Supplementary Information: Multi-Task Scattering-Model Classification and Parameter Regression of Nanostructures from Small-Angle Scattering Data

Batuhan Yildirim,[abc] James Doutch,[b] and Jacqueline M. Cole*[abc]

Implementation and Training Details

Input Embeddings

Before being processed by the encoder, the discretized log-quotient transformed $I(q)$ (of sequence length S and 256 bins), $\{\mathbf{x} \in \mathbb{Z}^S \mid 0 \leq x_i < 256\}$, is tokenized and position encoded.

In tokenizing an $I(q)$ sequence, we bin the intensity values globally across all q-values and assign a unique $F = 1024$ dimensional vector to each bin. These unique vectors are initialized by randomly sampling them from a unit Gaussian distribution then projecting with a learned transformation,

$$\mathbf{t}_i = \boldsymbol{W}_e \mathbf{t}_i^0,\ \mathbf{t}_i^0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{1}$$

where $\boldsymbol{W}_e$ is the embedding projection matrix and $\mathbf{t}_i^0$ and $\mathbf{t}_i$ are the initialized and projected token vectors, respectively. The tokens representing the binned $I(q)$ values at each sequence index are stacked to create a token matrix $\boldsymbol{T} \in \mathbb{R}^{S \times F}$. Positional embeddings[1] assign a fixed sinusoidal embedding to the sequence index of each q-value. These are fixed sine and cosine vectors whose frequencies depend on the index of the q-value they correspond to. Similar to the tokenized $I(q)$ matrix, these are stacked along the sequence dimension to create a position embedding matrix, $\boldsymbol{P} \in \mathbb{R}^{S \times F}$. The input embedding $\boldsymbol{X}$ is the sum of the tokenized and positional embeddings,

$$\boldsymbol{X} = \boldsymbol{T} + \boldsymbol{P}. \tag{2}$$

Including $\boldsymbol{P}$ in the input embedding allows the model to make use of information about the absolute and relative positions of different values in the input sequence. This is necessary when modeling sequences as transformers are invariant to permutations in the sequence dimension and simply express output tokens as a weighted linear combination of all input tokens. Hence there are no natural sequential or spatial inductive biases built into the architecture such as in recurrent and convolutional neural networks.

Multitask Objective Function

The objective function is constructed as follows. The scattering-model classification loss for a single instance is defined as

$$\mathscr{L}_{\text{clf}} = \sum_{k=1}^{K} m_k \log p(\hat{m}_k), \tag{3}$$

where K is the number of scattering-model classes in the training dataset, $p(\hat{m}_k)$ is the predicted probability that an input belongs to the kth scattering-model class, and m_k is the ground-truth which equals one for the correct scattering-model class and zero otherwise. The scattering-model parameter regression loss for a single instance is defined as

$$\mathscr{L}_{\text{reg}} = \frac{1}{R} \sum_{j=1}^{R} (y_j - \hat{y}_j)^2, \tag{4}$$

where R is the number of scattering-model parameters in the training dataset, $\hat{y}_j$ is the predicted value of the jth scattering-model parameter, and y_j is the corresponding ground-truth value. The index j in the regression loss spans all scattering-model parameters across all model classes in the training data. As a result, the regression decoder predicts all parameters in the training data for a given input. To deal with this during training, a given y has all parameter values that do not correspond to its scattering-model class set to NaN. When calculating the regression loss for a single instance, we take the mean over the non-NaN values in y, meaning that the predictions that correspond to the NaN values in the ground truth do not contribute to parameter updates during back-propagation of gradients in the neural network. Finally, the overall objective, averaged over many instances, is

$$L = \frac{1}{N} \sum_{i=1}^{N} \lambda_{\text{clf}} \mathscr{L}_{\text{clf}} + \lambda_{\text{reg}} \mathscr{L}_{\text{reg}}, \tag{5}$$

where N is the number of training samples in a batch during training or the total number of samples in the validation or test datasets. λ_{clf} and λ_{reg} are hyperparameters that control the weighting of the classification and regression terms of the objective function. We used values of 1.0 for both of these terms.

Model Hyperparameters

Table 1 Encoder hyperparameters.

latent dim.	1024
num. latents	48
dropout	0.05
cross att. widening factor	1
cross att. heads	4
self att. blocks	12
self att. widening factor	1
self att. heads	4

[a] Cavendish Laboratory, Department of Physics, University of Cambridge, J.J. Thomson Avenue, Cambridge, CB3 0HE, United Kingdom
[b] ISIS Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot, Oxfordshire, OX11 0QX, United Kingdom
[c] Research Complex at Harwell, Rutherford Appleton Laboratory, Didcot, Oxfordshire, OX11 0FA, United Kingdom
* Corresponding author

Table 2 Scattering-model decoder hyperparameters.

dropout	0.45
num. heads	4
widening factor	3

Table 3 Scattering-model parameter decoder hyperparameters.

dropout	0.05
num. heads	2
widening factor	1

Training Details

We trained for 200 epochs with a batch size of 1834 using the AdamW[2,3] optimizer with a learning rate of 5e-3 and a weight decay of 1e-7. A linear learning rate warm up schedule was applied for 20 epochs, before decaying it to 10% of the original learning rate using a cosine annealing schedule[4] over the remaining epochs.

Evaluation Details

Classification Metrics

Accuracy

The number of correctly classified instances divided by the total number of instances.

Top-3 Accuracy

SASformer outputs a categorical probability distribution over scattering-model classes. If the three largest probabilities in this distribution, which correspond to scattering model classes, contain the ground-truth scattering-model class, we count this instance as correctly classified. Top-3 accuracy is then computed the same as accuracy, where the only difference is the criteria for being correctly classified.

Regression Metrics

Mean Absolute Error

Mean absolute error (MAE) is defined as

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \tag{6}$$

where y_i is the ground-truth value and $\hat{y}_i$ is the prediction for the ith scattering-model parameter. MAE is a useful error metric since it preserves units, meaning the error can be interpreted in the units of the prediction and target values.

Mean Absolute Percentage Error

Mean absolute percentage error (MAPE) is defined as

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i}. \tag{7}$$

It is a unitless quantity and has the advantage that MAPE values can be compared between different scattering-model parameters.

Coefficient of Determination (R^2)

R^2 is defined as

$$R^2 = 1 - \frac{\mathrm{MSE}(y,\hat{y})}{\mathrm{Var}(y)}, \tag{8}$$

where $\mathrm{MSE}(y,\hat{y}) = \frac{1}{N}\sum_{i=1}^{N}(y_- \hat{y}_i)^2$ is the mean-squared error between the prediction and the target, and $\mathrm{Var}(y)$ is the variance of the target variable. It explains the fraction of the variance in the target variable (y) explained by the regression.

k-Nearest Neighbors Classification Baseline Results

We used the k-nearest neighbors (kNN) algorithm as a baseline for the scattering-model classification task. Using the same training and test sets as for SASformer, we used scikit-learn's implementation of kNN with $k = 5$ to predict the model classes. Averaged over all scattering-model classes, accuracy and top-3 accuracy are 0.713 and 0.857, respectively. Full results showing the classification accuracy and top-3 accuracy for each scattering-model are shown in Table 4.

Table 4 Classification accuracy and top-3 accuracy of the kNN baseline on the SAS-55m-20k test set. Both metrics range between 0 and 1. Arrows in column headers indicate whether higher or lower is better.

Scattering Model	Acc ↑	Top 3 Acc ↑	Scattering Model	Acc ↑	Top 3 Acc ↑
Adsorbed layer	0.994	0.998	Lamellar stack paracrys.	0.894	0.967
Binary hard sphere	0.306	0.427	Linear pearls	0.687	0.972
Broad peak	0.971	0.992	Lorentz	1.0	1.0
Core multi shell	0.135	0.254	Mass fractal	0.999	1.0
Core shell bicelle	0.272	0.42	Mass surface fractal	1.0	1.0
Core shell cylinder	0.262	0.46	Mono Gauss coil	0.933	0.992
Core shell ellipsoid	0.473	0.632	Multilayer vesicle	0.878	0.943
Core shell sphere	0.279	0.485	Onion	0.22	0.558
Correlation length	0.999	1.0	Parallelepiped	0.581	0.81
Cylinder	0.284	0.544	Peak Lorentz	0.892	0.969
Dab	1.0	1.0	Pearl necklace	0.815	0.987
Ellipsoid	0.444	0.68	Poly Gauss coil	0.89	0.978
Elliptical cylinder	0.377	0.668	Polymer excl volume	0.96	0.99
Flexible cylinder	0.334	0.6	Porod	1.0	1.0
Flexible cylinder elliptical	0.992	0.995	Power law	1.0	1.0
Fractal	0.823	0.95	Raspberry	0.898	0.946
Fractal core shell	0.659	0.819	Rectangular prism	0.428	0.778
Fuzzy sphere	0.462	0.787	Sphere	0.389	0.814
Gauss Lorentz gel	0.818	0.952	Stacked disks	0.764	0.923
Gaussian peak	0.977	0.988	Star polymer	0.994	1.0
Gel fit	0.787	0.94	Surface fractal	0.965	0.991
Guinier	1.0	1.0	Teubner strey	0.894	0.972
Hollow cylinder	0.334	0.679	Triaxial ellipsoid	0.497	0.806
Hollow rect. prism thin walls	0.826	0.944	Two Lorentzian	0.958	0.984
Lamellar	0.926	0.994	Two power law	0.899	0.953
Lamellar hg	0.774	0.9	Unified power rg	0.932	0.992
Lamellar hg stack caille	0.513	0.826	Vesicle	0.175	0.994
Lamellar stack caille	0.628	0.874	-	-	-

Scattering-Model Parameter Regression Results

Table 5 Scattering-model parameter regression results part 1 of 4. MAE is the mean absolute error; AE_{σ} is the standard deviation of the absolute errors; AE_{IQR} is the inter-quartile range of the absolute errors; MAPE is the mean absolute percentage error, reported in fractional form; and R^2 is the coefficient of determination. Arrows in column headers indicate whether higher or lower is better.

Scattering Model	Parameter	MAE ↓	AE_{σ}	AE_{IQR}	MAPE ↓	R^2 ↑
Adsorbed layer	Adsorbed amount	0.862	0.554	0.822	2.808	0.255
	Density shell	0.251	0.148	0.249	0.290	0.044
	Radius	169.103	205.555	165.773	1.632	-0.037
	Second moment	0.986	1.026	0.976	0.035	0.997
	Volfraction	0.062	0.036	0.058	14.358	0.067
Binary hard sphere	Radius lg	0.638	0.815	0.605	0.176	0.846
	Radius sm	26.520	58.706	22.264	0.116	0.940
	Volfraction lg	0.147	0.156	0.152	1.840	0.869
	Volfraction sm	0.032	0.030	0.034	0.899	0.644
Broad peak	Lorentz exp	0.052	0.046	0.051	0.028	0.986
	Lorentz length	4.011	4.360	3.970	0.071	0.938
	Lorentz scale	12.433	10.842	13.880	0.342	0.652
	Peak pos	0.016	0.026	0.016	0.082	0.986
	Porod exp	0.032	0.026	0.033	0.011	0.995
	Porod scale	0.000	0.000	0.000	0.374	0.703
Core multi shell	N	1.310	1.052	1.412	0.308	0.593
	Radius	145.896	190.091	162.747	1.450	0.114
	Thickness1	177.719	206.226	198.250	2.148	-0.157
	Thickness2	186.422	214.861	214.198	2.354	-0.214
	Thickness3	187.377	211.698	209.021	2.438	-0.259
	Thickness4	190.852	223.132	216.725	2.229	-0.277
	Thickness5	195.579	217.730	222.531	2.506	-0.261
	Thickness6	190.307	217.657	213.596	2.339	-0.300
	Thickness7	197.992	210.381	214.292	2.761	-0.239
	Thickness8	191.228	208.704	210.022	2.450	-0.255
	Thickness9	195.301	220.654	223.486	2.243	-0.317
	Thickness10	200.640	221.206	221.734	2.390	-0.291
Core shell bicelle	Length	69.649	100.546	76.930	0.568	0.766
	Radius	38.753	59.264	42.805	0.257	0.921
	Thick face	16.333	16.300	19.172	0.506	0.168
	Thick rim	11.695	13.348	13.586	0.361	0.516
Core shell cylinder	Length	413.824	898.812	354.404	0.448	0.833
	Radius	307.195	679.054	253.536	0.277	0.906
	Thickness	13.004	13.956	13.174	0.204	0.886
Core shell ellipsoid	Radius equat core	31.362	54.297	30.452	0.172	0.947
	Thick shell	4.690	8.358	4.320	0.312	0.860
	X core	0.716	0.784	0.703	0.396	0.869
	X polar shell	1.525	1.331	1.693	0.902	0.517
Core shell sphere	Radius	27.392	47.397	27.654	0.150	0.956
	Thickness	4.274	7.868	3.910	0.256	0.883
Correlation length	Cor length	1.207	1.051	1.243	0.024	0.995
	Lorentz exp	0.015	0.012	0.016	0.006	0.997
	Lorentz scale	4.452	5.067	4.326	0.087	0.936
	Porod exp	0.019	0.016	0.019	0.008	0.993
	Porod scale	0.000	0.000	0.000	0.123	0.851
Cylinder	Length	1698.570	4163.153	1084.535	0.191	0.960
	Radius	21.786	37.560	22.131	0.097	0.972
Dab	Cor length	8.309	15.362	7.192	0.027	0.996
Ellipsoid	Radius equatorial	20.264	33.426	21.469	0.092	0.979
	Radius polar	27.992	49.168	28.525	0.127	0.952
Elliptical cylinder	Axis ratio	52.091	110.994	41.087	0.349	0.748
	Length	2302.328	6575.117	1393.437	0.404	0.894
	Radius minor	35.539	67.944	33.574	0.183	0.911
Flexible cylinder	Kuhn length	38.042	66.663	38.915	0.187	0.909
	Length	2511.259	6269.678	1521.593	0.263	0.906
	Radius	20.779	37.871	20.353	0.085	0.972
Flexible cylinder elliptical	Axis ratio	0.214	0.226	0.216	0.052	0.986
	Kuhn length	24.236	38.025	25.938	0.136	0.969
	Length	1581.836	4314.192	830.649	0.157	0.952

Table 6 Scattering-model parameter regression results part 2 of 4. MAE is the mean absolute error; AE_{σ} is the standard deviation of the absolute errors; AE_{IQR} is the inter-quartile range of the absolute errors; MAPE is the mean absolute percentage error, reported in fractional form; and R^2 is the coefficient of determination. Arrows in column headers indicate whether higher or lower is better.

Scattering Model	Parameter	MAE ↓	AE_{σ}	AE_{IQR}	MAPE ↓	R^2 ↑
Fractal core shell	Cor length	2.280	1.485	2.323	0.688	-0.070
	Fractal dim	0.399	0.243	0.388	0.186	-0.094
	Radius	3.500	3.658	3.454	0.188	0.859
	Thickness	2.973	3.196	2.939	0.151	0.892
	Volfraction	0.021	0.014	0.020	0.582	0.145
Fractal	Cor length	1.253	1.110	1.337	0.381	0.605
	Fractal dim	0.341	0.227	0.364	0.158	0.143
	Radius	0.935	0.798	0.959	0.048	0.992
	Volfraction	0.022	0.013	0.021	0.677	0.073
Fuzzy sphere	Fuzziness	0.058	0.066	0.057	0.079	0.971
	Radius	13.248	23.446	13.381	0.060	0.989
Gauss Lorentz gel	Cor length dynamic	21.497	40.551	20.276	0.080	0.969
	Cor length static	27.348	55.642	26.081	0.119	0.941
	Gauss scale	462.762	894.316	467.135	0.425	0.821
	Lorentz scale	528.680	1080.449	454.921	0.358	0.736
Gaussian peak	Peak pos	0.011	0.013	0.011	0.178	0.997
	Sigma	0.002	0.001	0.002	0.184	0.995
Gel fit	Cor length	25.259	45.994	22.563	0.097	0.960
	Fractal dim	0.070	0.064	0.071	0.031	0.989
	Guinier scale	7.748	8.943	8.489	0.213	0.787
	Lorentz scale	8.106	9.770	8.737	0.208	0.756
	Rg	22.843	44.256	21.524	0.089	0.964
Guinier	Rg	11.136	18.799	11.377	0.035	0.994
Hollow cylinder	Length	272.465	616.818	211.554	0.196	0.922
	Radius	450.768	942.475	302.335	0.230	0.817
	Thickness	9.548	10.913	9.419	0.163	0.933
Hollow rect. prism thin walls	B2a ratio	0.183	0.138	0.190	1.169	0.382
	C2a ratio	0.184	0.139	0.179	1.091	0.363
	Length a	23.331	38.720	24.063	0.111	0.969
Lamellar	Thickness	1.216	1.462	1.258	0.030	0.995
Lamellar hg	Length head	4.644	5.356	4.481	0.141	0.937
	Length tail	5.849	6.059	5.863	0.179	0.909
Lamellar hg stack caille	Caille parameter	0.048	0.044	0.050	0.557	0.806
	D spacing	0.309	0.326	0.296	0.043	0.903
	Length head	10.278	10.871	10.150	0.350	0.714
	Length tail	10.139	10.143	10.041	0.356	0.737
	Nlayers	39.664	37.169	40.098	0.154	0.965
Lamellar stack caille	Caille parameter	0.059	0.057	0.064	0.987	0.682
	D spacing	0.460	0.518	0.443	0.122	0.930
	Nlayers	59.351	64.066	59.459	0.204	0.909
	Thickness	3.345	3.648	3.240	0.105	0.969
Lamellar stack paracrys.	D spacing	0.598	0.577	0.640	0.148	0.901
	Nlayers	68.483	72.548	70.012	0.239	0.884
	Sigma d	0.048	0.049	0.049	0.707	0.947
	Thickness	2.972	3.546	2.921	0.091	0.973
Linear pearls	Edge sep	1.441	1.524	1.472	0.499	0.917
	Num pearls	0.307	0.235	0.313	0.051	0.966
	Radius	11.784	20.275	11.836	0.051	0.992
Lorentz	Cor length	11.196	20.121	10.597	0.034	0.993
Mass fractal	Cutoff length	0.331	0.310	0.341	0.064	0.964
	Fractal dim mass	0.080	0.071	0.079	0.036	0.923
	Radius	1.009	1.190	1.073	0.024	0.996
Mass surface fractal	Fractal dim mass	0.048	0.046	0.047	0.022	0.970
	Fractal dim surf	0.038	0.033	0.039	0.015	0.972
	Rg cluster	1.811	2.244	1.815	0.045	0.988
	Rg primary	203.827	434.901	191.485	0.078	0.966
Mono Gauss coil	I zero	373.734	1003.459	120.201	0.279	0.758
	Rg	8.919	15.069	8.770	0.039	0.996
Multilayer vesicle	N shells	0.814	0.763	0.819	0.111	0.976
	Radius	48.477	68.633	49.300	0.451	0.894
	Thick shell	3.308	4.083	3.401	0.088	0.956
	Thick solvent	3.363	4.651	3.358	0.094	0.948

Table 7 Scattering-model parameter regression results part 3 of 4. MAE is the mean absolute error; AE_{σ} is the standard deviation of the absolute errors; AE_{IQR} is the inter-quartile range of the absolute errors; MAPE is the mean absolute percentage error, reported in fractional form; and R^2 is the coefficient of determination. Arrows in column headers indicate whether higher or lower is better.

Scattering Model	Parameter	MAE ↓	AE_{σ}	AE_{IQR}	MAPE ↓	R^2 ↑
Onion	A1	2.397	1.782	2.541	2.714	-0.055
	A2	2.751	1.826	2.709	2.378	-0.299
	A3	2.764	1.777	2.775	2.235	-0.271
	A4	2.734	1.785	2.755	6.144	-0.250
	A5	2.689	1.760	2.679	2.844	-0.248
	A6	2.686	1.717	2.631	3.223	-0.211
	A7	2.728	1.730	2.680	4.075	-0.229
	A8	2.733	1.763	2.714	2.757	-0.241
	A9	2.730	1.747	2.704	3.473	-0.243
	A10	2.738	1.762	2.735	6.417	-0.226
	Radius core	42.121	68.370	44.345	0.205	0.904
	Thickness1	5.121	4.683	5.194	0.157	0.939
	Thickness2	26.134	16.770	25.898	1.027	-0.236
	Thickness3	26.088	16.983	25.996	1.025	-0.252
	Thickness4	26.218	17.096	26.058	0.984	-0.280
	Thickness5	26.216	16.845	25.815	1.038	-0.251
	Thickness6	25.774	16.787	25.698	1.000	-0.241
	Thickness7	26.246	16.998	26.143	0.978	-0.246
	Thickness8	26.638	17.695	27.026	1.106	-0.314
	Thickness9	26.455	16.989	26.541	0.989	-0.267
	Thickness10	26.450	17.005	26.235	1.003	-0.278
Parallelepiped	Length a	137.189	170.439	158.590	1.467	0.280
	Length b	146.516	182.981	161.079	1.136	0.263
	Length c	58.388	112.911	51.739	0.362	0.770
Peak Lorentz	Peak hwhm	0.165	0.453	0.089	0.109	0.957
	Peak pos	0.013	0.026	0.011	0.094	0.985
Pearl necklace	Edge sep	1.146	1.393	1.062	0.938	0.941
	Num pearls	0.332	0.261	0.345	0.055	0.958
	Radius	15.205	29.026	14.651	0.059	0.984
	Thick string	0.258	0.146	0.250	1.904	-0.026
Poly Gauss coil	I zero	686.649	1380.500	608.009	0.606	0.619
	Rg	5.217	8.087	5.741	0.036	0.997
Polymer excl volume	Porod exp	0.029	0.023	0.029	0.014	0.998
	Rg	12.619	23.078	11.683	0.048	0.990
Power law	Power	0.030	0.024	0.032	0.013	0.999
Raspberry	Penetration	0.250	0.159	0.259	3.335	-0.001
	Radius lg	8292.868	13153.885	10607.428	0.305	0.643
	Radius sm	9.858	18.550	8.633	0.117	0.938
	Surface fraction	0.137	0.105	0.143	1.103	0.445
	Volfraction lg	0.053	0.037	0.054	2.491	0.235
	Volfraction sm	0.022	0.015	0.021	2.495	0.205
Rectangular prism	B2a ratio	0.178	0.132	0.186	1.214	0.435
	C2a ratio	0.176	0.131	0.186	0.986	0.426
	Length a	27.168	47.821	26.042	0.121	0.954
Sphere	Radius	50.658	112.679	35.561	0.073	0.977
Stacked disks	N stacking	6.853	6.508	7.067	0.221	0.892
	Radius	1.761	1.570	1.832	0.076	0.968
	Sigma d	0.032	0.027	0.033	3.802	0.676
	Thick core	7.524	6.438	7.941	0.388	0.443
	Thick layer	6.000	5.547	6.358	0.321	0.615

Table 8 Scattering-model parameter regression results part 4 of 4. MAE is the mean absolute error; AE_{σ} is the standard deviation of the absolute errors; AE_{IQR} is the inter-quartile range of the absolute errors; MAPE is the mean absolute percentage error, reported in fractional form; and R^2 is the coefficient of determination. Arrows in column headers indicate whether higher or lower is better.

Scattering Model	Parameter	MAE ↓	AE_{σ}	AE_{IQR}	MAPE ↓	R^2 ↑
Star polymer	Arms	0.145	0.131	0.152	0.047	0.995
	Rg squared	11.879	21.198	11.261	0.048	0.992
Surface fractal	Cutoff length	3.899	6.085	3.105	0.143	0.933
	Fractal dim surf	0.048	0.047	0.048	0.026	0.987
	Radius	1.255	2.126	0.994	0.060	0.965
Teubner strey	D	5.049	11.467	4.768	0.048	0.977
	Xi	2.286	2.941	2.088	0.078	0.983
Triaxial ellipsoid	Radius equat major	855.134	1705.382	780.210	2.302	0.332
	Radius equat minor	425.736	885.682	358.434	0.522	0.811
	Radius polar	801.364	1462.445	883.625	1.080	0.493
Two Lorentzian	Lorentz exp 1	0.035	0.027	0.036	0.018	0.994
	Lorentz exp 2	0.330	0.286	0.418	0.183	0.434
	Lorentz length 1	18.639	22.285	18.624	0.046	0.988
	Lorentz length 2	11.854	15.996	15.152	1.004	0.366
	Lorentz scale 1	970.034	1514.220	1076.143	0.691	0.511
	Lorentz scale 2	0.022	0.014	0.022	0.658	0.074
Two power law	Coefficent 1	1454.772	2246.504	1199.095	5.652	-0.238
	Crossover	0.020	0.027	0.018	0.183	0.947
	Power 1	0.044	0.037	0.045	0.021	0.996
	Power 2	0.065	0.053	0.068	0.031	0.991
Unified power rg	B1	0.000	0.000	0.000	1.245	0.412
	B2	0.000	0.000	0.000	5.019	-0.241
	B3	0.000	0.000	0.000	5.839	-0.208
	B4	0.000	0.000	0.000	5.152	-0.239
	B5	0.000	0.000	0.000	5.105	-0.248
	B6	0.000	0.000	0.000	4.961	-0.245
	G1	9034.578	15833.418	8468.681	2.547	0.301
	G2	0.248	0.156	0.252	1.694	-0.067
	G3	0.252	0.156	0.248	1.888	-0.090
	G4	0.254	0.154	0.249	1.871	-0.075
	G5	0.254	0.154	0.245	1.888	-0.110
	G6	0.255	0.157	0.256	1.690	-0.100
	Level	0.897	0.667	1.008	0.297	0.426
	Power1	0.191	0.197	0.186	0.094	0.903
	Power2	0.772	0.471	0.777	0.383	-0.085
	Power3	0.799	0.479	0.763	0.396	-0.099
	Power4	0.786	0.477	0.769	0.390	-0.098
	Power5	0.780	0.482	0.799	0.394	-0.098
	Power6	0.791	0.488	0.791	0.399	-0.101
	Rg1	5237.482	10346.564	4683.649	0.312	0.743
	Rg2	0.182	0.152	0.217	1.246	0.313
	Rg3	0.229	0.153	0.230	1.486	0.072
	Rg4	0.249	0.155	0.239	1.531	-0.050
	Rg5	0.256	0.156	0.254	1.843	-0.081
	Rg6	0.258	0.153	0.247	1.759	-0.070
Vesicle	Radius	6381.184	10996.375	7044.523	0.236	0.767
	Thickness	1.882	2.301	1.943	0.048	0.987
	Volfraction	0.229	0.144	0.224	16.524	0.163

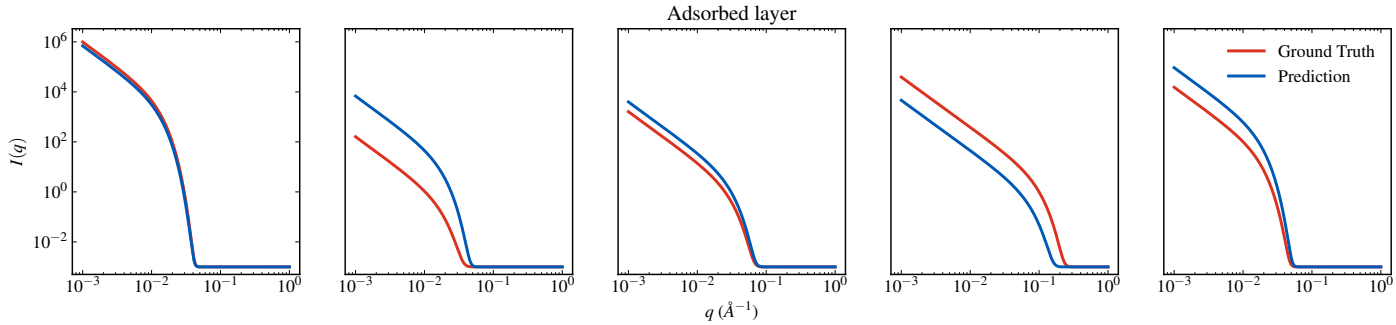

Fig. 1 Comparing predicted and ground-truth adsorbed layer models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

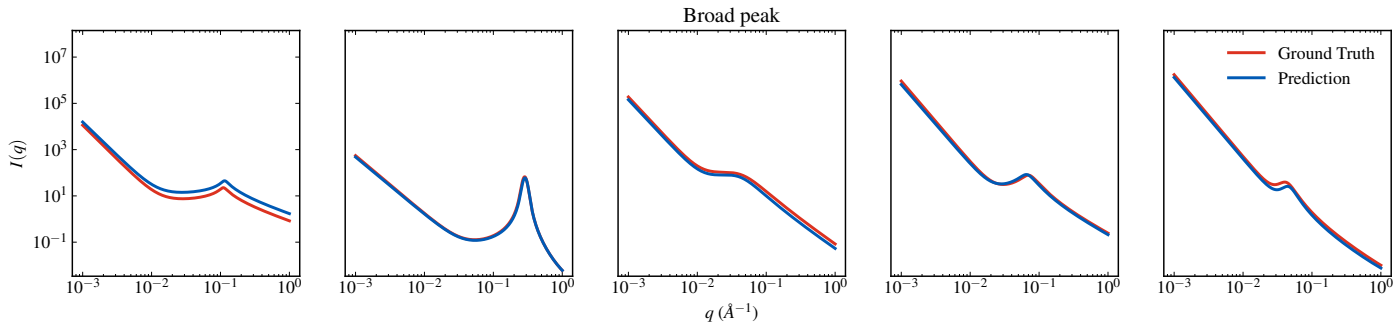

Fig. 2 Comparing predicted and ground-truth broad peak models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

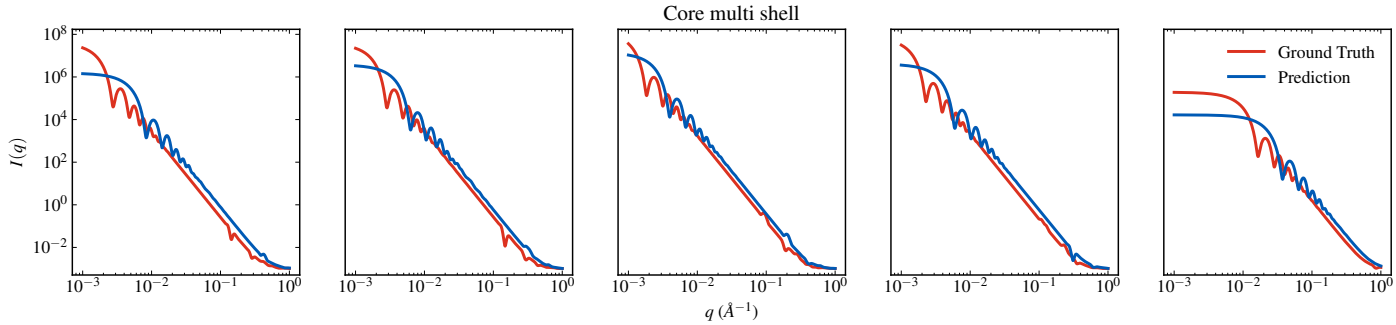

Fig. 3 Comparing predicted and ground-truth core multi shell models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

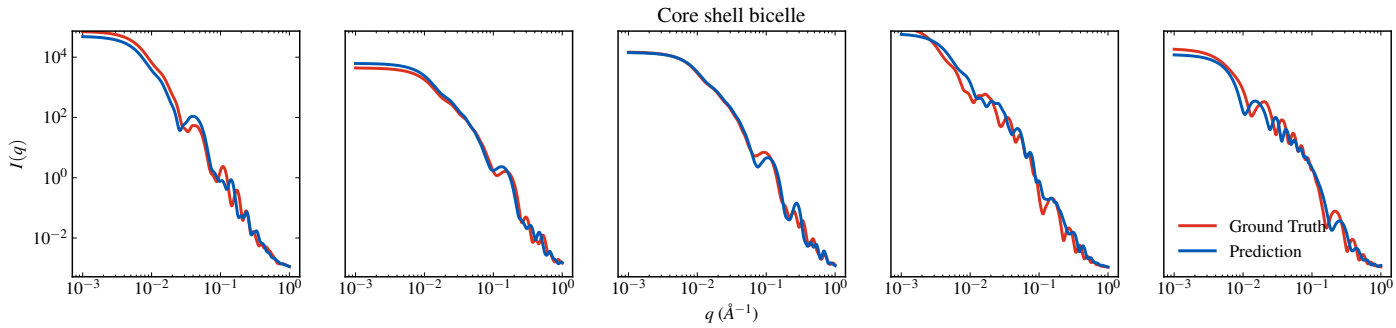

Fig. 4 Comparing predicted and ground-truth core shell bicelle models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

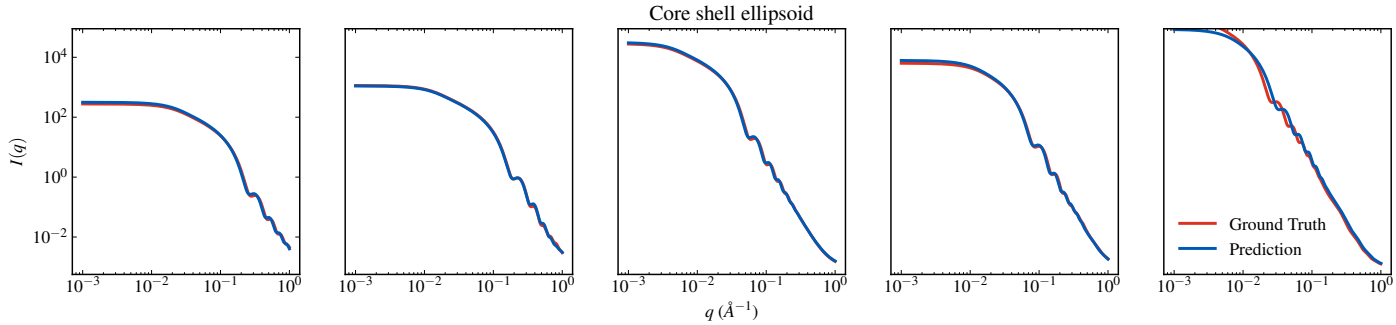

Fig. 5 Comparing predicted and ground-truth core shell ellipsoid models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

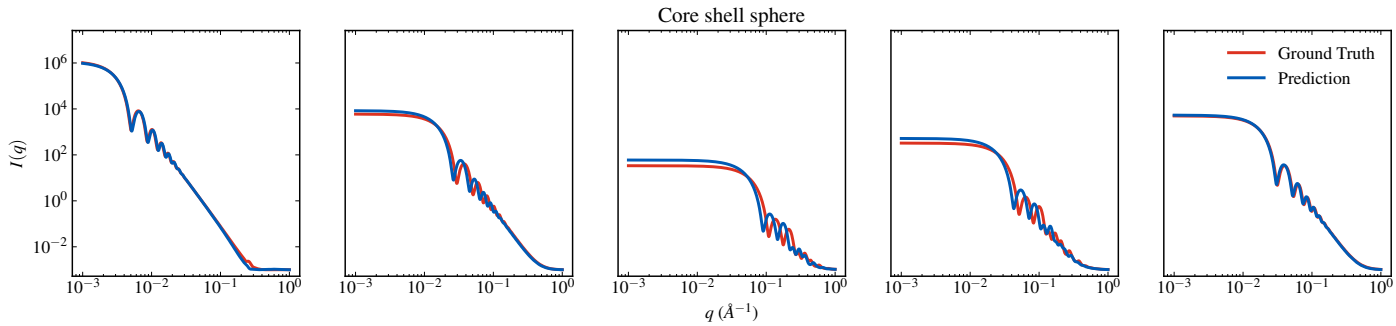

Fig. 6 Comparing predicted and ground-truth core shell sphere models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

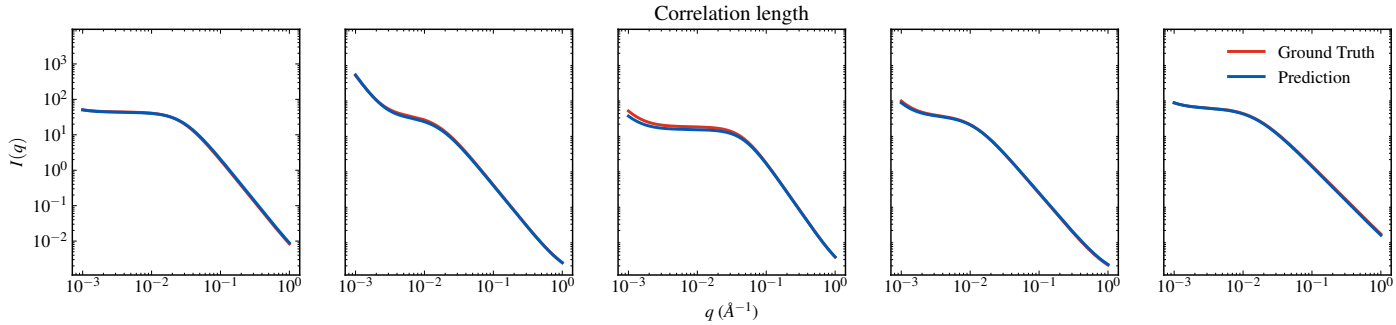

Fig. 7 Comparing predicted and ground-truth correlation length models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

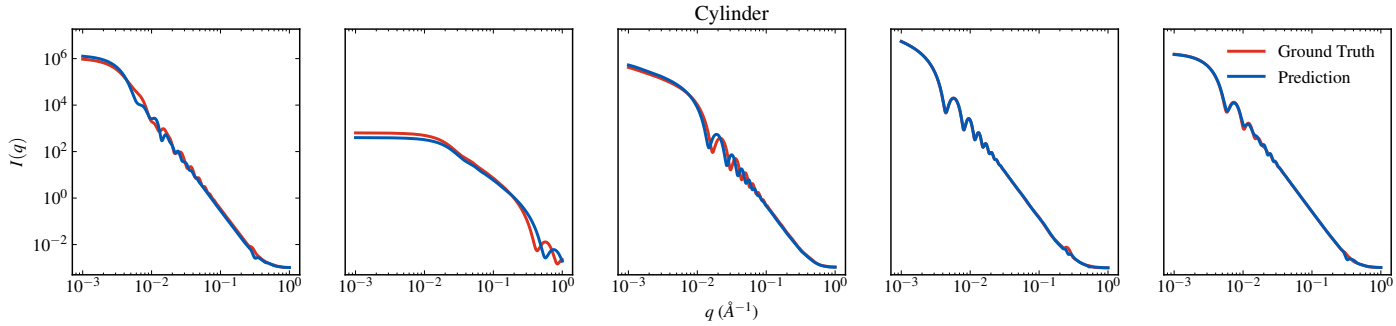

Fig. 8 Comparing predicted and ground-truth cylinder models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

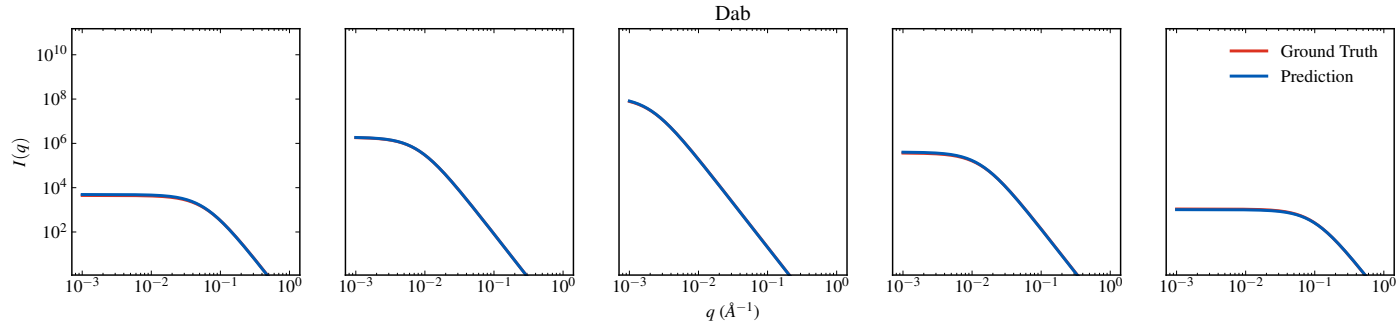

Fig. 9 Comparing predicted and ground-truth dab models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

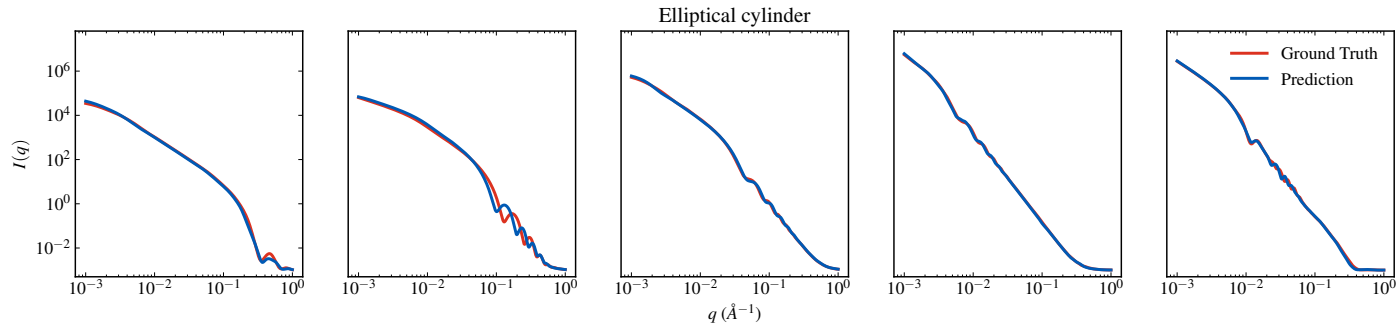

Fig. 10 Comparing predicted and ground-truth elliptical cylinder models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

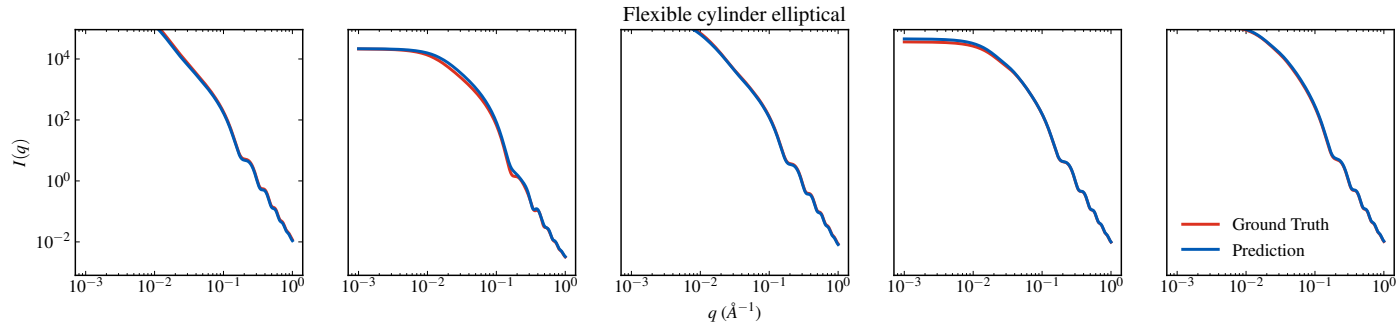

Fig. 11 Comparing predicted and ground-truth flexible cylinder elliptical models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

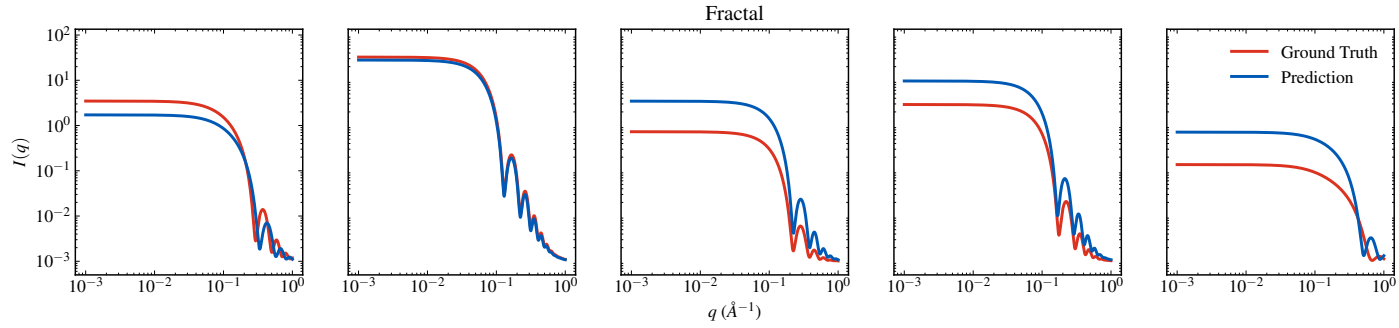

Fig. 12 Comparing predicted and ground-truth fractal models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

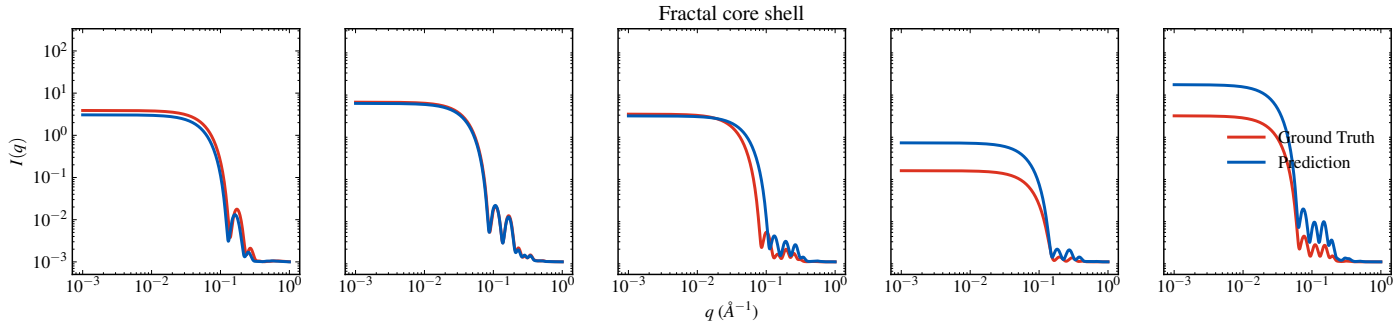

Fig. 13 Comparing predicted and ground-truth fractal core shell models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

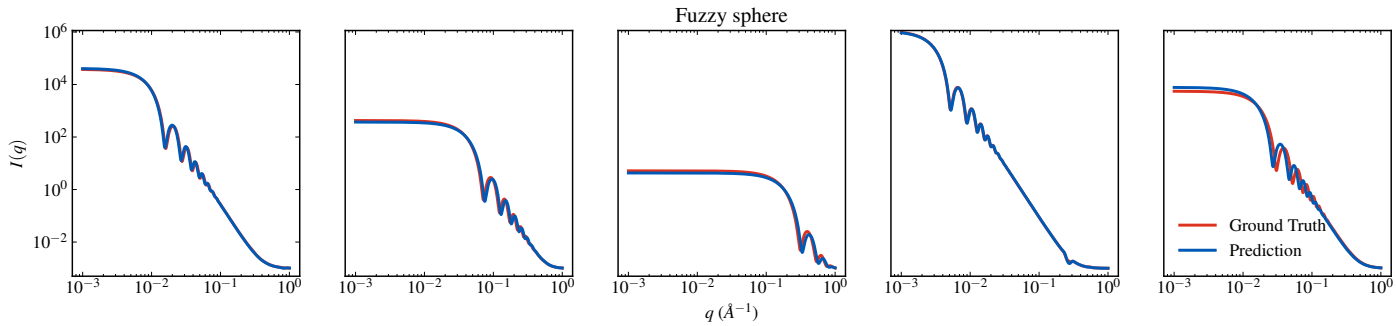

Fig. 14 Comparing predicted and ground-truth fuzzy sphere models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

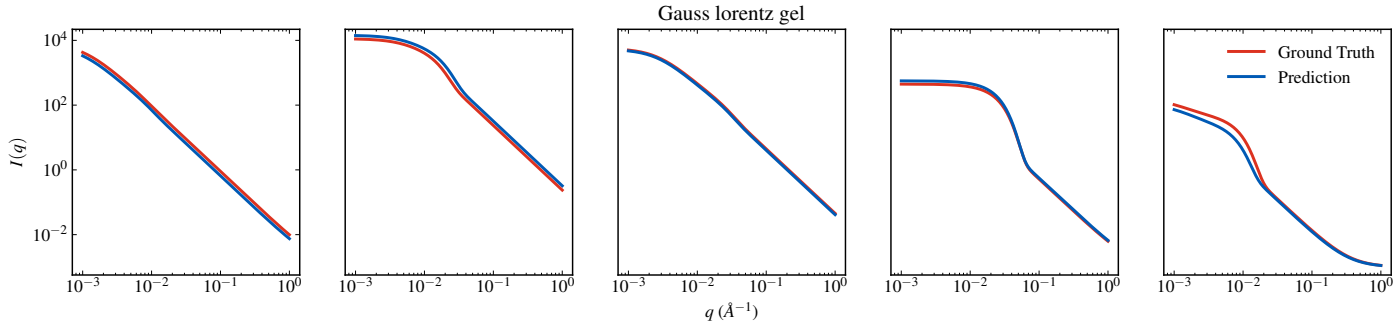

Fig. 15 Comparing predicted and ground-truth gauss lorentz gel models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

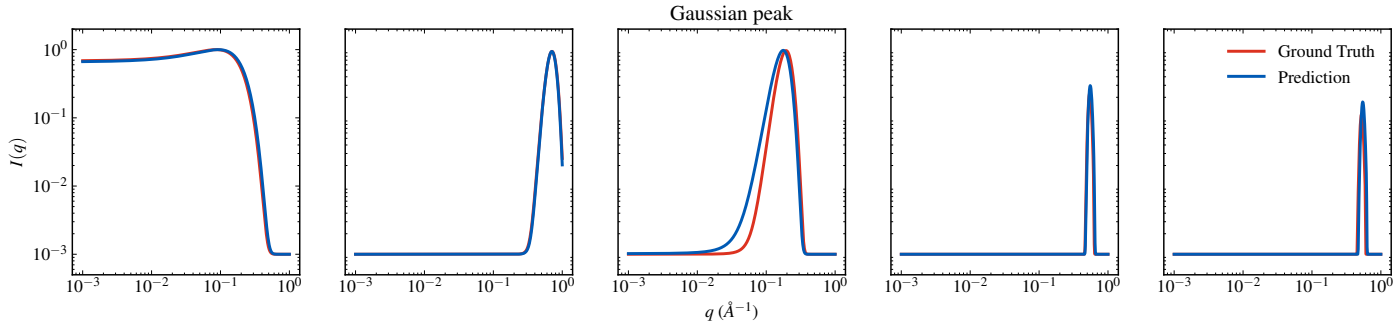

Fig. 16 Comparing predicted and ground-truth gaussian peak models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

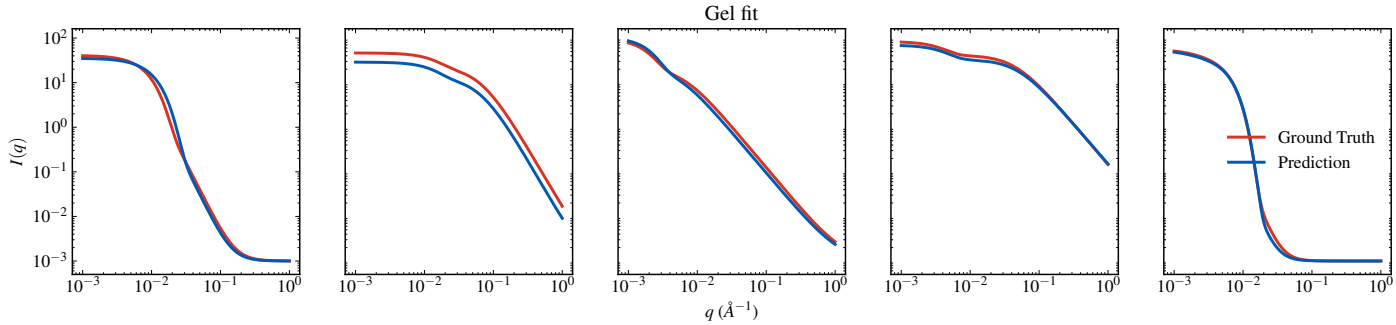

Fig. 17 Comparing predicted and ground-truth gel fit models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

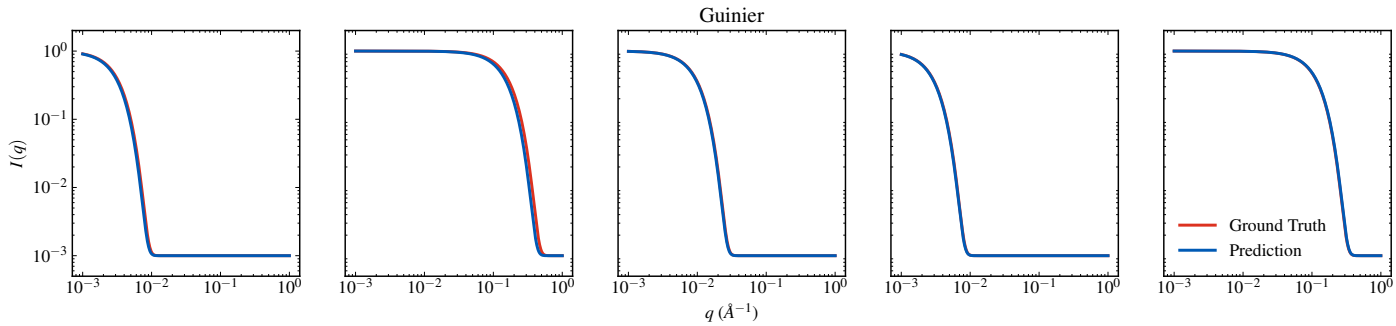

Fig. 18 Comparing predicted and ground-truth guinier models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

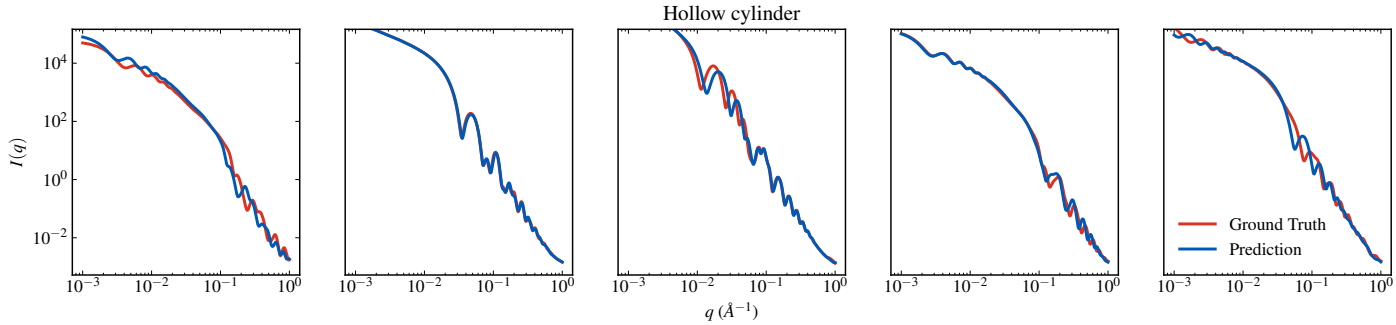

Fig. 19 Comparing predicted and ground-truth hollow cylinder models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

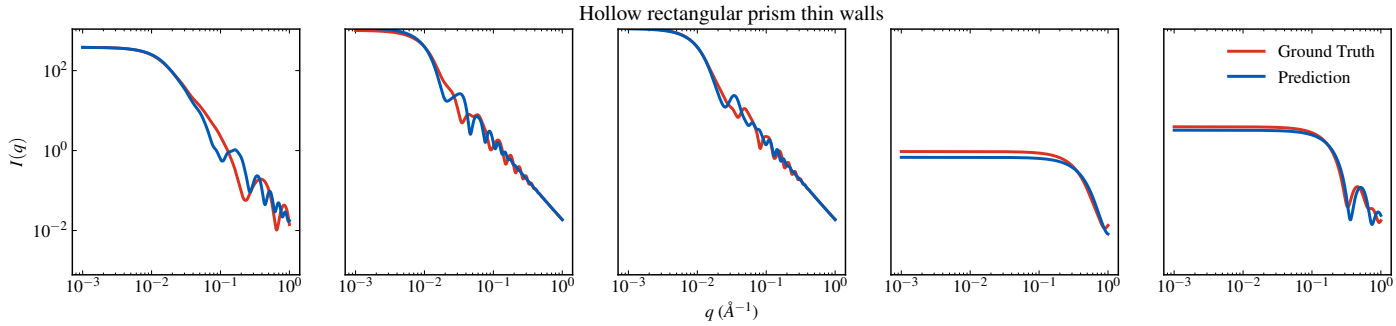

Fig. 20 Comparing predicted and ground-truth hollow rectangular prism thin walls models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

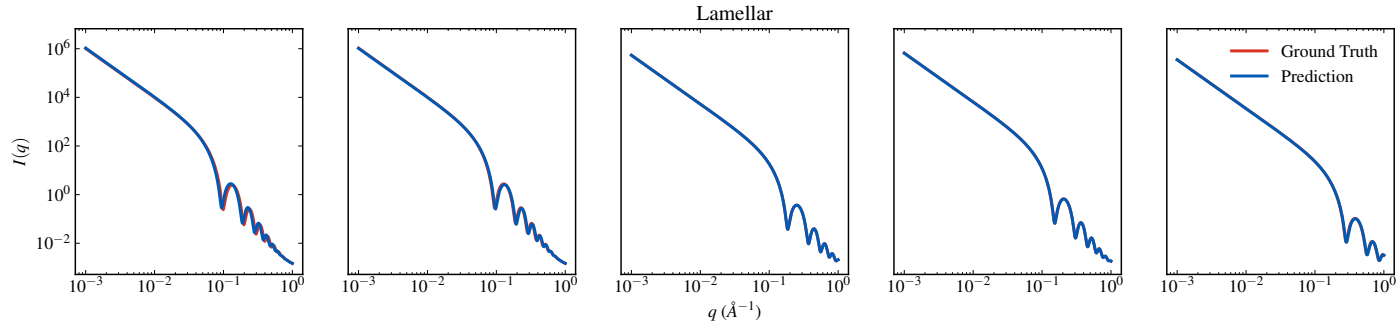

Fig. 21 Comparing predicted and ground-truth lamellar models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

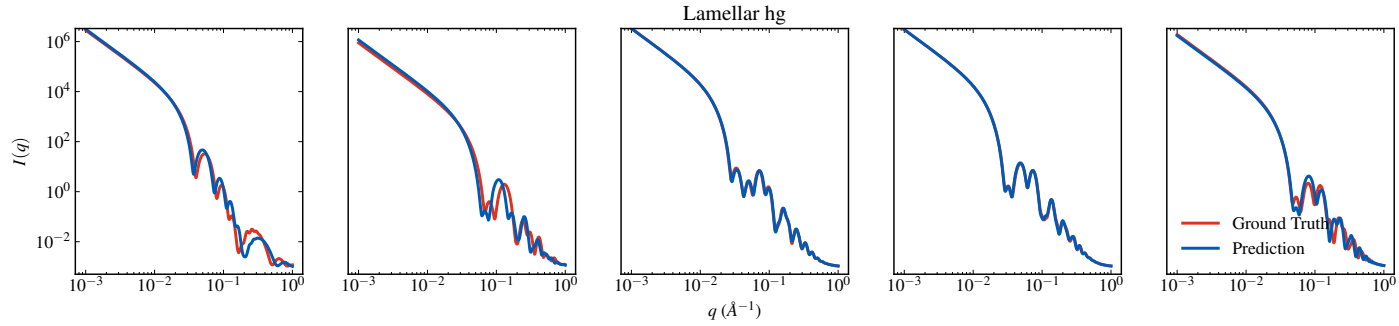

Fig. 22 Comparing predicted and ground-truth lamellar hg models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

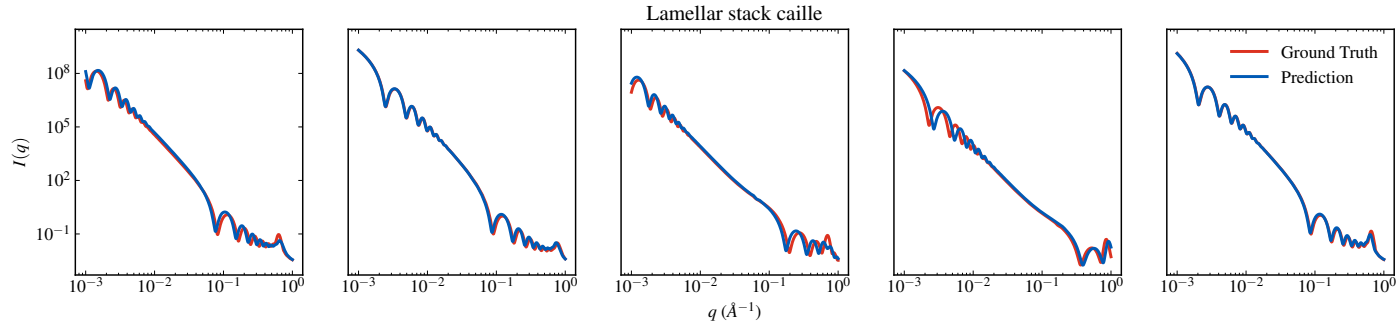

Fig. 23 Comparing predicted and ground-truth lamellar stack caille models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

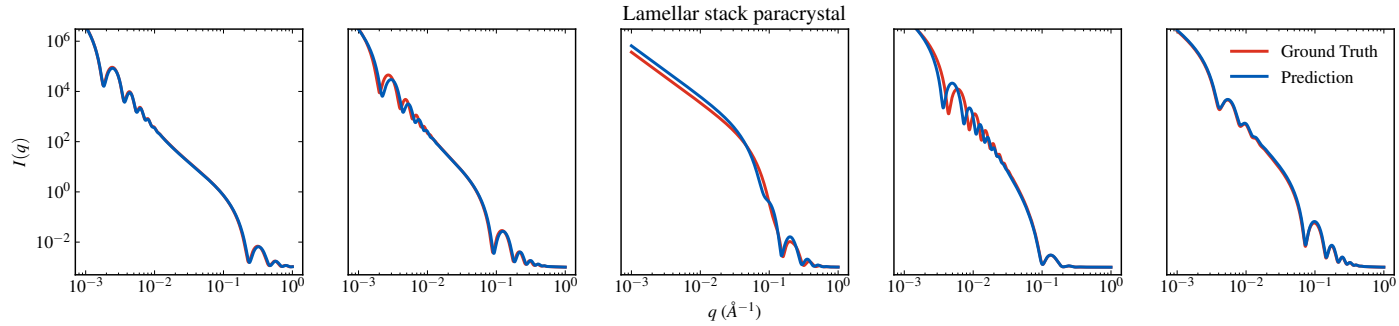

Fig. 24 Comparing predicted and ground-truth lamellar stack paracrystal models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

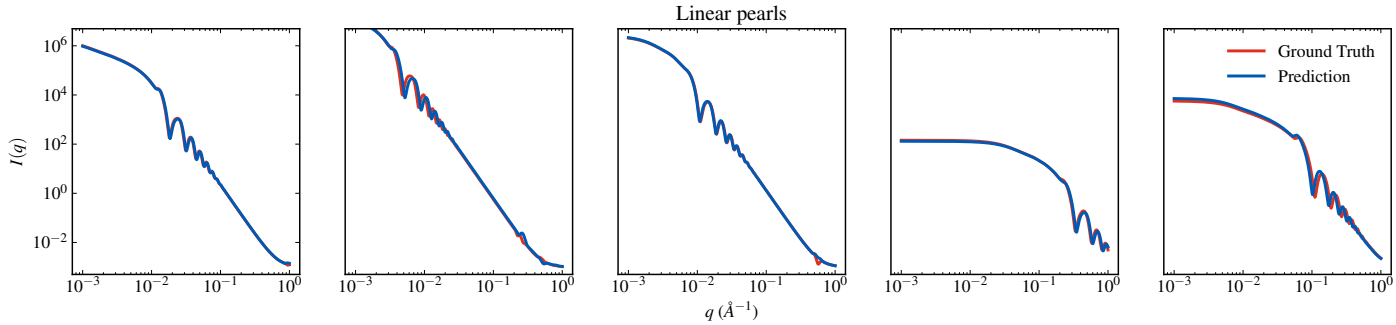

Fig. 25 Comparing predicted and ground-truth linear pearls models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

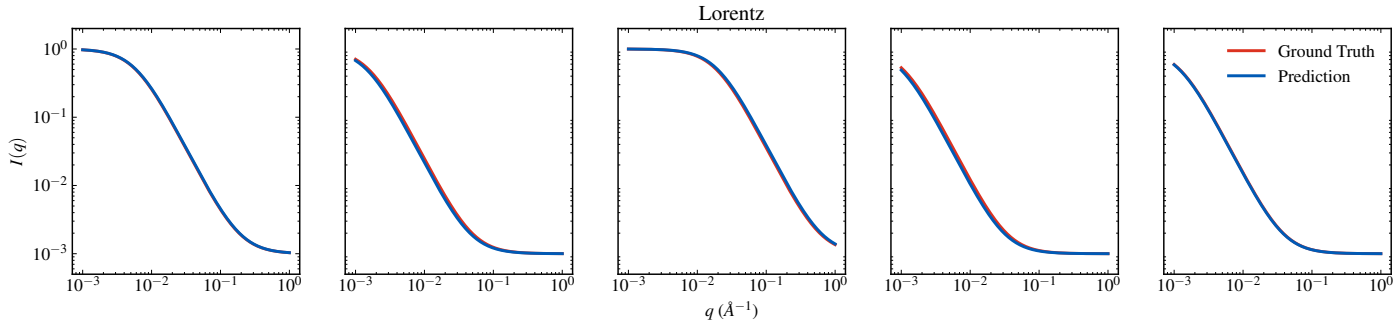

Fig. 26 Comparing predicted and ground-truth lorentz models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

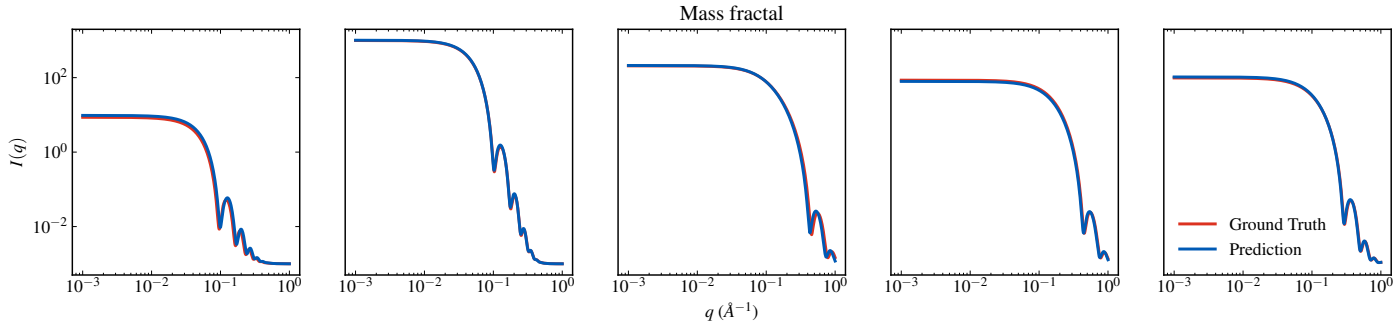

Fig. 27 Comparing predicted and ground-truth mass fractal models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

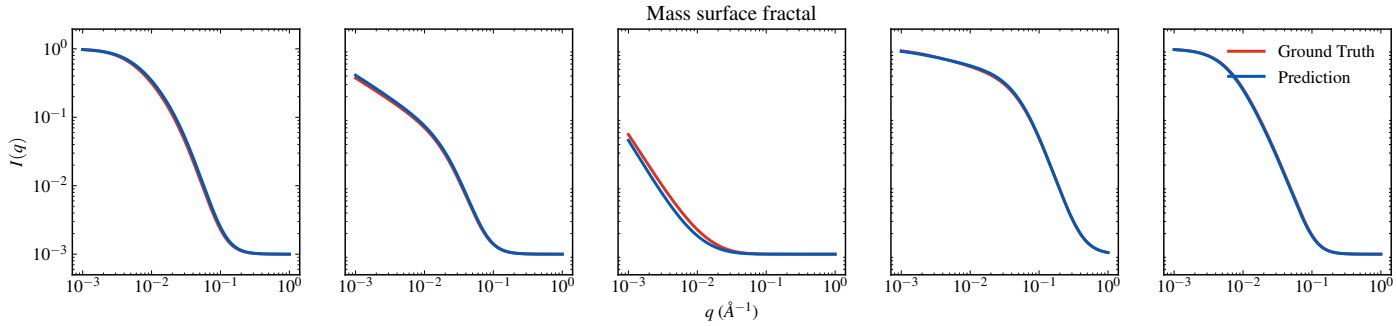

Fig. 28 Comparing predicted and ground-truth mass surface fractal models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

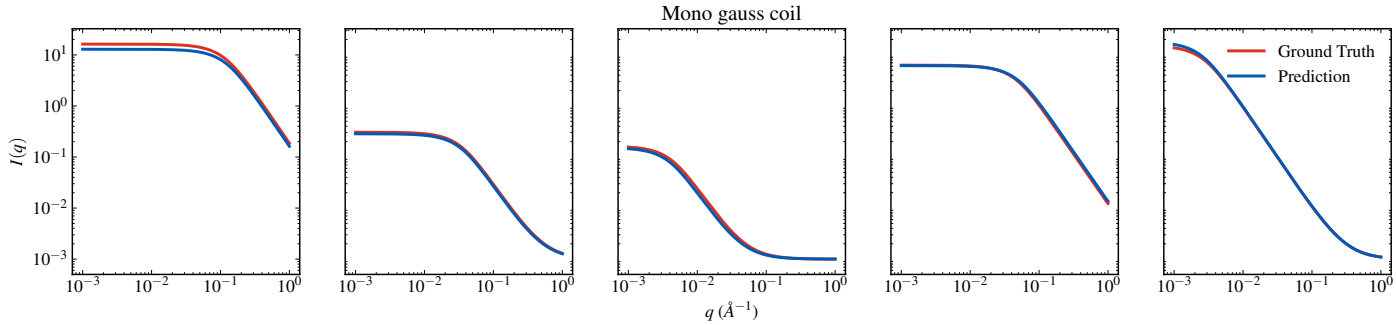

Fig. 29 Comparing predicted and ground-truth mono gauss coil models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

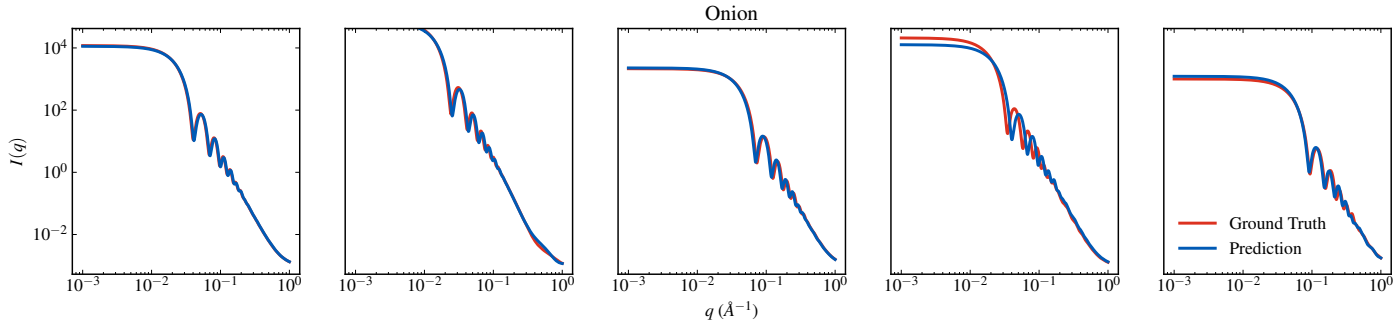

Fig. 30 Comparing predicted and ground-truth onion models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

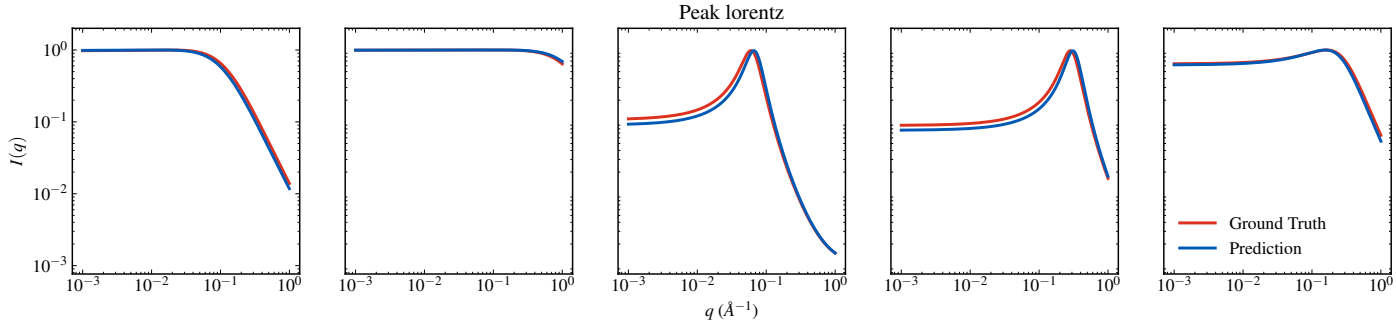

Fig. 31 Comparing predicted and ground-truth peak lorentz models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

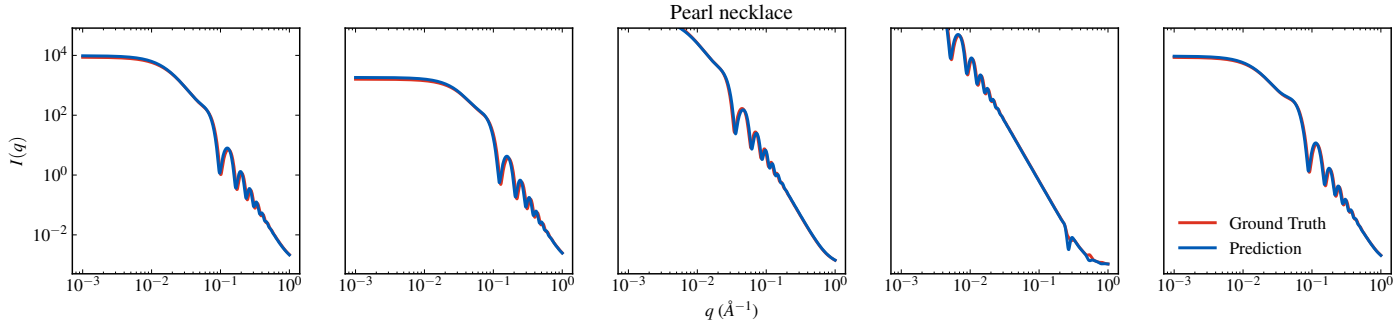

Fig. 32 Comparing predicted and ground-truth pearl necklace models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

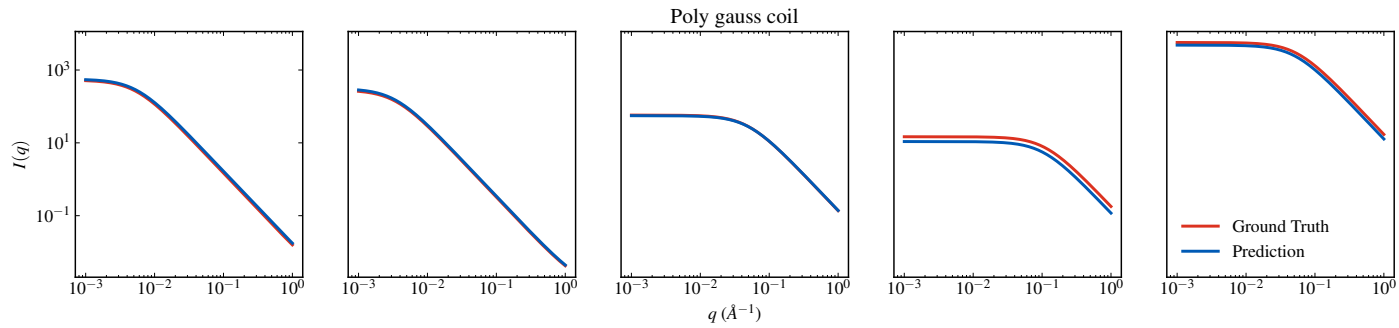

Fig. 33 Comparing predicted and ground-truth poly gauss coil models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

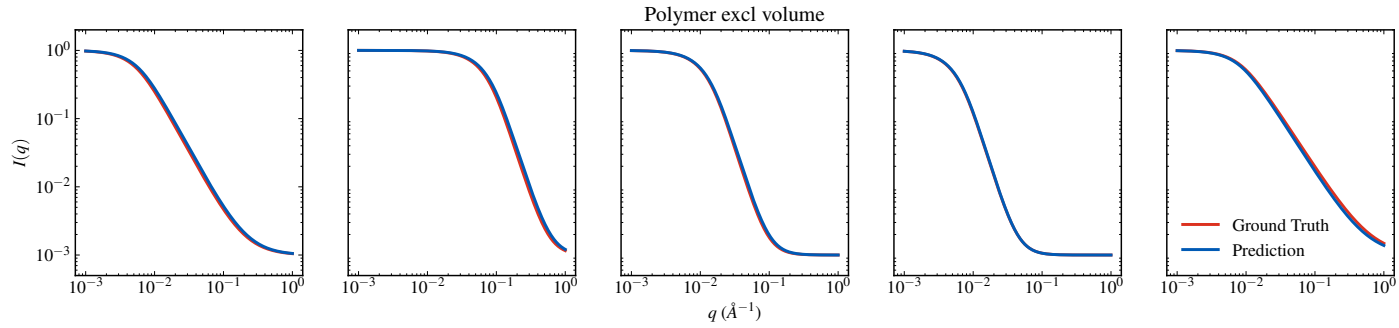

Fig. 34 Comparing predicted and ground-truth polymer excl volume models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

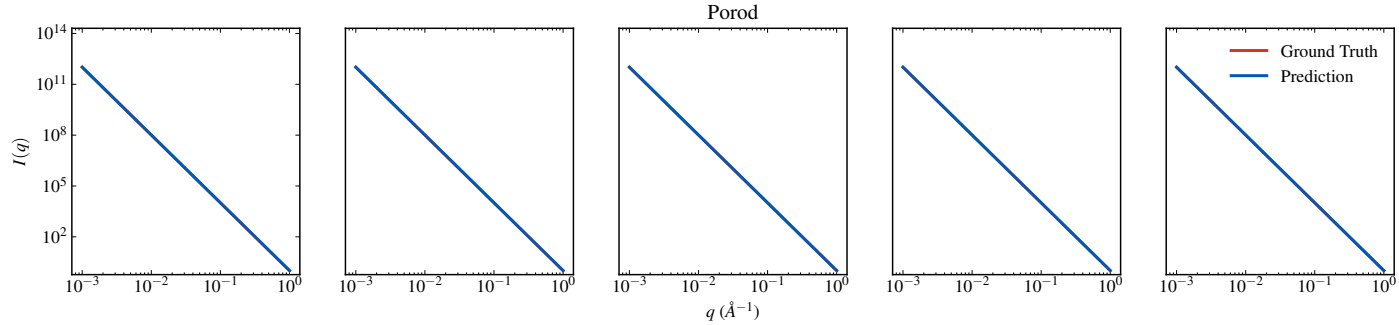

Fig. 35 Comparing predicted and ground-truth porod models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

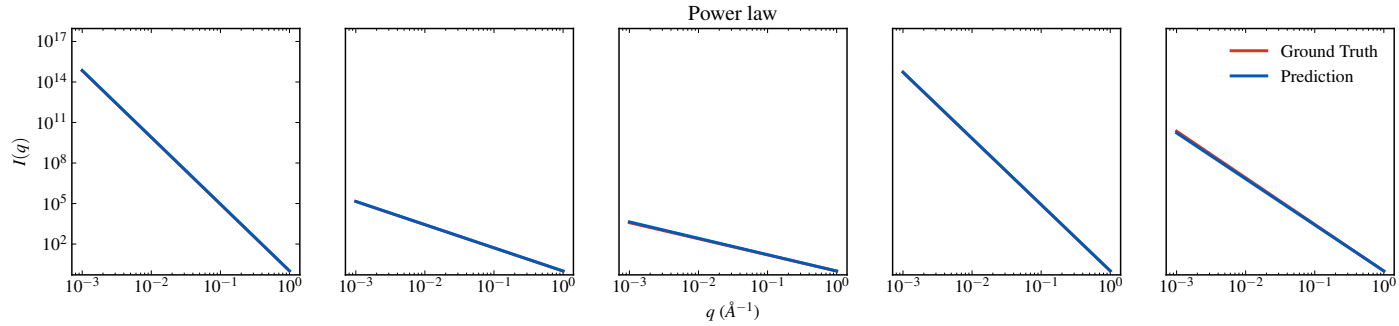

Fig. 36 Comparing predicted and ground-truth power law models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

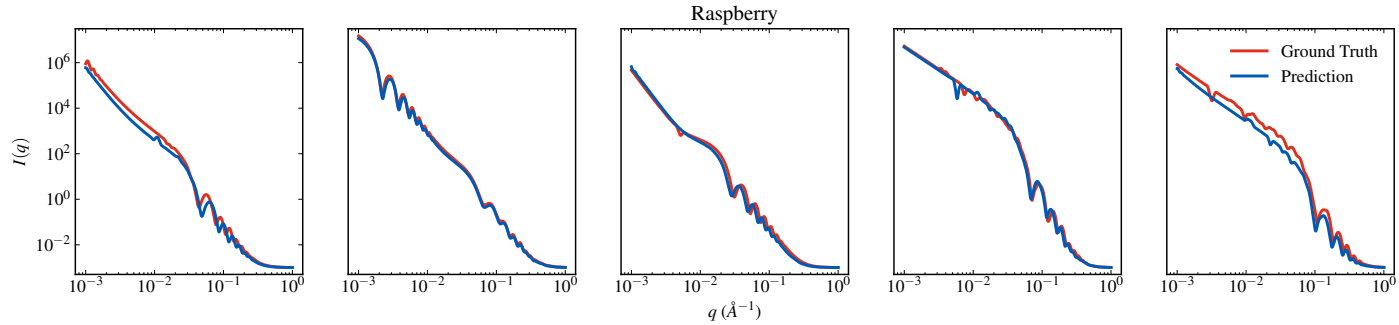

Fig. 37 Comparing predicted and ground-truth raspberry models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

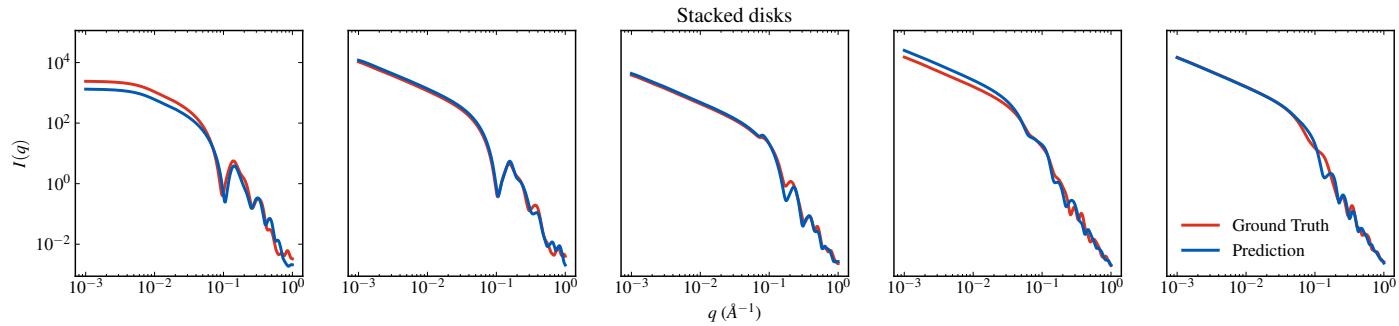

Fig. 38 Comparing predicted and ground-truth stacked disks models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

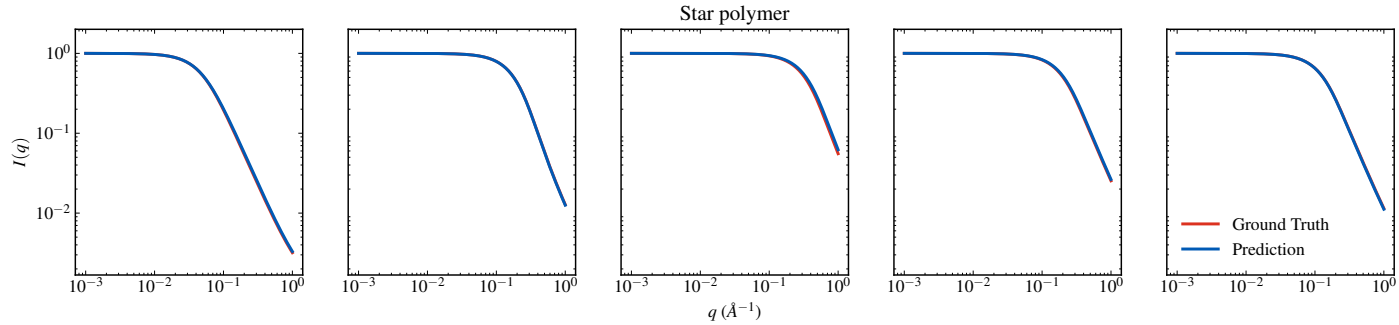

Fig. 39 Comparing predicted and ground-truth star polymer models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

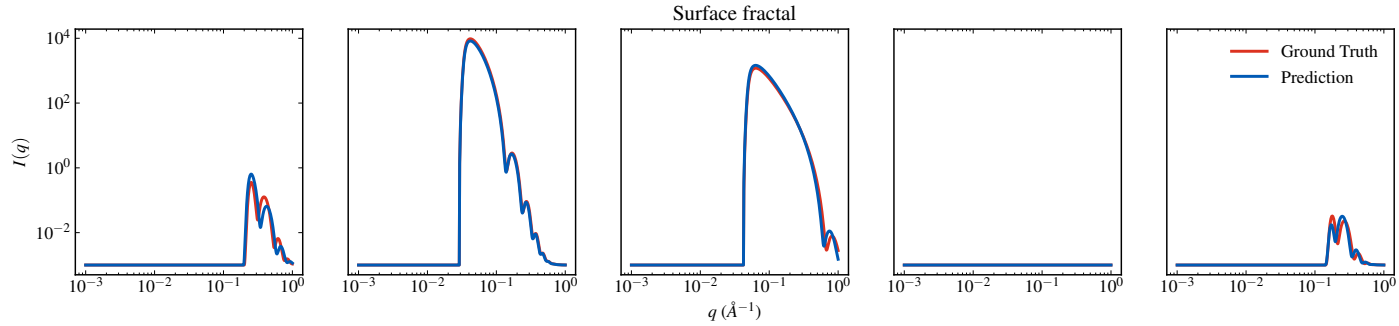

Fig. 40 Comparing predicted and ground-truth surface fractal models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

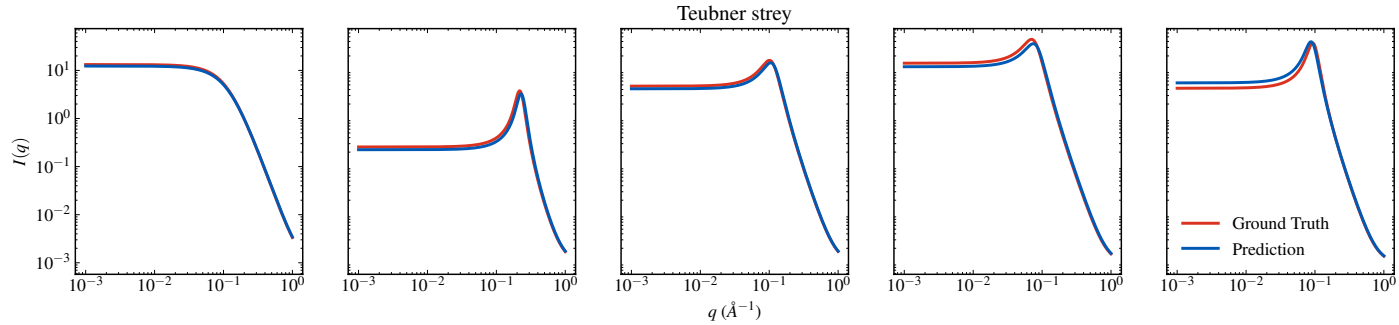

Fig. 41 Comparing predicted and ground-truth teubner strey models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

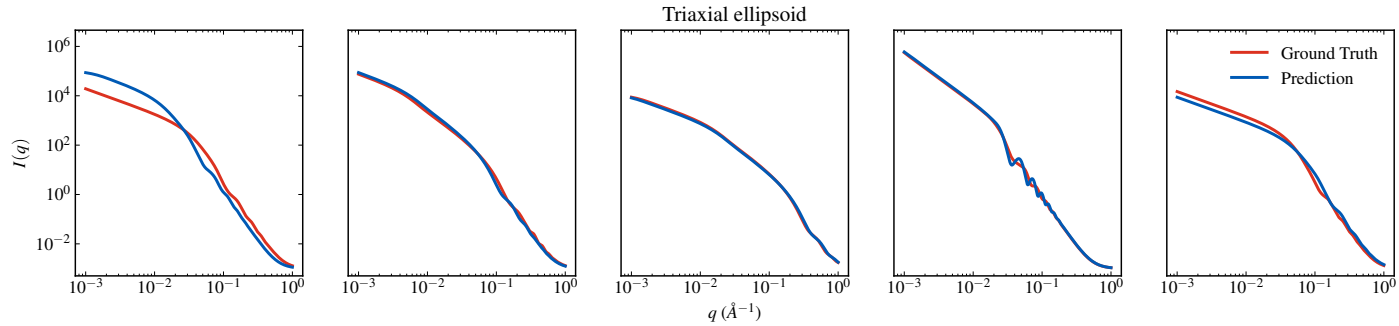

Fig. 42 Comparing predicted and ground-truth triaxial ellipsoid models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

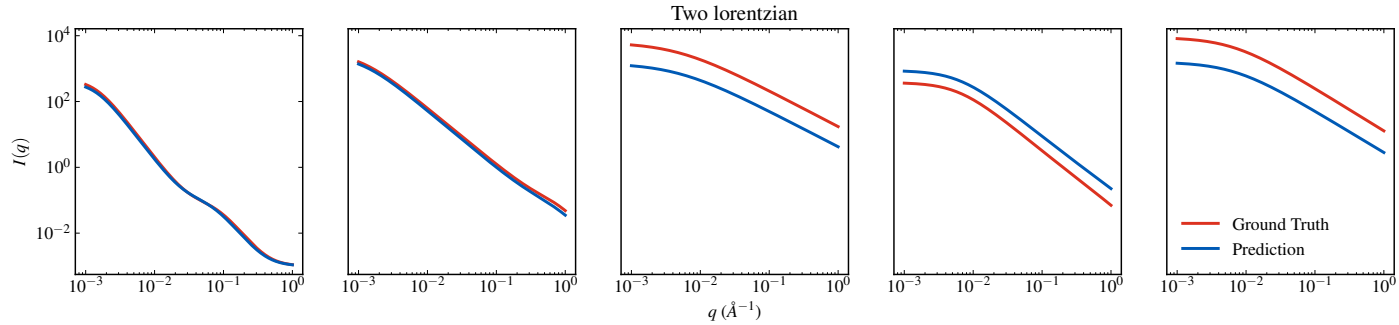

Fig. 43 Comparing predicted and ground-truth two lorentzian models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

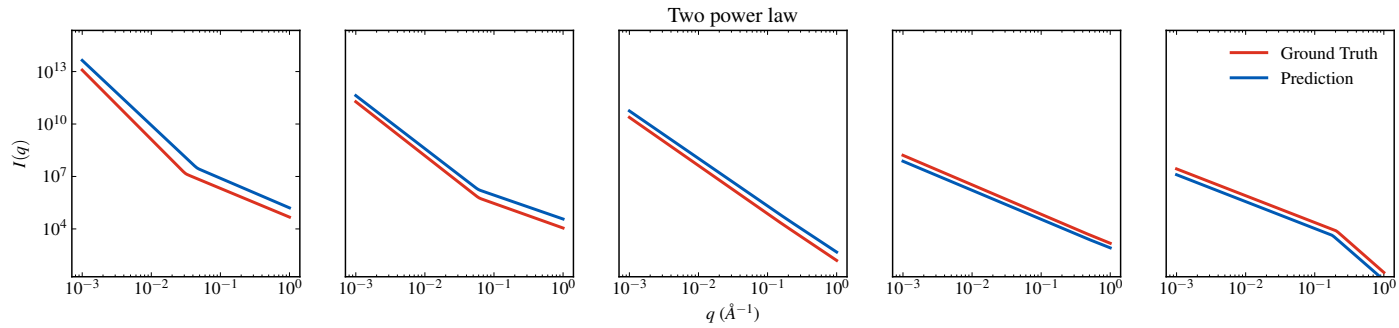

Fig. 44 Comparing predicted and ground-truth two power law models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

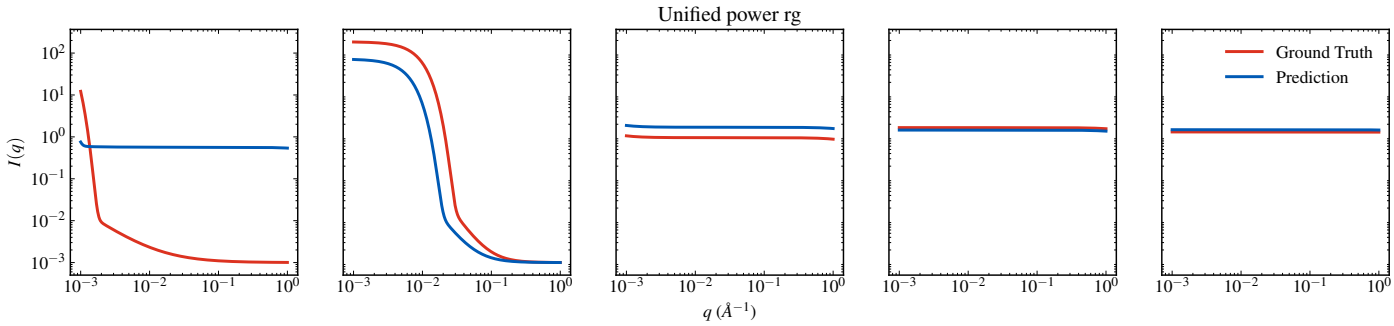

Fig. 45 Comparing predicted and ground-truth unified power Rg models. Predicted intensities are constructed from SASformer's predicted scattering-model parameters using the sasmodels Python package.

Notes and references

1 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, *Computing Research Repository (CoRR)*, 2017, **abs/1706.03762**,.

2 D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2014, https://arxiv.org/abs/1412.6980.

3 I. Loshchilov and F. Hutter, *Decoupled Weight Decay Regularization*, 2017, https://arxiv.org/abs/1711.05101.

4 I. Loshchilov and F. Hutter, *SGDR: Stochastic Gradient Descent with Warm Restarts*, 2016, https://arxiv.org/abs/1608.03983.