Supplementary Material for Inverse Rendering for Discrete X-Ray Computed Tomography

LOVRO NUIC, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
ZIYI ZHANG, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
KORBINIAN SAGER, Carl Zeiss AG, Germany
WENZEL JAKOB, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ACM Reference Format:
Lovro Nuic, Ziyi Zhang, Korbinian Sager, and Wenzel Jakob. 2026. Supplementary Material for Inverse Rendering for Discrete X-Ray Computed Tomography. *ACM Trans. Graph.* 45, 4, Article 134 (July 2026), 10 pages. https://doi.org/10.1145/3811391

1 Convergence Analysis

We analyze the convergence behavior of the probabilistic optimization in Section 3. We show that the optimization converges to a deterministic assignment that is a 1-neighbor local minimum where no single-variable change can reduce the loss. Throughout this analysis, surrogate evaluations are treated as fixed during each per-variable update, so there is no inter-variable coupling through the surrogate within a single update step.

1.1 Definitions

We optimize the global objective defined in Equation 7:

$$\mathcal{J}(\boldsymbol{\Theta}) \coloneqq \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{m}\theta_{i,k}\,\tilde{\mathcal{L}}_i(k), \tag{1}$$

where $\tilde{\mathcal{L}}_i(k) = \mathcal{L}(\mathcal{S}^{i\to k}(\boldsymbol{\Theta}))$ evaluates the loss with variable X_i fixed to k and a surrogate applied to the remaining variables.

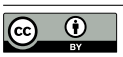
Definition 1.1 (1-neighbor local minimum). A configuration $\boldsymbol{X}^* = (X_1^*, \ldots, X_N^*) \in \{1, \ldots, m\}^N$ is a *1-neighbor local minimum* if no single-variable change reduces the loss:

$$\mathcal{L}(\boldsymbol{X}^*) \le \mathcal{L}(X_i = k, \boldsymbol{X}^*_{-i}) \quad \text{for all } i \in \{1, \ldots, N\},\ k \in \{1, \ldots, m\}, \tag{2}$$

where $\boldsymbol{X}^*_{-i} = (X_1^*, \ldots, X_{i-1}^*, X_{i+1}^*, \ldots, X_N^*)$ denotes all variables except i.

Definition 1.2 (Deterministic distribution). A distribution $p(\boldsymbol{X} \mid \boldsymbol{\Theta})$ is *deterministic* if each $\boldsymbol{\theta}_i$ is a one-hot vector, i.e., $\boldsymbol{\theta}_i = \mathbf{e}_{k_i^*}$ for some $k_i^* \in \{1, \ldots, m\}$. The corresponding configuration is $\boldsymbol{X}^* = (k_1^*, \ldots, k_N^*)$.

Authors' Contact Information: Lovro Nuic, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, lovro.nuic@epfl.ch; Ziyi Zhang, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, ziyi.zhang@epfl.ch; Korbinian Sager, Carl Zeiss AG, Oberkochen, Germany; Wenzel Jakob, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, wenzel.jakob@epfl.ch.



ACM 1557-7368/2026/7-ART134
https://doi.org/10.1145/3811391

1.2 Convergence to Deterministic Assignments

We first establish that the optimization drives each distribution toward a deterministic assignment.

Lemma 1.3 (Optimality at simplex vertices). *For any fixed surrogate evaluations $\{\tilde{\mathcal{L}}_i(k)\}_{k=1}^m$, the per-variable objective*

$$\mathcal{J}_i(\boldsymbol{\theta}_i) = \sum_{k=1}^{m}\theta_{i,k}\tilde{\mathcal{L}}_i(k) \tag{3}$$

is minimized at a one-hot vector $\boldsymbol{\theta}_i = \mathbf{e}_{k_i^*}$ where $k_i^* = \arg\min_k \tilde{\mathcal{L}}_i(k)$.

Proof. The function $\mathcal{J}_i(\boldsymbol{\theta}_i)$ is linear in $\boldsymbol{\theta}_i$: it is a weighted sum of the fixed values $\{\tilde{\mathcal{L}}_i(k)\}_k$ with weights $\{\theta_{i,k}\}_k$.

The probability simplex $\Delta_{m-1} = \{\boldsymbol{\theta} \in \mathbb{R}^m \mid \theta_k \ge 0,\ \sum_k \theta_k = 1\}$ is a convex polytope whose vertices are the one-hot vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$. Since the objective is linear over a simplex, a minimizer occurs at an extreme point.

Therefore, $\mathcal{J}_i(\boldsymbol{\theta}_i)$ is minimized at $\boldsymbol{\theta}_i = \mathbf{e}_{k_i^*}$ where $k_i^* = \arg\min_k \tilde{\mathcal{L}}_i(k)$. □

Theorem 1.4 (Convergence to a deterministic distribution). *Gradient descent on $\mathcal{J}(\boldsymbol{\Theta})$, with surrogate evaluations treated as fixed during each update, converges to the set of minimizers; if the minimizer is unique then $\boldsymbol{\theta}_i$ converges to a one-hot vector.*

Proof. When surrogate evaluations are treated as fixed, the global objective decomposes as

$$\mathcal{J}(\boldsymbol{\Theta}) = \frac{1}{N}\sum_{i=1}^{N}\mathcal{J}_i(\boldsymbol{\theta}_i) = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{m}\theta_{i,k}\,\tilde{\mathcal{L}}_i(k), \tag{4}$$

so each $\mathcal{J}_i(\boldsymbol{\theta}_i)$ is linear in $\boldsymbol{\theta}_i$ over the probability simplex.

Let $S_i \coloneqq \arg\min_k \tilde{\mathcal{L}}_i(k)$. Then for any $\boldsymbol{\theta}_i \in \Delta_{m-1}$,

$$\mathcal{J}_i(\boldsymbol{\theta}_i) = \sum_{k=1}^{m}\theta_{i,k}\tilde{\mathcal{L}}_i(k) \ \ge\ \sum_{k=1}^{m}\theta_{i,k}\min_{k'}\tilde{\mathcal{L}}_i(k') \ =\ \min_{k'}\tilde{\mathcal{L}}_i(k'). \tag{5}$$

Equality holds if and only if $\theta_{i,k} = 0$ for all candidates $k \notin S_i$, so all probability mass must be distributed among the minimizing candidates.

In the simplest case, where a single candidate k_i^* is strictly better than the others, the distribution $\boldsymbol{\theta}_i$ is forced to the vertex $\mathbf{e}_{k_i^*}$ (a one-hot vector). If multiple candidates tie for the minimum, any point on the face of the simplex formed by those candidates is an optimal solution.

Since projected gradient descent consistently moves the parameters toward lower objective values, the iterates converge to this optimal face which corresponds to the convex hull of $\{\mathbf{e}_k : k \in S_i\}$. □

1.3 Surrogate Equivalence at Deterministic Distributions

Lemma 1.5 (Surrogate equivalence). *At a deterministic distribution with configuration $\boldsymbol{X}^*$, all three surrogates (mode, sample, expectation) yield:*

$$\tilde{\mathcal{L}}_i(k) = \mathcal{L}(X_i = k, \boldsymbol{X}^*_{-i}). \tag{6}$$

Proof. We show that $\mathcal{S}_j(\boldsymbol{\theta}_j) = X_j^*$ when $\boldsymbol{\theta}_j = \mathbf{e}_{X_j^*}$ for each surrogate:

Mode: $\mathcal{S}_j(\boldsymbol{\theta}_j) = \arg\max_{k'} \theta_{j,k'} = X_j^*$ since $\theta_{j,X_j^*} = 1$.

Sample: $\mathcal{S}_j(\boldsymbol{\theta}_j) \sim \mathrm{Cat}(\boldsymbol{\theta}_j)$ yields X_j^* with probability 1.

Expectation: $\mathcal{S}_j(\boldsymbol{\theta}_j) = \sum_{k'} \theta_{j,k'} \xi^{(k')} = \xi^{(X_j^*)}$.

Thus $\mathcal{S}^{i\to k}(\boldsymbol{\Theta})$ produces $(X_i = k, \boldsymbol{X}^*_{-i})$ for all surrogates. □

1.4 Main Result

Theorem 1.6 (Convergence to 1-neighbor local minimum). *When the optimization converges to a deterministic distribution $\boldsymbol{\Theta}^*$, the corresponding configuration $\boldsymbol{X}^* = (k_1^*, \ldots, k_N^*)$ is a 1-neighbor local minimum.*

Proof. By Lemma 1.5, at a deterministic point $\boldsymbol{\Theta}^*$ the surrogate losses coincide with the true losses under single-site perturbations:

$$\tilde{\mathcal{L}}_i(k) = \mathcal{L}(X_i = k, \boldsymbol{X}^*_{-i}) \qquad \forall i,\ \forall k. \tag{7}$$

In particular, $\tilde{\mathcal{L}}_i(k_i^*) = \mathcal{L}(\boldsymbol{X}^*)$.

Assume for contradiction that $\boldsymbol{X}^*$ is not a 1-neighbor local minimum. Then there exist an index i and a label $k \neq k_i^*$ such that

$$\mathcal{L}(X_i = k, \boldsymbol{X}^*_{-i}) < \mathcal{L}(\boldsymbol{X}^*), \tag{8}$$

i.e., $\tilde{\mathcal{L}}_i(k) < \tilde{\mathcal{L}}_i(k_i^*)$.

Consider a perturbed distribution $\tilde{\boldsymbol{\Theta}}$ that only modifies block i by shifting an ϵ-fraction of probability mass from k_i^* to k:

$$\tilde{\boldsymbol{\theta}}_i = (1-\epsilon)\mathbf{e}_{k_i^*} + \epsilon \mathbf{e}_k, \qquad \tilde{\boldsymbol{\theta}}_j = \boldsymbol{\theta}_j^* \text{ for } j \neq i, \tag{9}$$

for some $\epsilon \in (0, 1)$.

Since the surrogate terms are evaluated at $\boldsymbol{\Theta}^*$ and are locally unchanged for sufficiently small ϵ, the change in the objective is governed by the i-th block:

$$\mathcal{J}(\tilde{\boldsymbol{\Theta}}) - \mathcal{J}(\boldsymbol{\Theta}^*) = \frac{1}{N}\Big(\sum_{k'} \tilde{\theta}_{i,k'}\, \tilde{\mathcal{L}}_i(k') - \tilde{\mathcal{L}}_i(k_i^*) \Big) \tag{10}$$

$$= \frac{\epsilon}{N}\big(\tilde{\mathcal{L}}_i(k) - \tilde{\mathcal{L}}_i(k_i^*) \big) \tag{11}$$

$$< 0. \tag{12}$$

Thus, for sufficiently small ϵ, we obtain $\mathcal{J}(\tilde{\boldsymbol{\Theta}}) < \mathcal{J}(\boldsymbol{\Theta}^*)$, contradicting the fact that $\boldsymbol{\Theta}^*$ is a limit point of a descent procedure.

Therefore no such improving 1-neighbor move exists, and $\boldsymbol{X}^*$ is a 1-neighbor local minimum. □

1.5 Discussion

Surrogate coupling. The expectation surrogate causes $\tilde{\mathcal{L}}_i(k)$ to depend on $\boldsymbol{\Theta}$ through expected material properties at other variables. As a result, the local linearity argument that favors one-hot minimizers no longer results in a global guarantee of convergence to a deterministic distribution.

Global vs. local minima. Our analysis guarantees convergence to a 1-neighbor local minimum, not a global minimum. Discrete optimization problems such as material assignment in CT are often NP-hard [Gardner et al. 1999; Gerard 2009]. The quality of the solution depends on initialization and the optimization path.

Ties. When multiple candidates achieve equal minimum loss, the optimization may converge to any mixture of tied candidates. In practice, exact ties are rare, and regularization or small perturbations break symmetry.

2 Ablation Study: Probabilistic Model Design Choices

This section provides a detailed analysis of the design choices in our probabilistic framework, supplementing the experimental evaluation presented in Section 3 of the main paper. The convergence properties analyzed in Section 1 motivate several of the design decisions discussed here. We conduct systematic ablation studies using 2D phantoms with full 360° angular coverage, allowing us to isolate the effects of different algorithmic components. Reconstruction performance is evaluated by increasing the number of projections based on golden ratio angular sampling.

For the reader's convenience, we reproduce Figure 4 from the main paper below.

Surrogate selection. We compare the three surrogate strategies introduced in Section 3 of the main paper: mode, sample, and expectation. Figure 1 shows reconstruction accuracy and mean intersection-over-union (IoU) as a function of the number of input projections.

The **mode surrogate** (Figure 1f) selects the most probable material at each voxel, ignoring uncertainty in the distribution. While computationally simple, this approach suffers from premature commitment: early in optimization, when distributions are nearly uniform, the mode is unstable and can trap the optimization in poor local configurations. Performance improves with more projections but remains consistently below other methods.

The **sample surrogate** (Figure 1g) draws random configurations according to current probabilities. To prevent early convergence to suboptimal discrete choices, we add Gaussian noise to the probabilities that decays according to a schedule similar to simulated annealing. Specifically, at iteration t, we perturb each $\theta_{i,k}$ by $\mathcal{N}(0, \sigma_t^2)$ where $\sigma_t = \sigma_0 \cdot \exp(-t/\tau)$ before sampling. This maintains exploration early in optimization while allowing convergence to discrete solutions. Despite this regularization, the inherent variance of sampling leads to noisy gradients and slower convergence compared to the expectation surrogate.

The **expectation surrogate** (Figure 1h–i) computes the expected material properties at each voxel and evaluates how switching to each candidate material would affect the loss. This approach incorporates the full distribution without introducing variance, leading to the most stable optimization. Our experiments consistently show that the expectation surrogate achieves the highest accuracy, reaching perfect reconstruction with fewer projections than other methods.

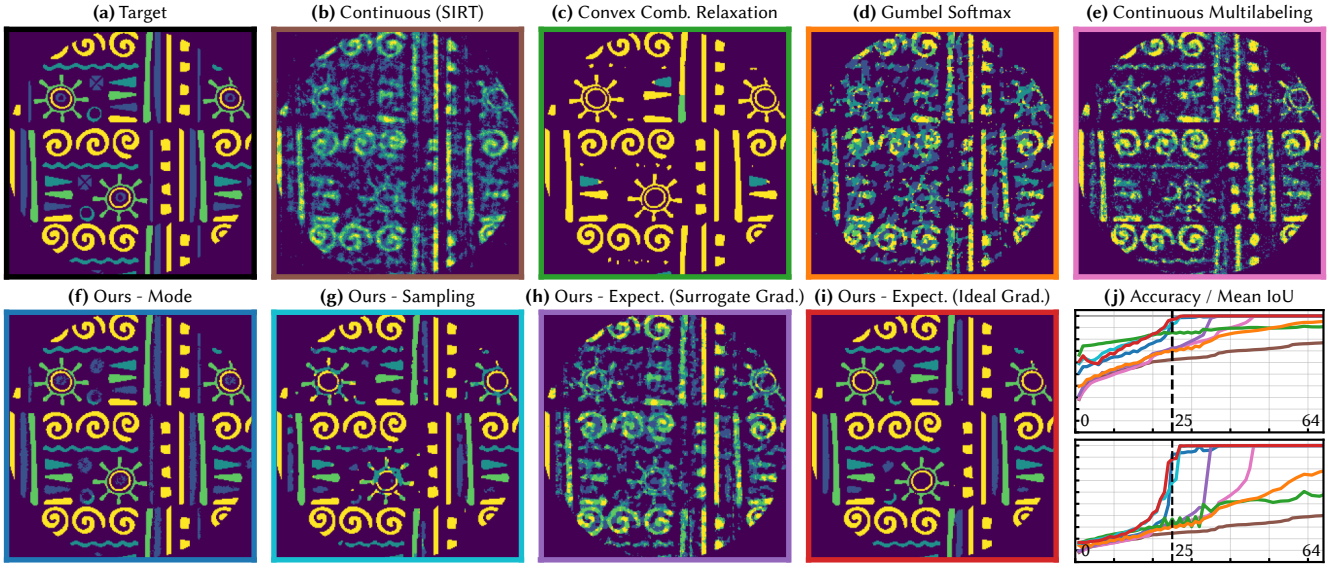


Fig. 1. **Problem formulations.** *(Reproduced from Figure 4 in the main paper for convenience.)* We compare the effectiveness of different continuous and probabilistic formulations in a synthetic 5-material CT reconstruction: (**a**) reference (2D), (**b**) continuous optimization of extinction with rounding [Gilbert 1972], (**c**) convex combinations of reference materials with a discreteness-promoting entropy loss [Gkioulekas et al. 2013; Nindel et al. 2021], (**d**) Gumbel-Softmax with a tuned temperature schedule, (**e**) continuous multilabeling [Zisler et al. 2016], **(f–i)** four variants of the proposed method using mode, sampling, and expectation-based surrogates; the latter further admits two gradient formulations (surrogate and ideal). Plot **(j)** shows reconstruction accuracy (% of correct assignments) and IoU (intersection-over-union, i.e., the average per-class overlap with ground truth) as a function of projection count. The shown reconstructions are for 25 projections (dashed vertical line). We find that the expectation-based surrogate with ideal gradients in (**i**) produces the fastest and most reliable convergence. The trends shown here also generalize to more challenging 3D reconstructions.

2.1 Comparison with Alternative Probabilistic Formulations

We compare our approach against three alternative methods for handling discrete optimization, as shown in Figure 1b–e: convex relaxation with entropy regularization, Gumbel-Softmax, and the continuous multilabeling approach of Zisler et al. [2016].

Convex relaxation with entropy. (Figure 1c). A natural baseline represents each voxel as a convex combination of material properties and propagates gradients through the expected material back to the mixing weights. We augment this with an entropy regularizer to promote sparse probability vectors:

$$\mathcal{L}_{\text{entropy}} = -\lambda_H \sum_{i=1}^{N} \sum_{k=1}^{m} \theta_{i,k} \log(\theta_{i,k} + \epsilon). \tag{13}$$

Contrary to our initial expectation that this would drive convergence toward discrete assignments, we observe a characteristic failure mode: the optimization preferentially selects the two extreme materials (lowest and highest attenuation) and interpolates between them to approximate intermediate values. Once trapped in this local minimum, the optimization has difficulty escaping to the true material labels. This behavior manifests in the accuracy curves (Figure 1j), where the method initially appears competitive, since many voxels in our test phantoms do contain the extreme materials. But the mean intersection-over-union (IoU) metric reveals performance comparable to other baselines, as the method systematically misclassifies intermediate materials.

Gumbel-Softmax. (Figure 1e). This reparameterization technique [Jang et al. 2017], discussed briefly in Section 3 of the main paper, replaces discrete sampling with a differentiable approximation, using the straight-through gradient estimator to update underlying logits. We anneal the temperature from $\tau_0 = 1.0$ to $\tau_f = 0.1$ according to an exponential schedule. This method exhibits notably slower convergence compared to alternatives, a consequence of the biased gradients inherent to the straight-through estimator. Furthermore, the temperature schedule presents a fundamental trade-off: reducing temperature too aggressively causes premature collapse to incorrect discrete solutions before the optimization has adequately explored the configuration space, while annealing too slowly prolongs reconstruction time without commensurate quality improvements. While Gumbel-Softmax has proven effective in contexts requiring hard selection, its characteristics are less well-suited to the spatially-structured discrete tomography setting.

Continuous Multilabeling [Zisler et al. 2016]. (Figure 1d). This approach, also discussed in Section 2 of the main paper, augments convex relaxation with a phase data term that biases probabilities

toward materials similar to the current expected reconstruction:

$$D(\mathbf{z}) = \sum_{i=1}^{N} \sum_{k=1}^{m} z_{ik} \left((W\mathbf{z})_i - c_k \right)^2, \tag{14}$$

where $(W\mathbf{z})_i$ is the expected intensity at voxel i and c_k is the intensity of material k. The original formulation uses indicator (box) constraints $b^{(-)} \leq AW\mathbf{z} \leq b^{(+)}$ for data fidelity, where the interval width reflects assumed noise levels. We found that for small intervals, the optimization converges extremely slowly as the feasible region severely restricts navigation through solution space. For larger intervals, the optimization may converge to discrete choices that satisfy the relaxed constraint but do not accurately represent the underlying object. Replacing the box constraint with an L_2 norm yields substantially better convergence, and we use this improved variant in our comparison.

However, this modification introduces a new challenge: the relative weighting between the data fidelity term and the phase data term becomes a sensitive hyperparameter. Emphasizing data fidelity can pull the solution toward continuous blends, while overweighting the phase term may drive premature discrete commitment before the reconstruction has converged. Despite this ambiguity, among the prior methods evaluated, this approach achieves the best performance, demonstrating the value of explicitly encouraging discrete solutions during optimization.

2.2 Analysis of Our Surrogate Strategies

Comparing our three surrogate strategies (Figure 1f–i), we find that the sampling and expectation surrogates achieve similar final reconstruction quality. However, we prefer the expectation surrogate as it avoids the inherent stochasticity of sampling, which introduces gradient noise and can destabilize optimization, particularly in later iterations when probability distributions have concentrated. The mode surrogate performs slightly worse, which we attribute to a less smooth loss landscape: small perturbations in probabilities near decision boundaries can cause discontinuous jumps in the surrogate evaluation, impeding gradient-based optimization.

2.3 Ideal Versus Surrogate Gradient

Comparing the ideal gradient (Equation 9 in the main paper) against the full surrogate gradient (Equation 8 in the main paper) with the expectation surrogate reveals an analogous issue to that observed with Zisler et al.'s method. The surrogate gradient includes coupling terms arising from the continuous dependence of the surrogate evaluation on neighboring voxels' probabilities. As noted in Section 1, this coupling means that the local linearity argument favoring one-hot minimizers no longer provides a global guarantee of convergence to deterministic distributions. These continuous components effectively introduce a competing objective that can pull the optimization away from discrete solutions toward blended configurations that minimize the surrogate's smooth approximation. The ideal gradient, by treating surrogate evaluations as fixed during each update step, avoids this interference and achieves faster, more stable convergence to discrete assignments. We therefore adopt this configuration as the default for all experiments presented in Section 4 of the main paper.

3 Real-data acquisition and materials

This section provides additional details about the real-data experiments in Section 4.3, including the CT acquisition setup, the filaments and printer used to fabricate the phantoms, and the hyperparameters specific to this experiment.

3.1 CT acquisition

We scanned both phantoms on a ZEISS Metrotom 800 industrial CT system. The X-ray tube was operated at 130 kV acceleration voltage and 300 μA tube current, with no additional pre-filtration beyond the source's intrinsic filter. The full acquisition comprises 2880 projections evenly distributed over 360°, with an integration time of 500 ms per projection. The detector has 1840×1456 pixels at a pixel pitch of 0.127 mm. The source–object and source–detector distances are 454.83 mm and 798.51 mm, respectively. Each scan is normalized using averaged flat-field (90 frames) and dark-field (10 frames) acquisitions, after which we apply the standard logarithmic transform to recover optical depths. We subsample this full set to create the sparse-view and limited-angle scenarios reported in Figures 12 and 13.

3.2 Phantom materials

The phantoms were fabricated on a Bambu Lab P2S multi-material FDM printer using the following four 1.75 mm filaments:

- **PLA** – DEEPLEE PLA (black).
- **LW-PLA** – eSUN PLA-LW.
- **S-PLA** – Formfutura StoneFil (granite).
- **ABS** – Creality CR-ABS (black).

Print parameters (nozzle temperature, bed temperature, print speed, layer height) were chosen per manufacturer recommendations for each filament.

3.3 Hyperparameters

Table 1 lists the per-material absorption coefficients $\mu^{(k)}$ and the offset prior standard deviations $\sigma_{\delta,k}$ used in the real-data experiments.

Table 1. Per-material attenuation values $\mu^{(k)}$ and offset prior standard deviations $\sigma_{\delta,k}$ used in the real-data experiments. Values are reported in the units of the FDK reference reconstruction and are not absolutely calibrated against a known reference standard.

Material	$\mu^{(k)}$	$\sigma_{\delta,k}$
Background (air)	0.000000	0.00200
LW-PLA	0.007050	0.00287
PLA	0.016509	0.00182
S-PLA	0.023067	0.00181
ABS	0.031143	0.00176

4 Additional Results

In the supplementary material, we show additional experiments on challenging 2D and 3D scenarios. All results from baseline methods

(SIRT [Gilbert 1972], DART [Batenburg and Sijbers 2011], TVR-DART [Zhuge et al. 2016]) are generated using the same settings as in the main document.

As in Figure 7 in the main document, we evaluate reconstruction quality under sparse-view (Figure 2, Figure 6), 120-degree limited-angle (Figure 3, Figure 7), 90-degree limited-angle (Figure 4, Figure 8), and noisy-input scenarios (Figure 5, Figure 9, Figure 10).

Across all scenarios—sparse views, limited-angle setups, and high-noise conditions—our method consistently outperforms baseline approaches. While traditional algorithms often degrade significantly under challenging conditions, our approach remains robust, delivering accurate reconstructions even with minimal or noisy input data.

References

Kees Joost Batenburg and Jan Sijbers. 2011. DART: A Practical Reconstruction Algorithm for Discrete Tomography. *IEEE Transactions on Image Processing* 20, 9 (2011).

R.J. Gardner, P. Gritzmann, and D. Prangenberg. 1999. On the computational complexity of reconstructing lattice sets from their X-rays. *Discrete Mathematics* 202, 1 (1999).

Yan Gerard. 2009. About the complexity of timetables and 3-dimensional discrete tomography: A short proof of NP-hardness. In *Combinatorial Image Analysis: 13th International Workshop, IWCIA 2009, Playa del Carmen, Mexico, November 24-27, 2009. Proceedings 13*. Springer, 289–301.

Peter Gilbert. 1972. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology* 36, 1 (1972), 105–117.

Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin. 2013. Inverse volume rendering with material dictionaries. *ACM Trans. Graph.* 32, 6, Article 162 (Nov. 2013), 13 pages. doi:10.1145/2508363.2508377

Eric Jang, Shixiang Gu, and Ben Poole. 2017. Categorical Reparameterization with Gumbel-Softmax. arXiv:1611.01144 [stat.ML] https://arxiv.org/abs/1611.01144

Thomas Klaus Nindel, Tomáš Iser, Tobias Rittig, Alexander Wilkie, and Jaroslav Křivánek. 2021. A gradient-based framework for 3D print appearance optimization. *ACM Trans. Graph.* 40, 4, Article 178 (July 2021), 15 pages. doi:10.1145/3450626.3459844

Xiaodong Zhuge, Willem Jan Palenstijn, and Kees Joost Batenburg. 2016. TVR-DART: A More Robust Algorithm for Discrete Tomography From Limited Projection Data With Automated Gray Value Estimation. *IEEE Transactions on Image Processing* 25, 1 (2016), 455–468.

Matthias Zisler, Stefania Petra, Claudius Schnörr, and Christoph Schnörr. 2016. Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints. In *Pattern Recognition*, Bodo Rosenhahn and Bjoern Andres (Eds.). Springer International Publishing, Cham, 261–272.

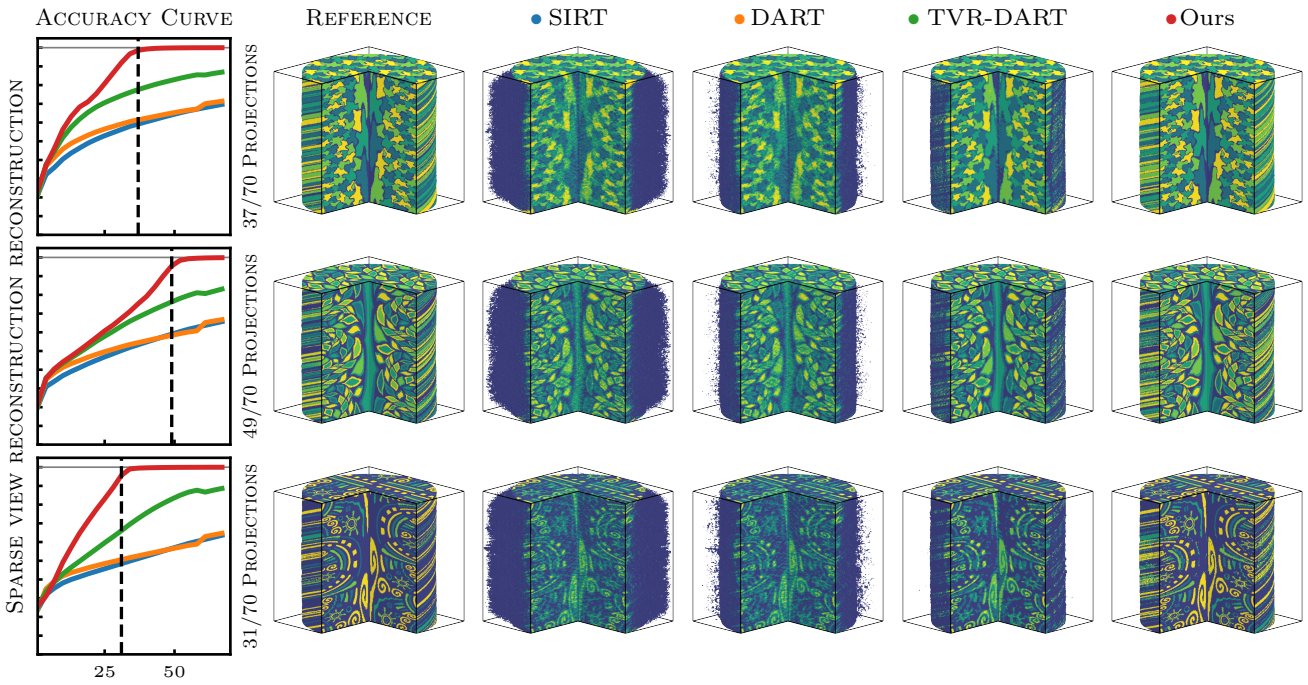


Fig. 2. **Sparse-view 3D reconstruction**. We evaluate the reconstruction quality using only a limited number of projections. The left column plots accuracy—measured as the percentage of correctly classified materials—versus the number of projections. The right column presents visual comparisons at a specific low projection count, marked by the dashed line, comparing our method with three baselines: SIRT, DART, and TVR-DART.

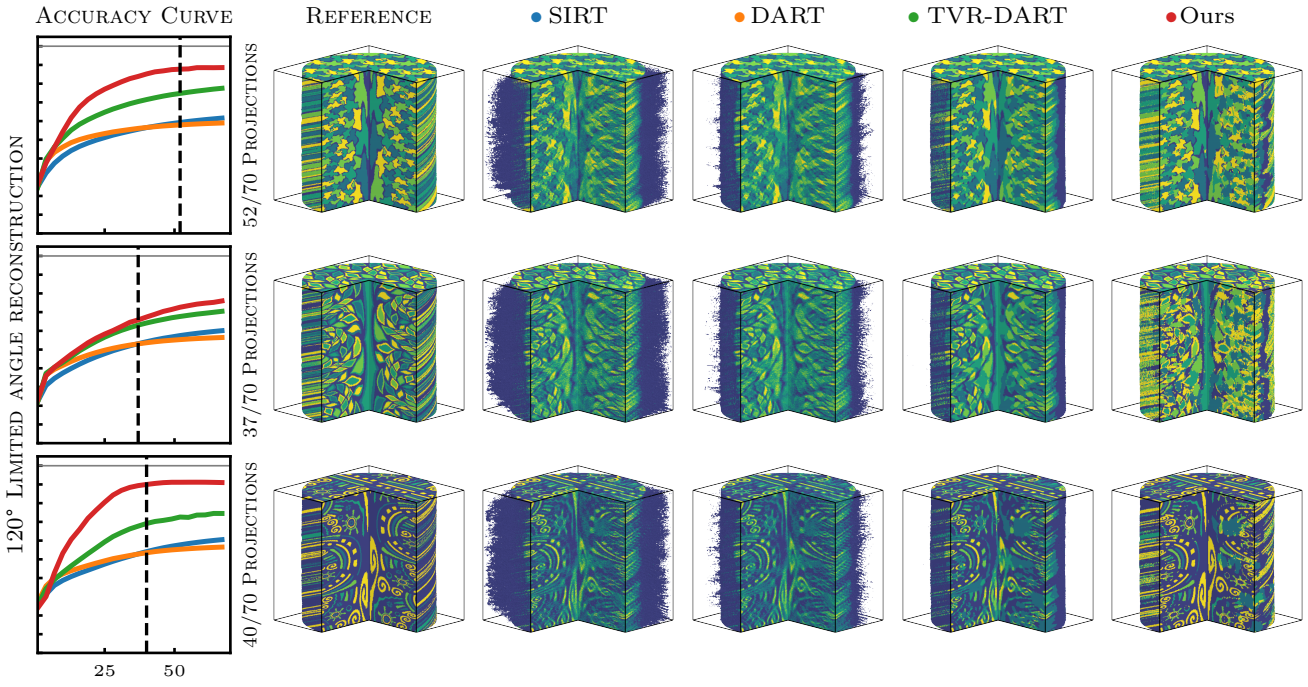


Fig. 3. **120-degree limited-angle 3D reconstruction**. We evaluate the reconstruction quality using a limited set of projections covering only 120 degrees.

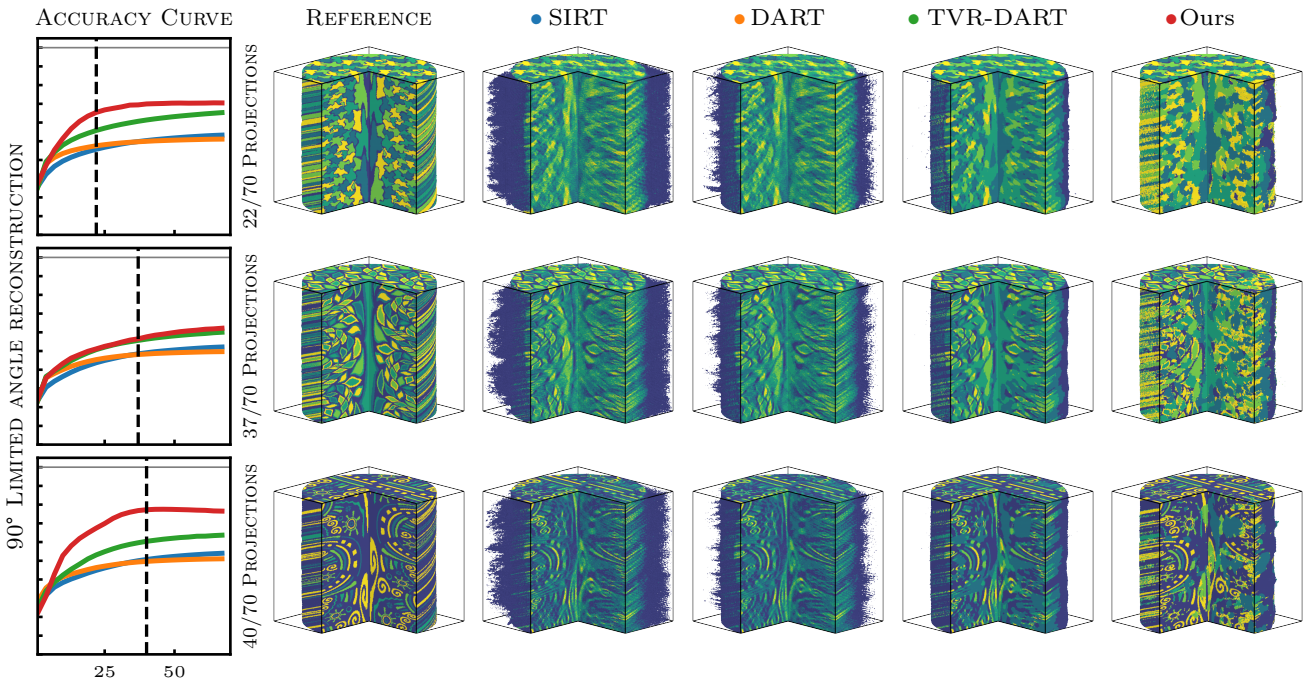


Fig. 4. **90-degree limited-angle 3D reconstruction**. We evaluate the reconstruction quality using projections covering only 90 degrees. This scenario is particularly challenging for traditional methods.

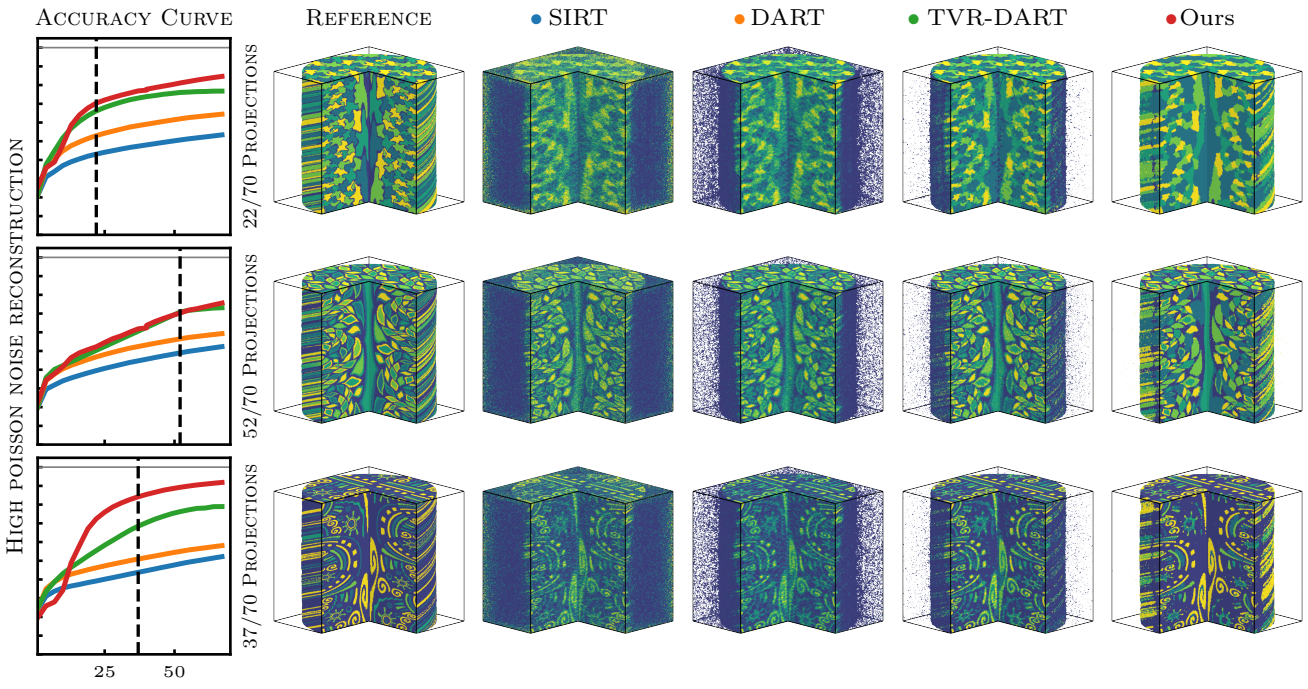


Fig. 5. **High-noise 3D reconstruction**. We evaluate the reconstruction quality in a high-noise scenario, where Poisson noise is added to the input projections.

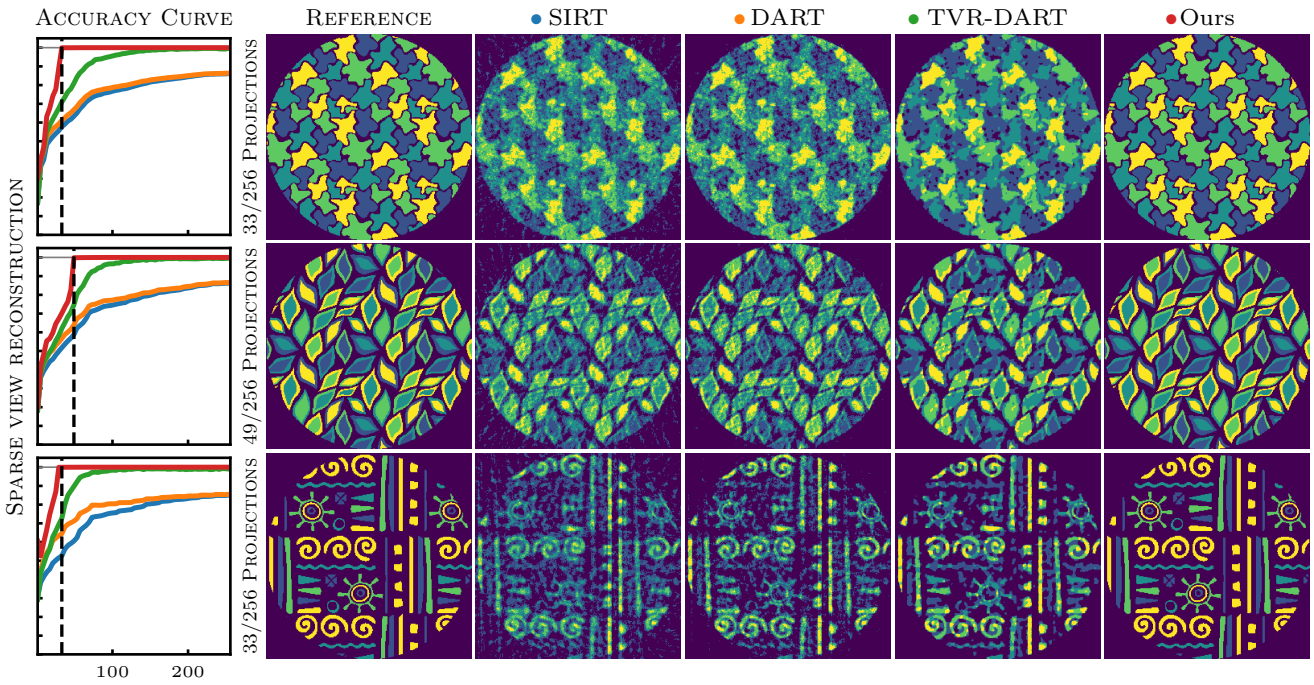


Fig. 6. **Sparse-view 2D reconstruction**. We evaluate the reconstruction quality using only a limited number of projections.

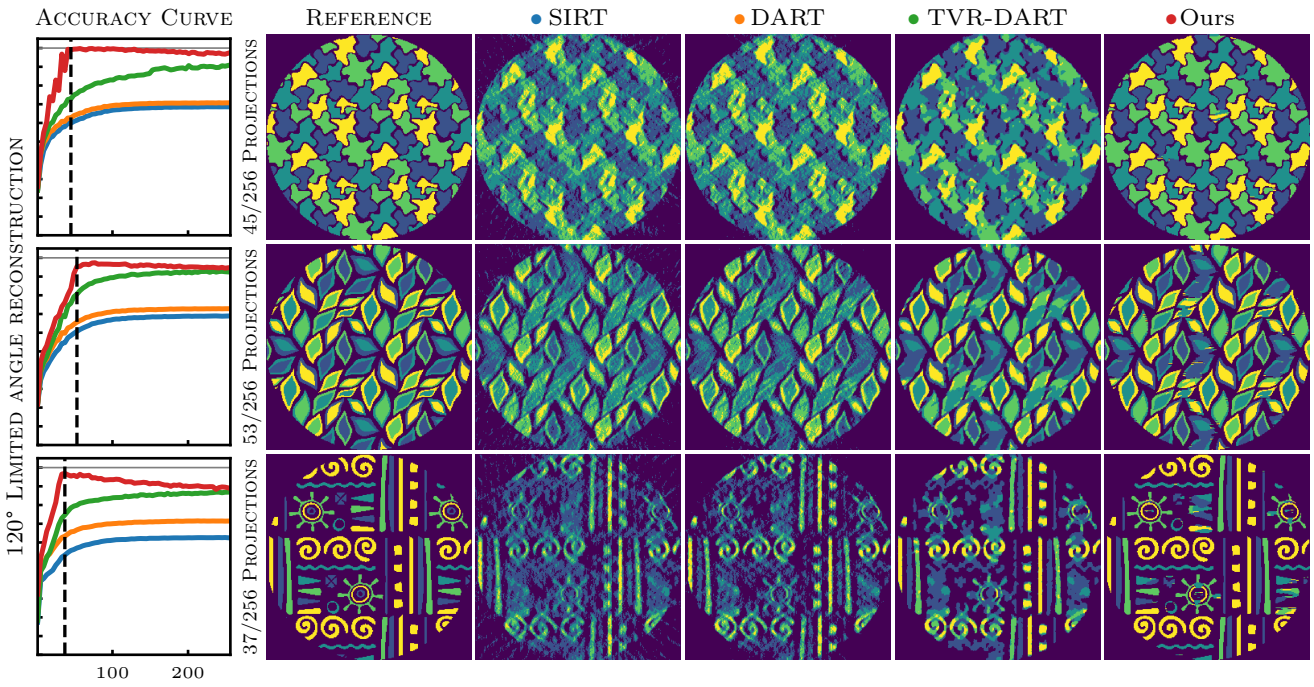


Fig. 7. **120-degree limited-angle 2D reconstruction**. We evaluate the reconstruction quality using projections covering only 120 degrees. Note that our accuracy slightly drops as we use an increasing number of projections in this scenario. This is due to overfitting to the limited-angle data, which can degrade overall reconstruction quality. This effect is not observed when using full 360-degree coverage.

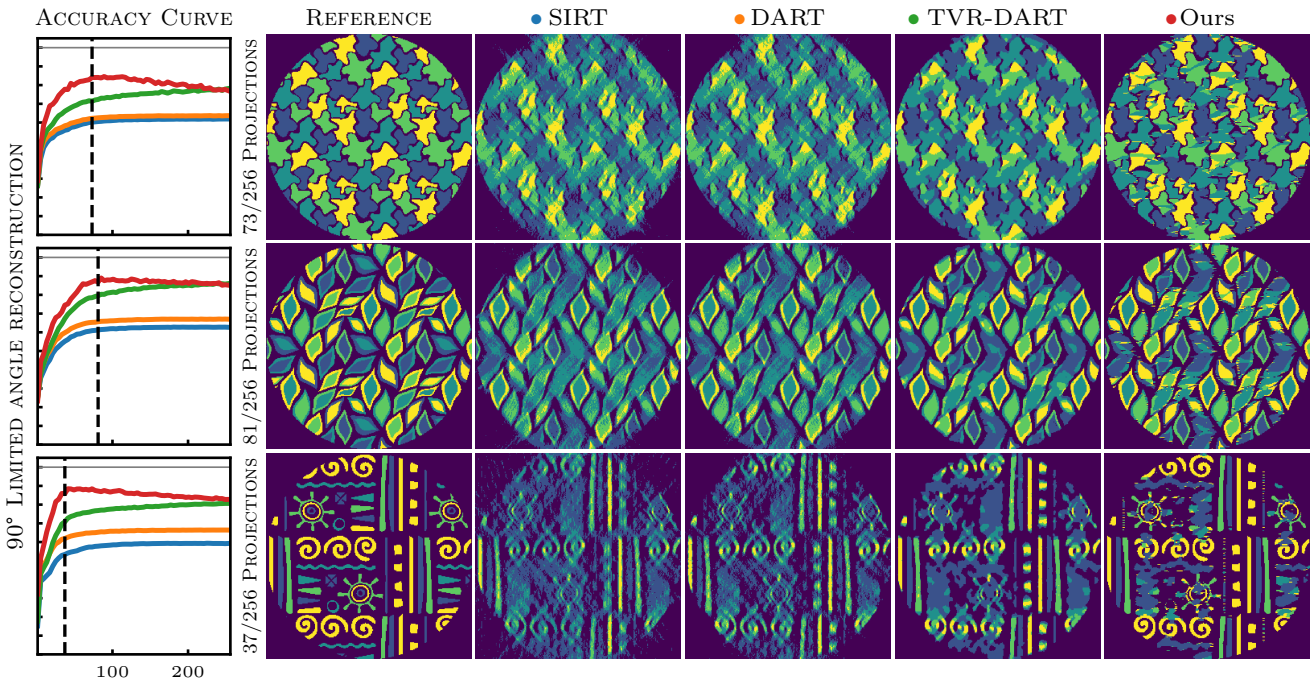


Fig. 8. **90-degree limited-angle 2D reconstruction**. We evaluate the reconstruction quality using projections covering only 90 degrees.

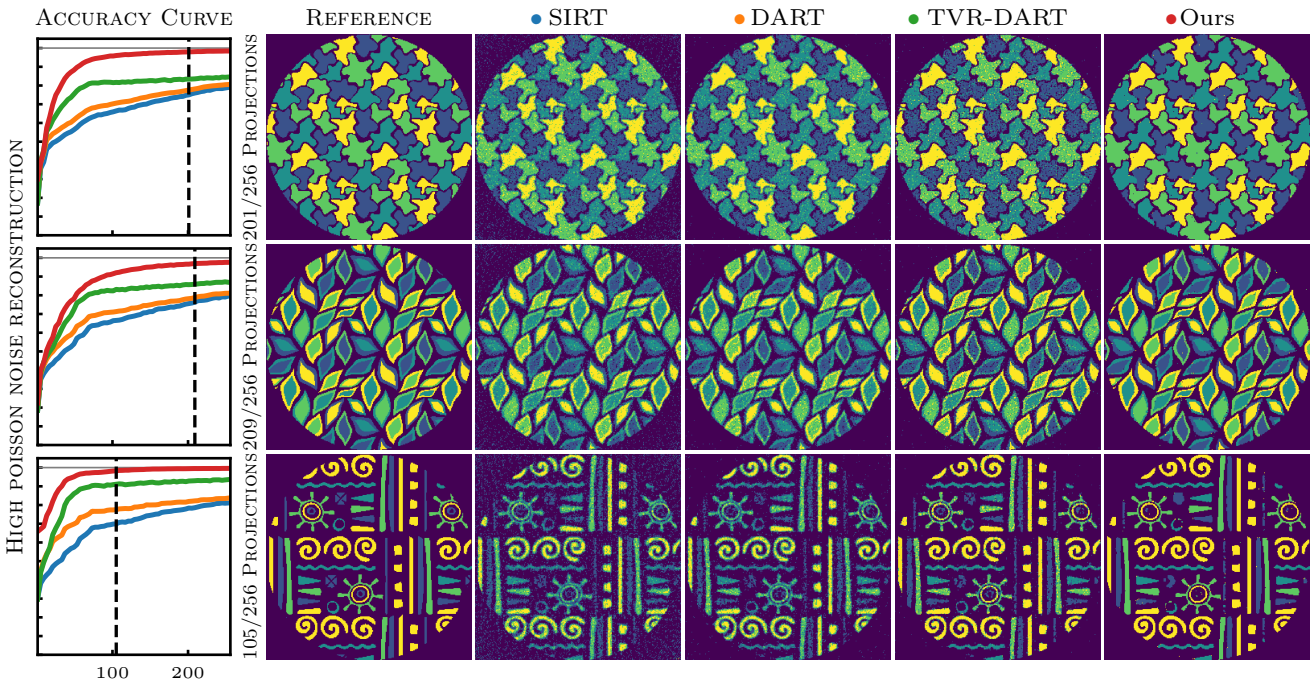


Fig. 9. **High-noise 2D reconstruction**. We evaluate the reconstruction quality in a high-noise scenario, where Poisson noise is added to the input projections.

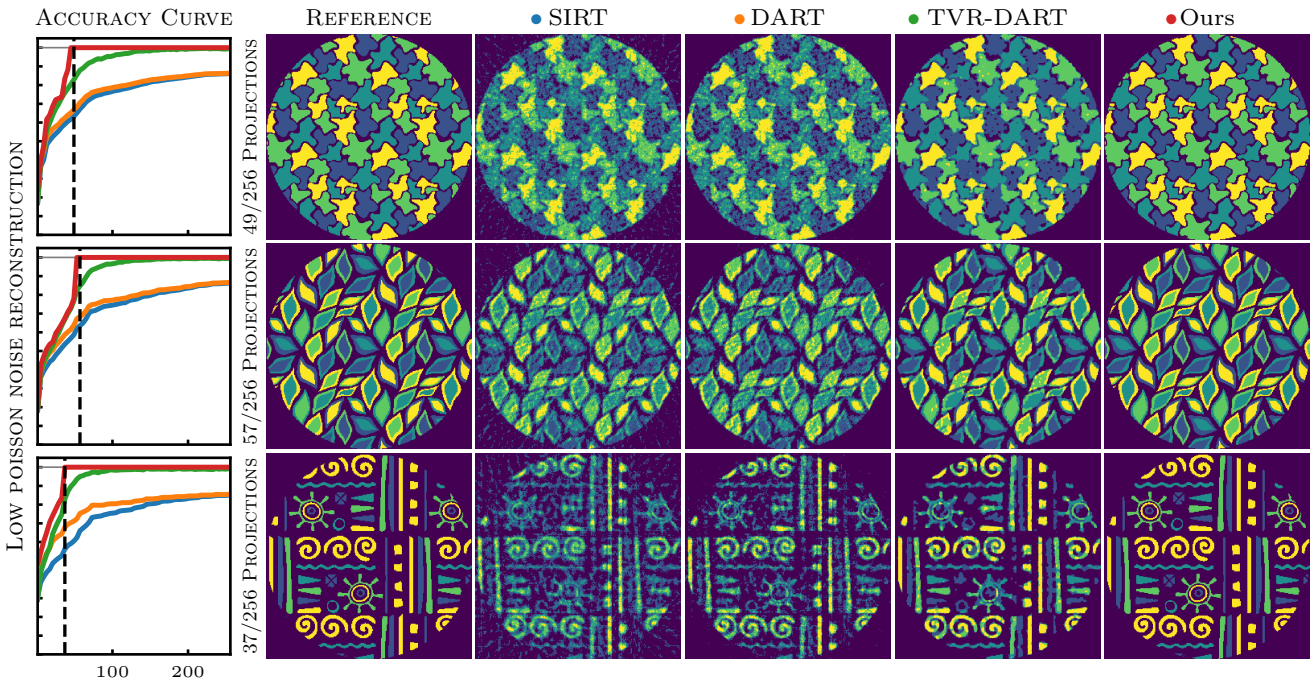


Fig. 10. **Low-noise 2D reconstruction**. This is the same scenario as in Figure 9, but with lower noise levels.