determining the ground state many-body wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$; however, within DFT the ground state can be understood entirely through the ground state electron density $\rho(\mathbf{r})$ which is otherwise considered an observable that can be calculated directly from the many-body wave function. While many-body wave functions describe the ground states of interacting electrons governed by the Schrödinger equation, most modern formulations of DFT use an auxiliary set of single-particle wave functions which obey the Kohn-Sham equations (Kohn & Sham, 1965). These single-particle wave functions describe a fictitious set of non-interacting electrons; however, the Kohn-Sham equations are cleverly constructed such that they (in theory) yield the same electron density as the true many-body wave function. Density functional theory thus discards the staggeringly complicated many-body wave function in favor of the much simpler set of single-particle wave functions as both will construct the true ground state electron particle density ρ.

Although DFT is a popular and versatile method for obtaining ground state information, trouble arises when considering excited states of the system as only the ground state can be described entirely by the density, while excited states need the true many-body wave function. This leads to a set of perturbation theory methods tiered by how diligently they account for many-body effects, which are otherwise missing in the non-interacting Kohn-Sham system. The simplest of which is the independent-particle approximation, which simply treats the optical response as a sum of optical responses of the independent-particle system.

B. Training Details

B.1. Training on the OptiMate3B Dataset

Table 5 shows the hyperparameters used for training on the OptiMate3B dataset. Training proceeds in two stages. In the first stage, the model is trained to predict IPA dielectric spectra. In the second stage, the best IPA checkpoint is used to initialise the model, which is then fine-tuned on RPA dielectric spectra.

Both stages use the AdamW optimiser with a learning rate of 1×10^{-5} and weight decay of 0.01. Gradients are clipped to a maximum norm of 10. The learning rate is reduced on plateau by a factor of 0.5, with a minimum learning rate of 10^{-8} and 10^{-9} for IPA and RPA training respectively. The plateau scheduler patience is set to 15 epochs. Early stopping is applied with a patience of 50 epochs in both stages.

The loss function is mean absolute error (MAE) in both stages. To stabilize early stopping, an exponential moving average (EMA) of the validation loss is maintained with smoothing factor $\alpha = 0.9$. Validation loss spikes exceeding $2\times$ the current smoothed value are ignored when updating the EMA, preventing transient instabilities from triggering premature learning rate decay or stopping.

B.2. Training on the GNNOpt dataset

Table 6 shows the hyperparameters used for training on the GNNOpt dataset. We follow the GNNOpt training setup, including their graph construction procedure and the scaling and interpolation of the target spectra. Unlike the OptiMate3B dataset setup, where the real and imaginary parts of the dielectric function are predicted jointly, here a separate model is trained and evaluated for each component. The optical spectra are interpolated onto a uniform grid of 251 points over the energy range $0 \leq \hbar\omega \leq 50\,\text{eV}$. Spectra are then normalized by a global scale factor computed as the median of the per-sample maximum values across the training set. The loss function is mean squared error (MSE) on the normalized spectra.

Graphs are constructed by placing edges between all atoms within a cutoff radius of $r_c = 6.0\,\text{Å}$. All experiments use a fixed random seed of 12 (same seed used in the original GNNOpt experiments). Training is performed for 200 epochs with a batch size of 8, and gradients are clipped to a maximum ℓ_2 norm of 1.0.

Table 5. Hyperparameters for the two-stage training procedure (IPA pre-training and RPA fine-tuning). Shared architecture parameters apply to both stages. (S) are parameter values used for the smaller variant.

Category	**Hyperparameter**	**Value**
Optimisation	Learning rate	1×10^{-5} (S: 1×10^{-4})
	Minimum learning rate	1×10^{-8} (RPA: 1×10^{-9})
	Weight decay	0.01
	Gradient clipping	10
	Batch size	32
	Max epochs	1000
Scheduler & Stopping	LR scheduler	ReduceLROnPlateau
	LR reduction factor	0.5
	LR patience	15 epochs
	Early stopping patience	50 epochs
Loss	Loss function	MAE
	Val. loss EMA α	0.9
	Spike filter factor	2.0
Data Split	Training fraction	80%
	Validation fraction	10%
	Test fraction	10%
Architecture	Feature embedding dim (F)	1024 (S: 256)
	Message passing rounds (T)	6
	Radial basis functions	64
	Cutoff radius	6.0 Å
	Max degree ($L_{\max}$)	2
	Attention heads (H)	8
	Dropout	0.1
	Readout depth (D)	4 (S: 2)
	Readout hidden dim (F_s)	8192 (S: 256)

Table 6. Hyperparameters for training on the GNNOpt dataset.

Category	Hyperparameter	Value
Optimisation	Optimizer	AdamW
	Learning rate	6×10^{-4}
	Weight decay	0.015
	Gradient clipping	1.0
	Batch size	8
	Epochs	200
Scheduler	LR scheduler	Exponential
	LR decay (γ)	0.98
Loss & Seed	Loss function	MSE
	Random seed	12
Architecture	Feature embedding dim (F)	256
	Message passing rounds (T)	3
	Radial basis functions	64
	Cutoff radius	6.0 Å
	Max degree ($L_{\max}$)	2
	Attention heads (H)	8
	Dropout	0.3
	Readout depth (D)	1
	Readout hidden dim (F_s)	256

C. Dataset Descriptions

C.1. OptiMate3B Dataset

The OptiMate3B datasets consists of an IPA dataset with 23,834 materials and an RPA dataset with 10,533 materials, which we obtain from Grunert et al. (2026a) and Grunert et al. (2026b). The structures were originally drawn from the Alexandria database and filtered according to the following criteria:

- no elements beyond Bismuth and no Lanthanides;
- a direct or indirect bandgap above 50 meV;
- an indirect bandgap below 100 meV or above 10 eV
- energy above hull below 150 meV per atom;
- symmetry of at least space group 75 (tetragonal or higher);
- at most 8 atoms per primitive unit cell.

For the RPA dataset, structures were further restricted to at most 4 atoms per unit cell. We adopt the same split procedure as Grunert et al. (2026c), splitting by unique chemical composition in an 80:10:10 ratio yielding the exact same splits as in the original paper of 19,036, 2413, and 2385 materials for training, validation and test sets respectively. The same compositional split is preserved for the RPA dataset, giving 8,414, 1,040, and 1,079 materials.

C.2. GNNOpt Dataset

The GNNOpt dataset consists of 944 crystal structures obtained from the GNNOpt GitHub repository. The optical spectra were originally computed using IPA by (Yang et al., 2022), with structures sourced from the Materials Project (Jain et al., 2013). The dataset was constructed by filtering for materials satisfying the following criteria:

- an energy band gap between 0 and 5.0 eV;
- at most 10 atoms per primitive unit cell.

We adopt the same train/validation/test split as the original GNNOpt paper, taken directly from the repository, corresponding to an 80%, 10%, 10% division of 733, 97, and 110 materials, respectively.

D. Additional Results and Analysis

D.1. Covalent Radius Embeddings Ablation

The effect of covalent radius embeddings is modest but generally positive, as shown in Table 7. The clearest gains are in the RPA targets, where MAE and MSE improve consistently (up to 3.13% and 4.00% respectively). For the IPA targets the picture is more mixed; MAE improves slightly but the median MSE is marginally worse ($+1.14\%$), suggesting that covalent radius information is more beneficial for the RPA computed dielectric function. Overall, the average median improvement across MAE and MSE is 1.18% and 0.14% for SC.

Table 7. IPA and RPA test set results ablating covalent radius embeddings in GotenNet$_{\text{Opt}}$. Each cell shows the mean with median in parentheses below. Best results are shown in bold. The $\Delta\%$ row shows the relative change of GotenNet$_{\text{Opt}}$ compared to the model without covalent radius embeddings based on the medians.

	Im($\bar{\varepsilon}_{\text{IPA}}$)			Re($\bar{\varepsilon}_{\text{IPA}}$)			Im($\bar{\varepsilon}_{\text{RPA}}$)			Re($\bar{\varepsilon}_{\text{RPA}}$)		
Model	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑
w/o cov. radius	0.241 (0.174)	0.664 (0.088)	0.893 (0.912)	0.251 (0.180)	0.623 (0.088)	0.893 (0.913)	0.167 (0.096)	0.465 (0.025)	0.920 (0.940)	0.171 (0.102)	0.332 (0.024)	0.923 (0.945)
GotenNet$_{\text{Opt}}$	**0.238** **(0.173)**	**0.643** **(0.089)**	**0.894** **(0.913)**	**0.247** **(0.178)**	**0.601** **(0.089)**	**0.894** **(0.914)**	**0.164** **(0.093)**	**0.453** **(0.024)**	**0.919** **(0.942)**	**0.168** **(0.099)**	**0.318** **(0.024)**	**0.923** **(0.946)**
$\Delta\%$	0.57%	+1.14%	0.11%	1.11%	+1.14%	0.11%	3.13%	4.00%	0.21%	2.94%	0.00%	0.11%

D.2. Effect of Model Size on Performance

Table 8 reports performance on the OptiMate3B dataset of GotenNet$_{\text{Opt}}$-S, a smaller variant (16M parameters) using readout depth 2, feature dimension 256, and readout hidden dimension 256, compared to GotenNet$_{\text{Opt}}$ (474M). Despite its reduced size, GotenNet$_{\text{Opt}}$-S trails GotenNet$_{\text{Opt}}$ by only 6.9% in median MAE/MSE and 0.3% in median SC on average, suggesting that performance gains stem primarily from architectural differences rather than increased parameter count. Nevertheless, since inference is run once per crystal in the intended screening use case, the computational overhead of the larger model is negligible, making it the preferred choice over the smaller variant.

Table 8. IPA and RPA test set results comparing GotenNet$_{\text{Opt}}$-S (small variant) and GotenNet$_{\text{Opt}}$ (main model). Each cell shows the mean with median in parentheses below. Best results are shown in bold. The $\Delta\%$ row shows the relative improvement of GotenNet$_{\text{Opt}}$ compared to GotenNet$_{\text{Opt}}$-S based on the medians.

	Im($\bar{\varepsilon}_{\text{IPA}}$)			Re($\bar{\varepsilon}_{\text{IPA}}$)			Im($\bar{\varepsilon}_{\text{RPA}}$)			Re($\bar{\varepsilon}_{\text{RPA}}$)		
Model	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑
GotenNet$_{\text{Opt}}$-S	0.243 (0.178)	0.679 (0.092)	0.893 (0.910)	0.252 (0.183)	0.640 (0.092)	0.893 (0.913)	0.170 (0.102)	0.452 (0.028)	0.918 (0.938)	0.173 (0.105)	0.326 (0.028)	0.921 (0.943)
GotenNet$_{\text{Opt}}$	**0.238** **(0.173)**	**0.643** **(0.089)**	**0.894** **(0.913)**	**0.247** **(0.178)**	**0.601** **(0.089)**	**0.894** **(0.914)**	**0.164** **(0.093)**	**0.453** **(0.024)**	**0.919** **(0.942)**	**0.168** **(0.099)**	**0.318** **(0.024)**	**0.923** **(0.946)**
$\Delta\%$	2.81%	3.26%	0.33%	2.73%	3.26%	0.11%	8.82%	14.29%	0.43%	5.71%	14.29%	0.32%

D.3. Effect of Graph Construction on Performance

We compare GotenNet$_{\text{Opt}}$ trained with Voronoi-based graph construction against the distance-based cutoff of 6 Å on the OptiMate3B dataset. As shown in Table 9, the cutoff-based method consistently outperforms the Voronoi-based method across all targets and metrics, with an average combined median improvement of 4.3% in MAE and MSE and 0.3% in SC. While the improvements are modest, they are consistent across all four targets and both the IPA and RPA datasets, suggesting that the fixed-radius cutoff provides a slightly better graph structure for this task than the Voronoi approach.

Table 9. IPA and RPA test set results comparing GotenNet$_{\text{Opt}}$ on Voronoi vs. distance based graph construction. Each cell shows the mean with median in parentheses below. Best results are shown in bold. The $\Delta\%$ row shows the relative change of the 6 Å cutoff method compared to Voronoi based on the medians.

	IM($\bar{\varepsilon}_{\text{IPA}}$)			RE($\bar{\varepsilon}_{\text{IPA}}$)			IM($\bar{\varepsilon}_{\text{RPA}}$)			RE($\bar{\varepsilon}_{\text{RPA}}$)		
MODEL	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑	MAE ↓	MSE ↓	SC ↑
VORONOI	0.243 (0.178)	0.678 (0.094)	0.892 (0.910)	0.254 (0.185)	0.635 (0.094)	0.892 (0.911)	0.169 (0.099)	0.497 (0.026)	0.917 (0.939)	0.173 (0.103)	0.352 (0.026)	0.921 (0.944)
6 Å CUTOFF	**0.238** **(0.173)**	**0.643** **(0.089)**	**0.894** **(0.913)**	**0.247** **(0.178)**	**0.601** **(0.089)**	**0.894** **(0.914)**	**0.164** **(0.093)**	**0.453** **(0.024)**	**0.919** **(0.942)**	**0.168** **(0.099)**	**0.318** **(0.024)**	**0.923** **(0.946)**
$\Delta\%$	2.81%	5.32%	0.33%	3.78%	5.32%	0.33%	6.06%	7.69%	0.32%	3.88%	7.69%	0.21%

D.4. Additional spectra plots

Top 5 GotenNet wins as meaured by ΔSC Figure 7 shows the predicted and target dielectric spectra for the five samples where GotenNet$_{\text{Opt}}$ most outperforms OptiMate3B as measured by ΔSC between the two models on the RPA test set.

Top 5 OptiMate3B wins as measured by ΔSC. Figure 8 shows the predicted and target dielectric spectra for the five samples where OptiMate3B most outperforms GotenNet$_{\text{Opt}}$ as measured by ΔSC.

Top 5 samples where both models fail. Figure 9 shows the predicted and target dielectric spectra for the five samples with the lowest average SC across both models, i.e. the cases where both models perform worst.

Top 5 samples where both models succeed. Figure 10 shows the predicted and target dielectric spectra for the five samples with the highest average SC across both models, i.e. the cases where both models perform best.

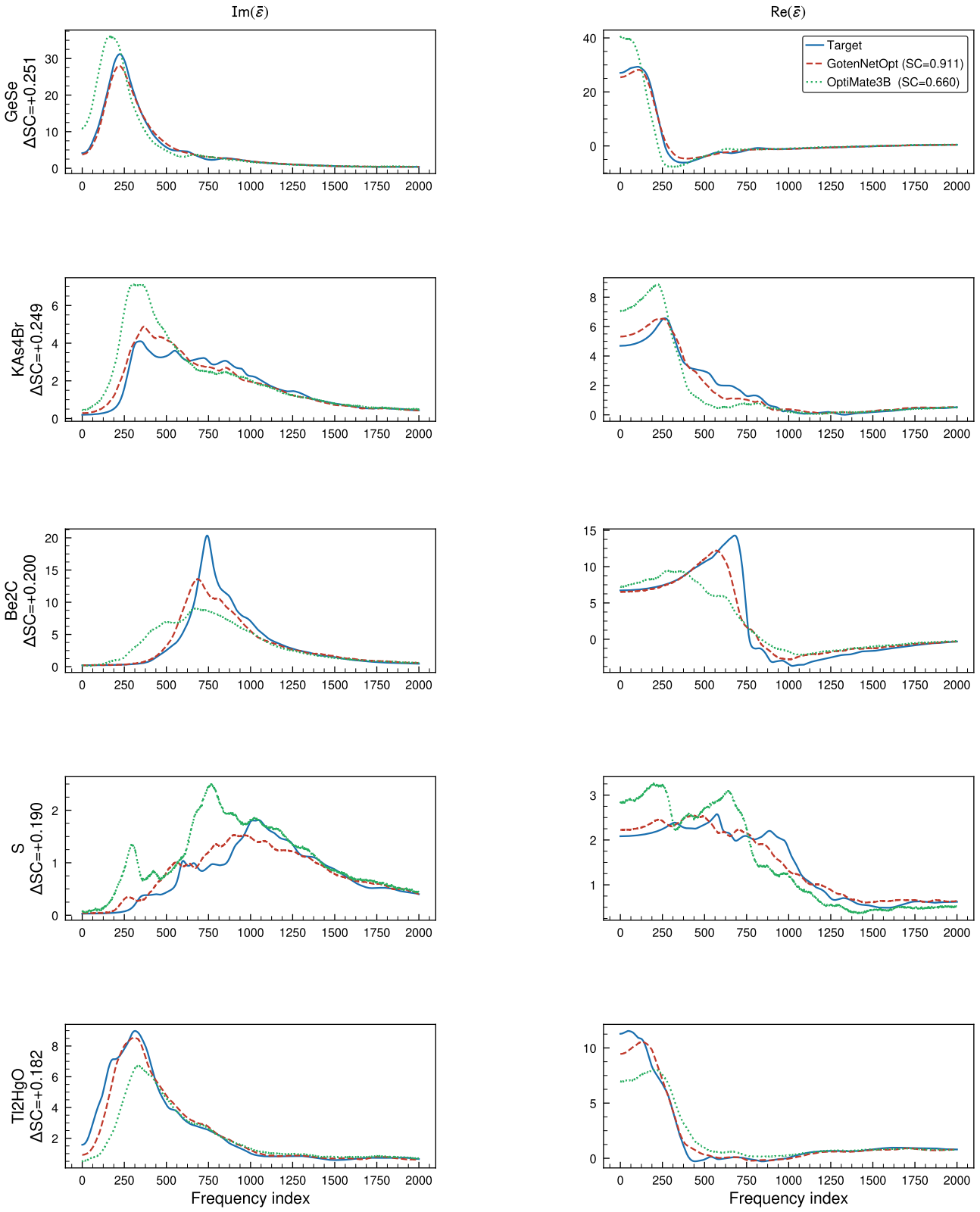

Figure 7. Predicted of GotenNet$_{\text{Opt}}$ (dashed, red) and OptiMate3B (dashed, green) vs. target (solid, blue) dielectric spectra for the top 5 wins of GotenNet$_{\text{Opt}}$ over OptiMate3B as measured by ΔSC between the two models on the RPA test set.

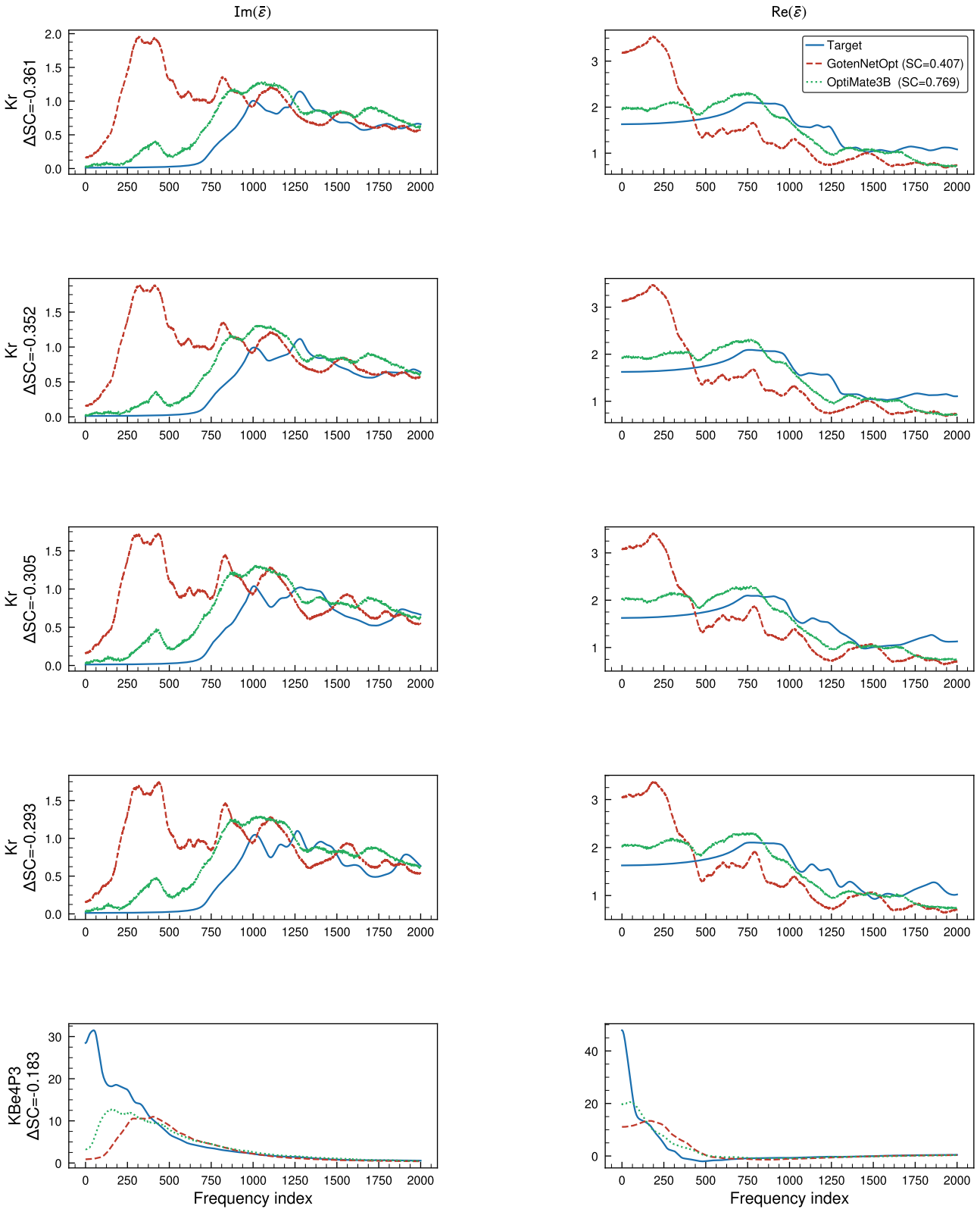

Figure 8. Predicted spectra of $\text{GotenNet}_{\text{Opt}}$ (dashed, red) and OptiMate3B (dashed, green) vs. target (solid, blue) dielectric spectra for the top 5 wins of OptiMate3B over $\text{GotenNet}_{\text{Opt}}$ as measured by ΔSC on the RPA test set.

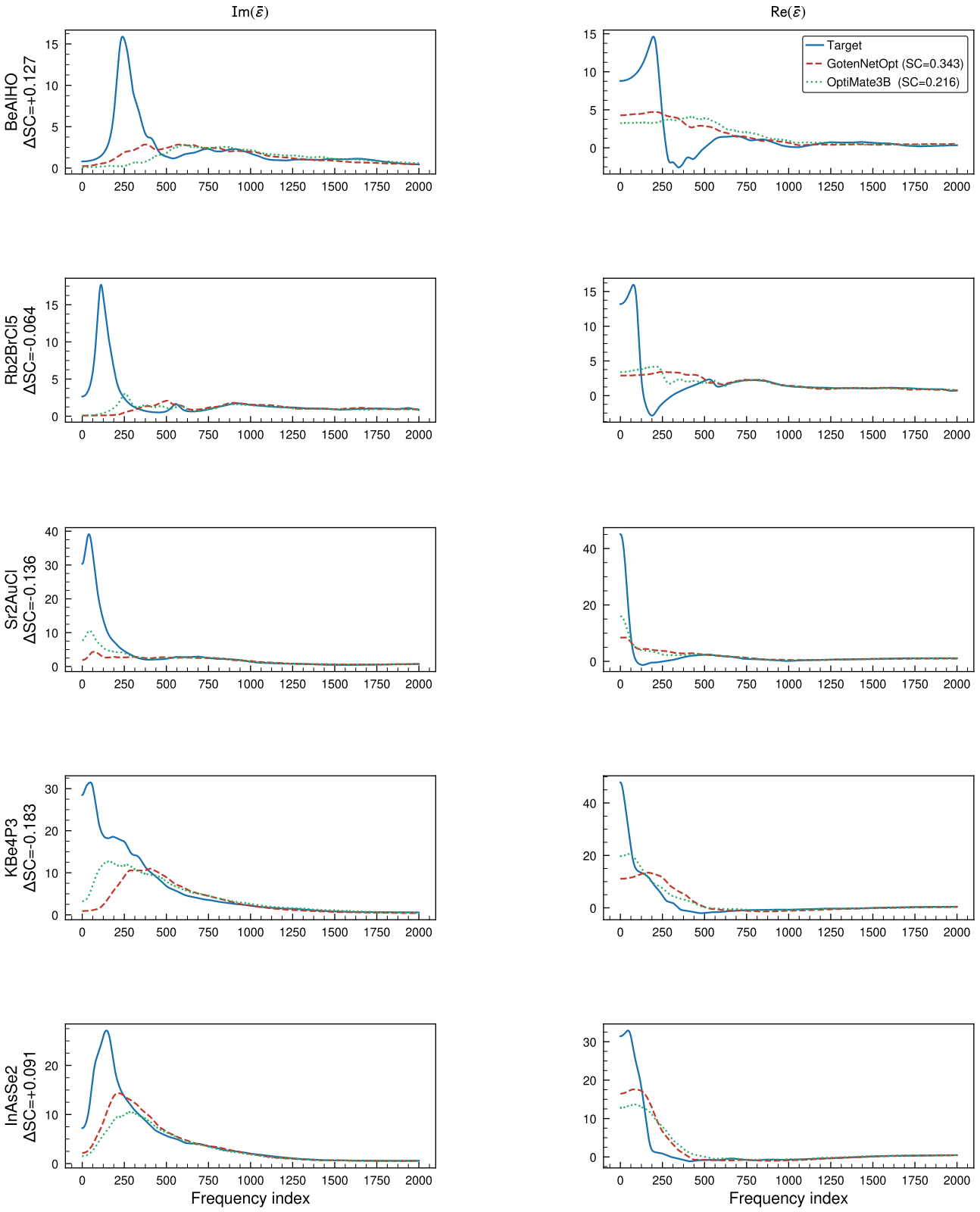

Figure 9. Predicted spectra of $\text{GotenNet}_{\text{Opt}}$ (dashed, red) and OptiMate3B (dashed, green) vs. target (solid, blue) for the 5 samples with the lowest average SC across both models (SC $<$ 0.80) on the RPA test set, illustrating failure cases shared by both models.

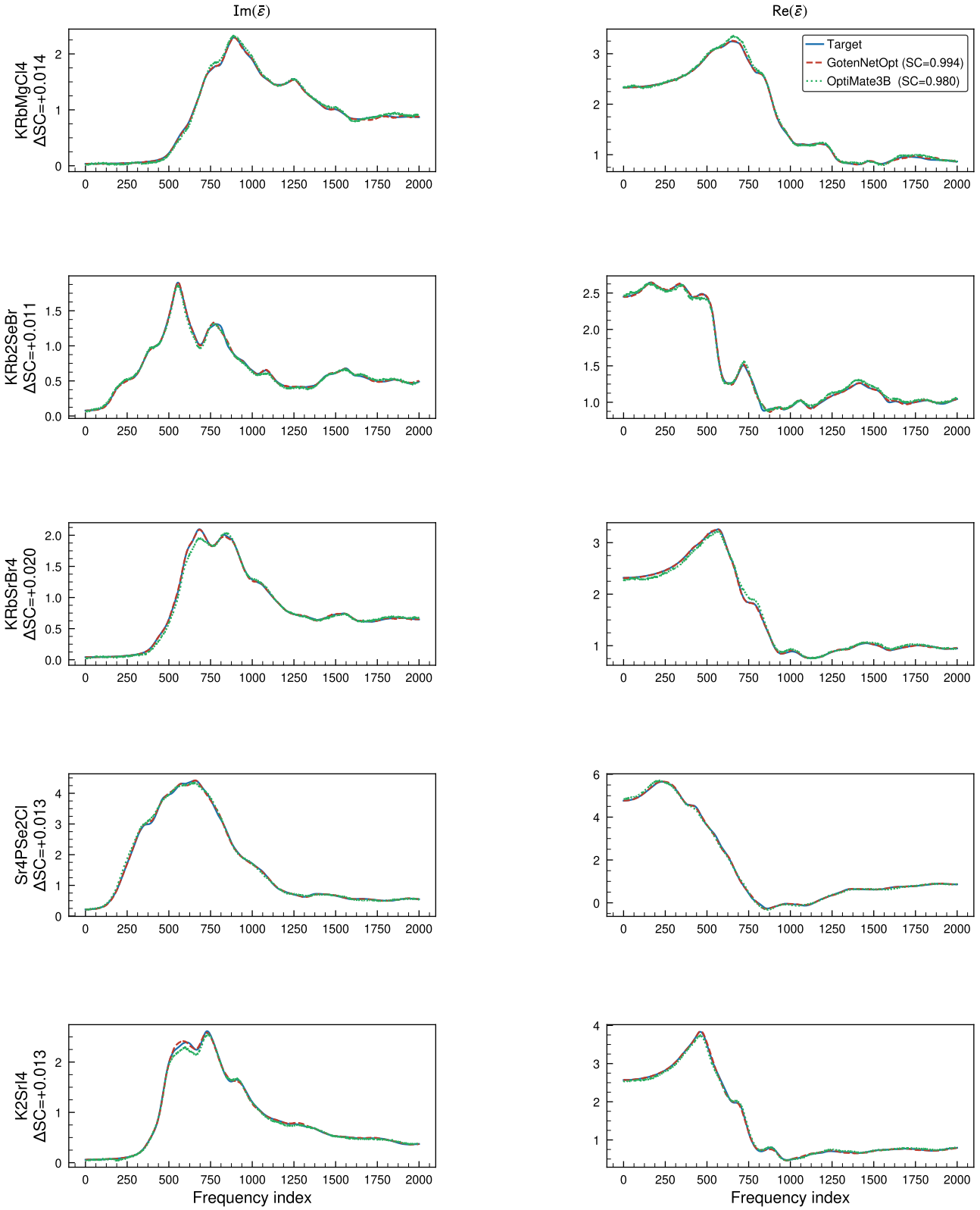

Figure 10. Predicted spectra of $\text{GotenNet}_{\text{Opt}}$ (dashed, red) and OptiMate3B (dashed, green) vs. target (solid, blue) for the 5 samples with the highest average SC across both models (SC > 0.95) on the RPA test set, illustrating representative cases where both models achieve high fidelity.