SPLR: The Sparse-Plus-Low-Rank Quasi-Newton Method for Entropic-Regularized Optimal Transport

Chenrui Wang[1], Yixuan Qiu[1]

[1] Shanghai University of Finance and Economics

Problem Definition

Our goal is to solve the entropic-regularized optimal transport problem, and our approach is to study its dual problem. After removing one redundant degree of freedom, we obtain the *smooth* and *unconstrained convex* optimization problem:

$$\begin{aligned} f(x) =& -\mathcal{L}(\alpha, \beta) \\ =& \eta \sum \exp\{\eta^{-1}(\alpha_i + \beta_j - M_{ij})\} \\ & -\alpha^T a - \beta^T b, \end{aligned}$$

where $x = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_{m-1})^T$.

Motivation

The main idea of SPLR is to provide an *accurate approximation* of the Hessian matrix while maintaining its *sparsity* as much as possible:

- **Sparse:** Some second-order methods have been proposed based on the idea of Hessian sparsification. SPLR also obtain an approximation of the true Hessian matrix by sparsification.
- **Low-Rank:** Similar to the BFGS update rule, we incorporate a low-rank correction term to enhance the approximation quality.

Contribution

Our main contributions are summarized as follows:

- New theoretical results are developed to understand Hessian sparsification.
- A new quasi-Newton method is proposed to solve entropic-regularized OT with *super-linear-like* convergence speed in practice.
- We provide convergence guarantees for the proposed method: SPLR enjoys a *global convergence* property and the convergence speed is at least *linear*.
- We conduct extensive numerical experiments to demonstrate the performance of SPLR on various entropic-regularized OT problems.

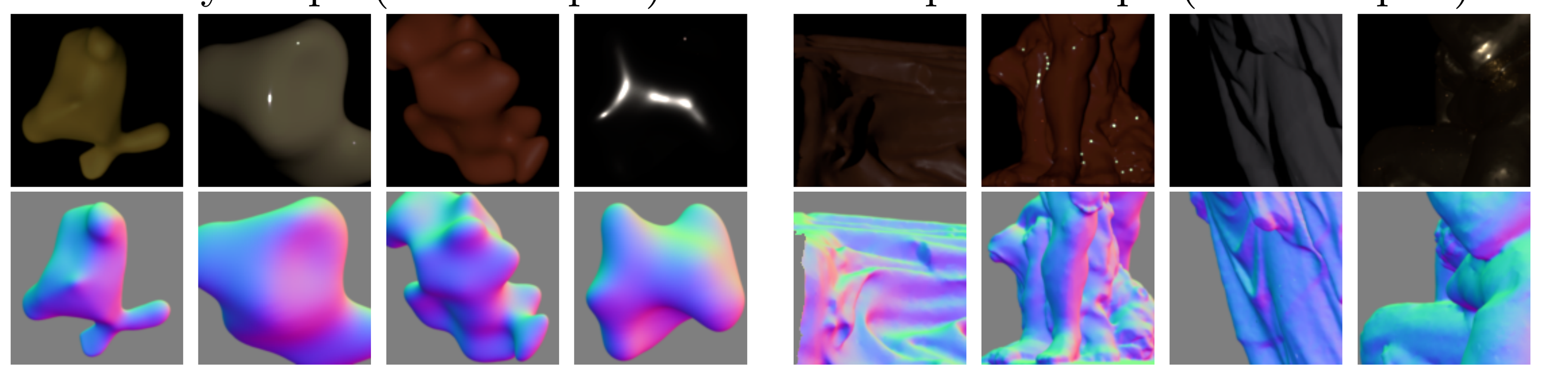
Hessian Sparsification

foobar

Experiments

Synthetic Datasets for Training:

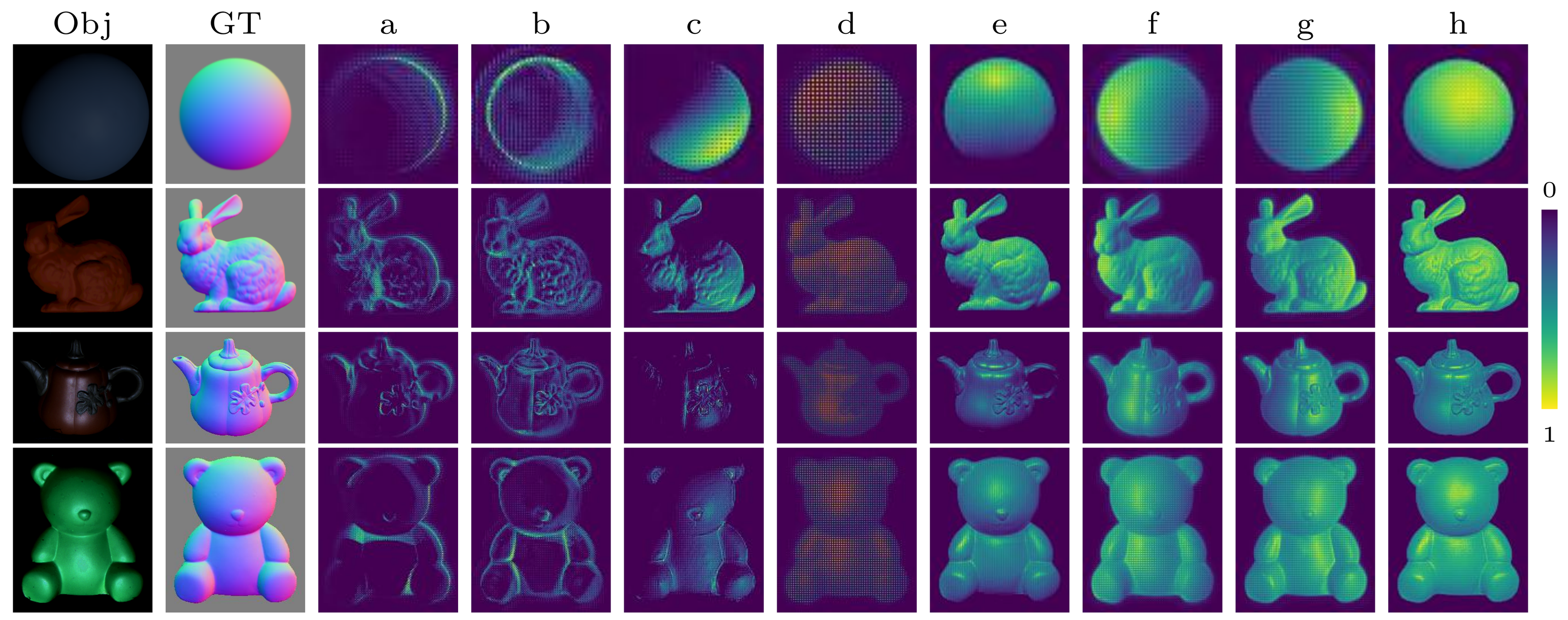
Blobby shape (26K samples). Sculpture shape (59K samples).



Quantitative Results on DiLiGenT Main Dataset:

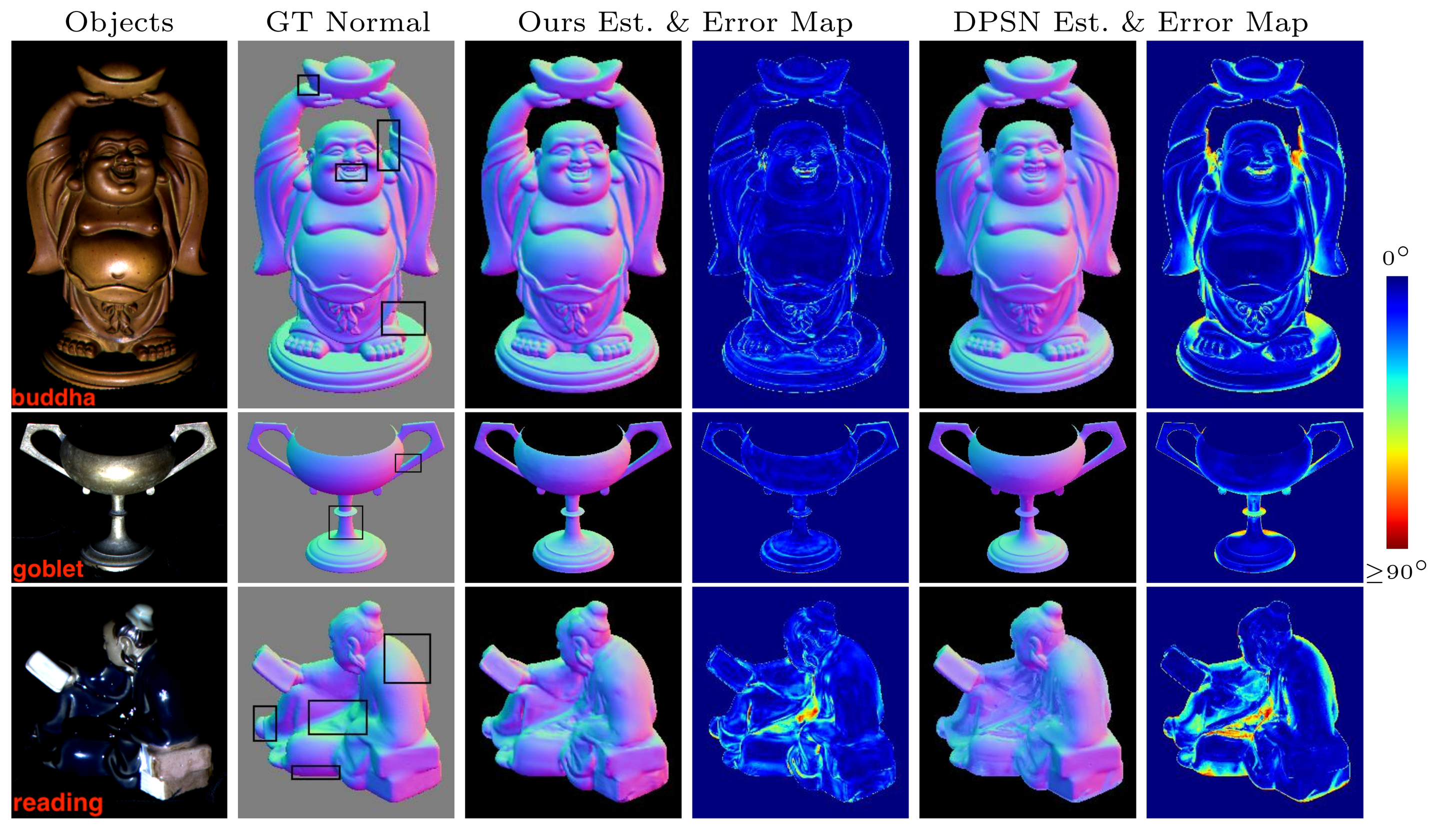
Method	ball	cat	pot1	bear	pot2	buddha	goblet	reading	cow	harvest	Avg.
L2	4.10	8.41	8.89	8.39	14.65	14.92	18.50	19.80	25.60	30.62	15.39
AZ08	2.71	6.53	7.23	**5.96**	11.03	12.54	13.93	14.17	21.48	30.50	12.61
WG10	2.06	6.73	7.18	6.50	13.12	10.91	15.70	15.39	25.89	30.01	13.35
IA14	3.34	6.74	6.64	7.11	8.77	10.47	9.71	14.19	13.05	25.95	10.60
ST14	**1.74**	**6.12**	**6.51**	6.12	8.78	10.60	10.09	13.63	13.93	25.44	10.30
DPSN	2.02	6.54	7.05	6.31	7.86	12.68	11.28	15.51	8.01	16.86	9.41
PS-FCN (B+S+32, 16)	3.31	7.64	8.14	7.47	8.22	8.76	9.81	14.09	8.78	17.48	9.37
PS-FCN (B+S+32, 96)	2.82	6.16	7.13	7.55	**7.25**	**7.91**	**8.60**	**13.33**	**7.33**	**15.85**	**8.39**

Feature Visualization:

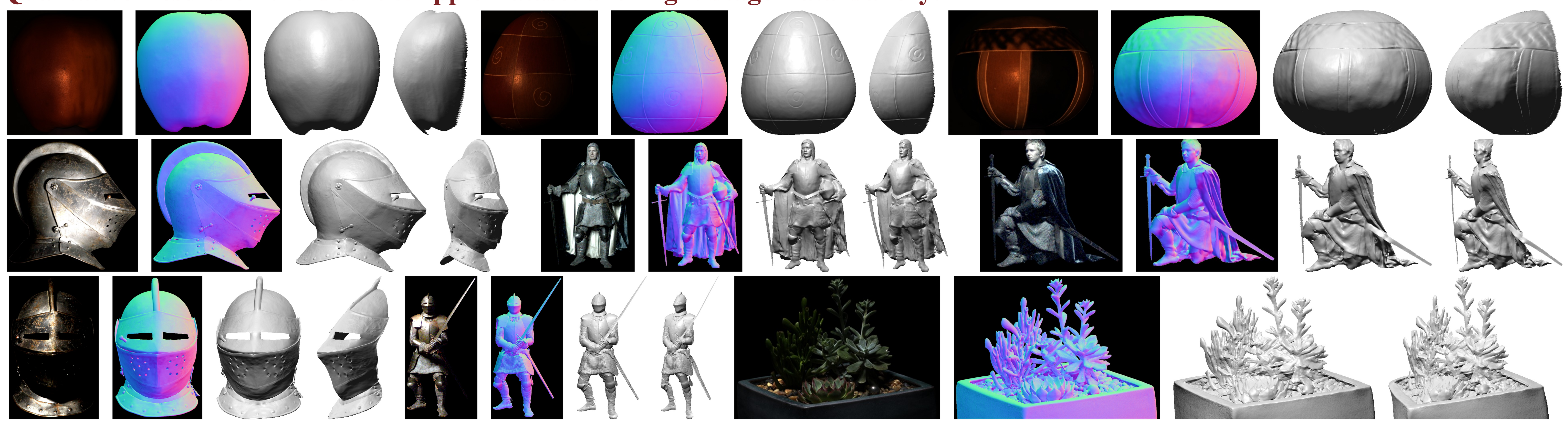


- Different regions with similar normal directions are fired in different channels. Each channel can therefore be interpreted as the probability of the normal belonging to a certain direction.

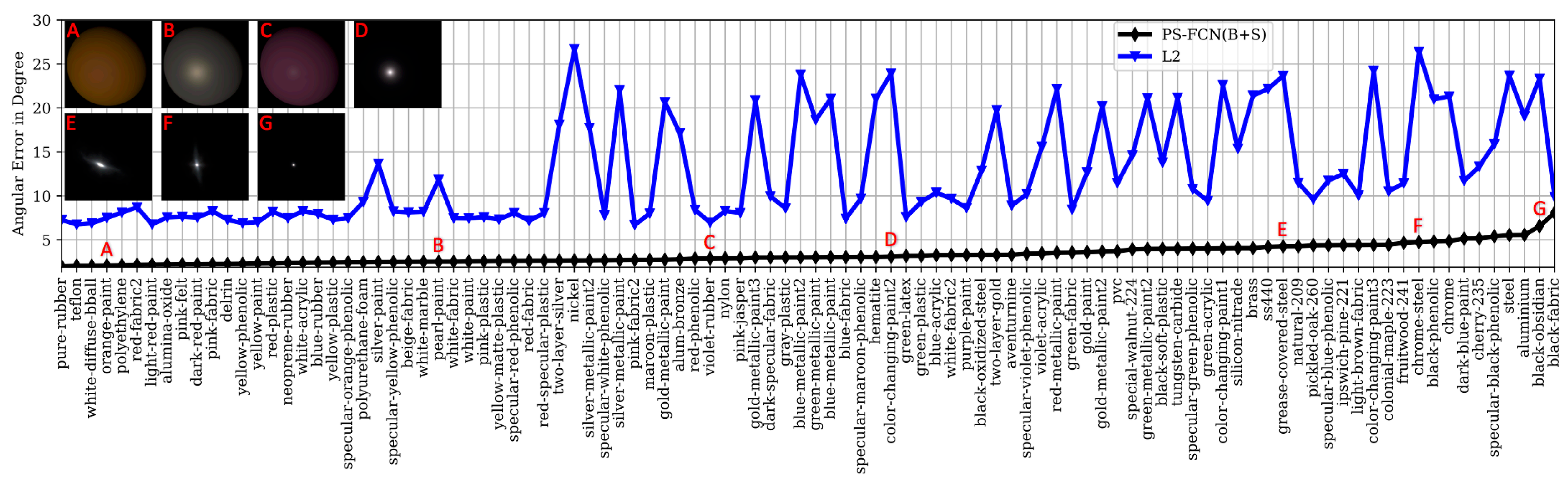
Qualitative Results on DiLiGenT Main Dataset:



Qualitative Results on the Gourd&Apple Dataset and Light Stage Data Gallery:



Quantitative Results on Spheres Rendered with 100 Different Materials:



Quantitative Results of Uncalibrated PS-FCN on DiLiGenT Main Dataset:

Method	ball	cat	pot1	bear	pot2	buddha	goblet	reading	cow	harvest	Avg.
AM07	7.27	31.45	18.37	16.81	49.16	32.81	46.54	53.65	54.72	61.70	37.25
SM10	8.90	19.84	16.68	11.98	50.68	15.54	48.79	26.93	22.73	73.86	29.59
WT13	**4.39**	36.55	**9.39**	**6.42**	14.52	**13.19**	20.57	58.96	19.75	55.51	23.93
PF14	4.77	**9.54**	9.51	9.07	15.90	14.92	29.93	24.18	19.53	29.21	16.66
LC18	9.30	12.60	12.40	10.90	15.70	19.00	**18.30**	**22.30**	15.00	28.00	16.30
UPS-FCN	6.62	14.68	13.98	11.23	**14.19**	15.87	20.72	23.26	**11.91**	**27.79**	**16.02**

Project Webpage:

Code & Dataset & Model