Supplementary Figures

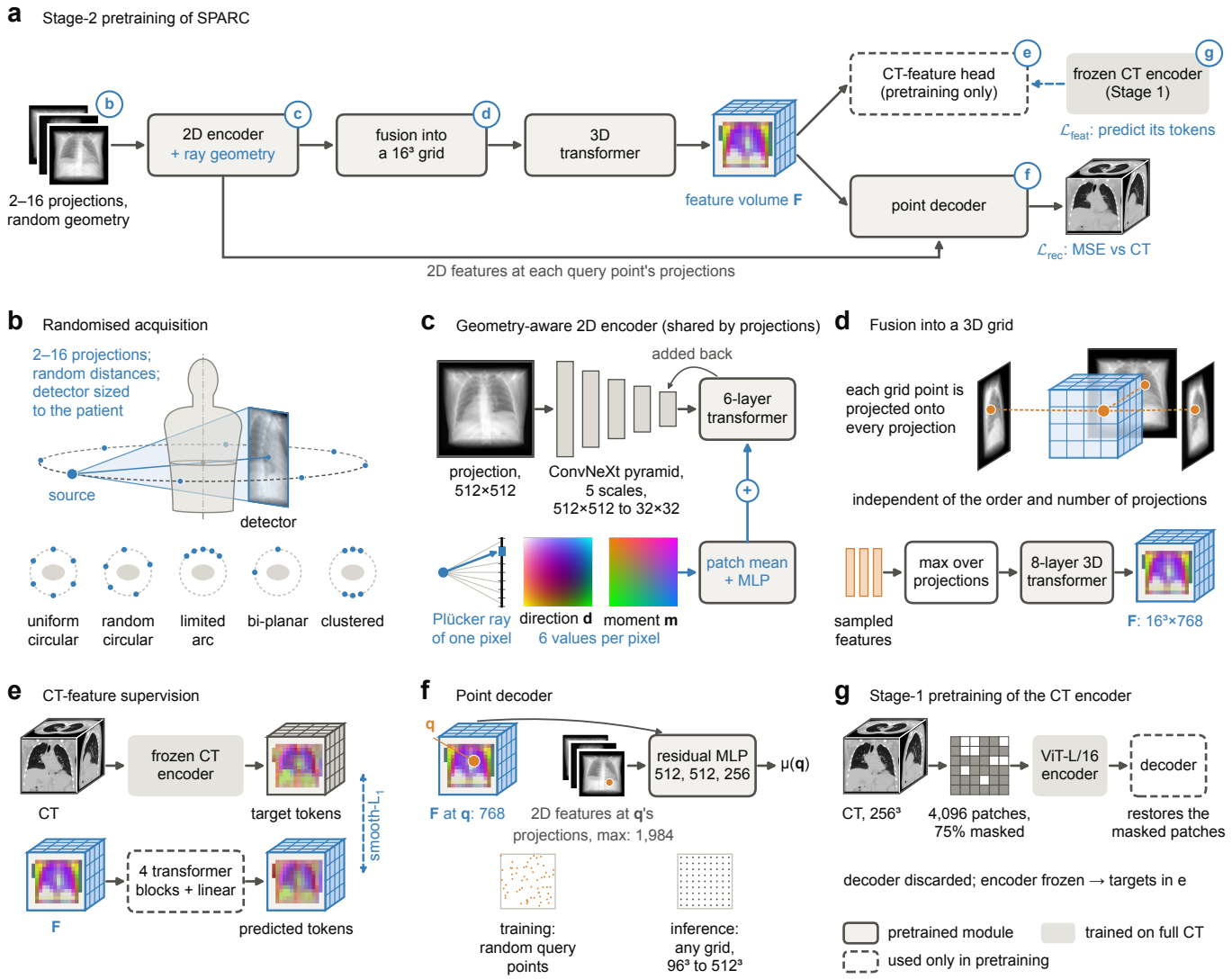
Title: A foundation model for three-dimensional anatomical reconstruction from sparse X-ray projections

Authors: Yiqun Lin, Jiayang Xu, Lie Ju, Jiarong Guo, Weiru Wang, Hualiang Wang, Huifeng Yao, Haoran Sun, Bin Pu, Kenli Li, Daniel C. Alexander, Hejie Cui, Yukun Zhou

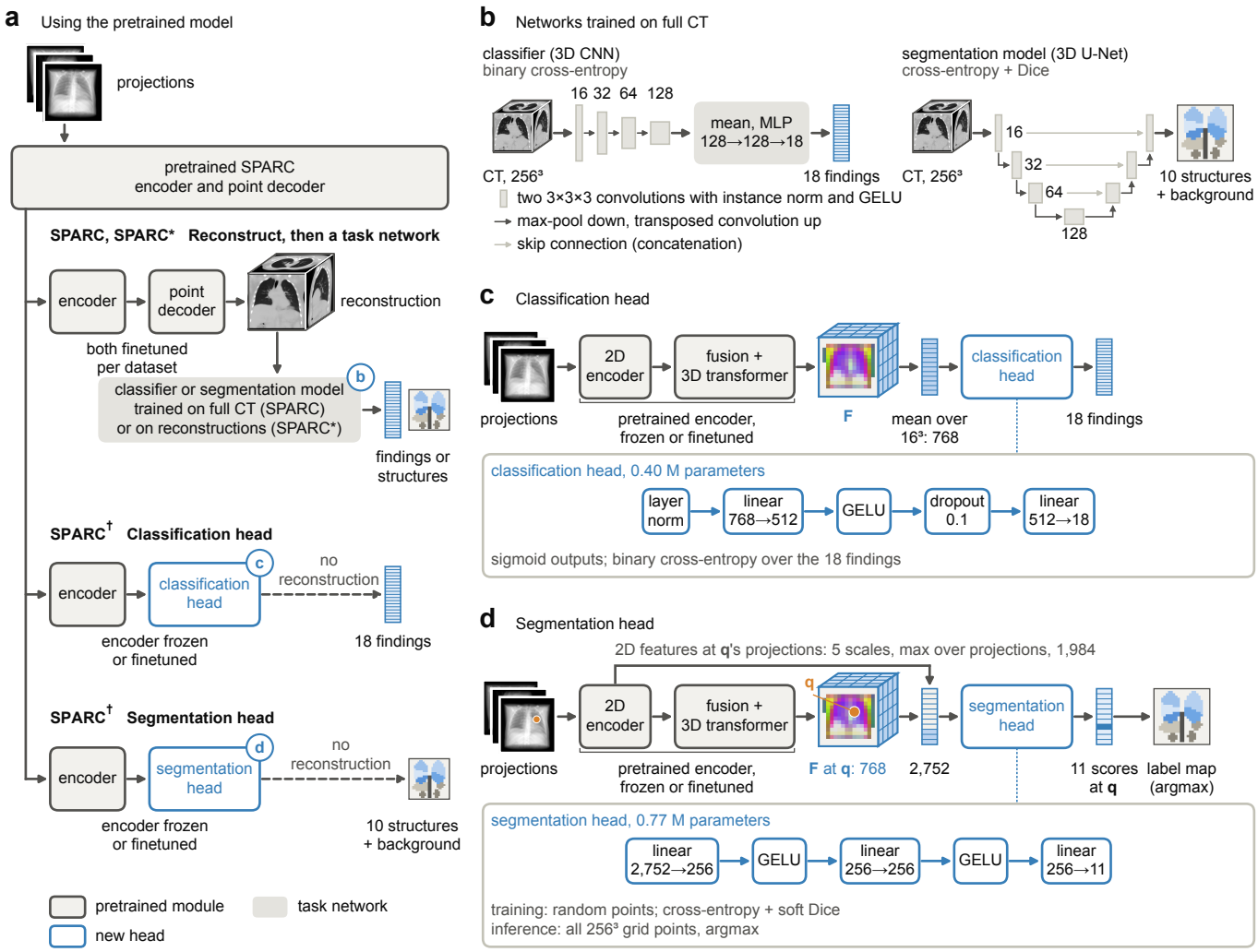
Contents. Supplementary Fig. 1–14. Unless stated otherwise, every statistic is a mean over the full test set with a 95% confidence interval from a bootstrap over patients (10,000 resamples); differences between methods use paired resampling of the same patients (Methods, Confidence intervals and statistical tests). PSNR and SSIM are computed over $[-1{,}024, 1{,}024]$ HU, scaled to $[0, 1]$.

List of Supplementary Figures

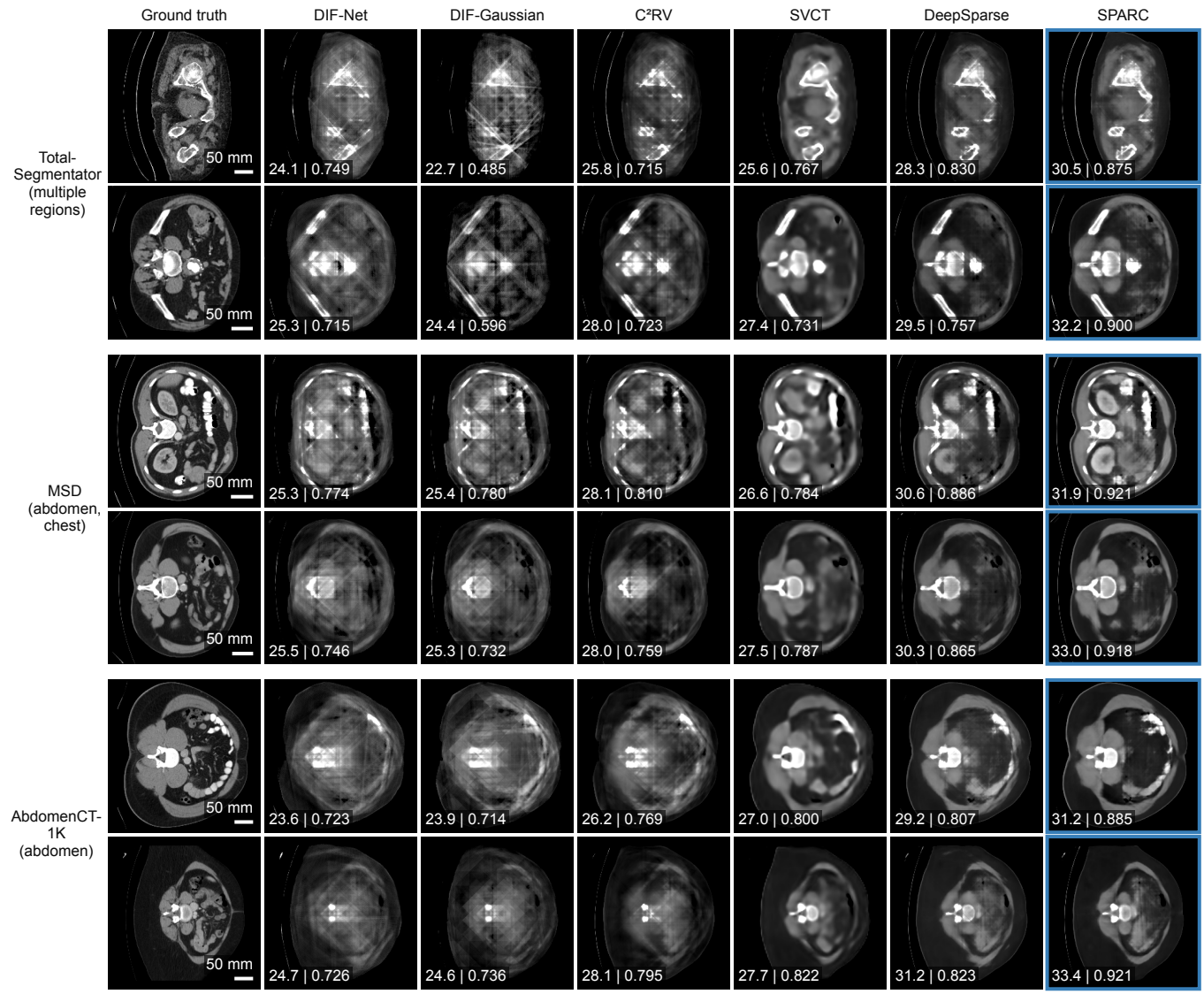
1 The SPARC model . . . 2
2 Using the pretrained model for the clinical tasks . . . 3
3 Reconstruction from eight projections: TotalSegmentator, MSD and AbdomenCT-1K . . . 4
4 Reconstruction fidelity with four projections . . . 5
5 Reconstruction from four projections: CT-RATE, TotalSegmentator, MSD and AbdomenCT-1K . . . 6
6 Reconstruction from four projections: AMOS, CQ500, ToothFairy3 and VerSe . . . 7
7 Where SPARC's reconstruction error lies . . . 8
8 Reconstruction time and memory . . . 9
9 Robustness to acquisition errors with eight projections . . . 10
10 Robustness to acquisition errors with four projections . . . 11
11 SPARC under acquisition errors . . . 12
12 Abnormality classification with four projections . . . 13
13 Other ways of using SPARC for the clinical tasks . . . 14
14 Multi-organ segmentation with four projections . . . 15



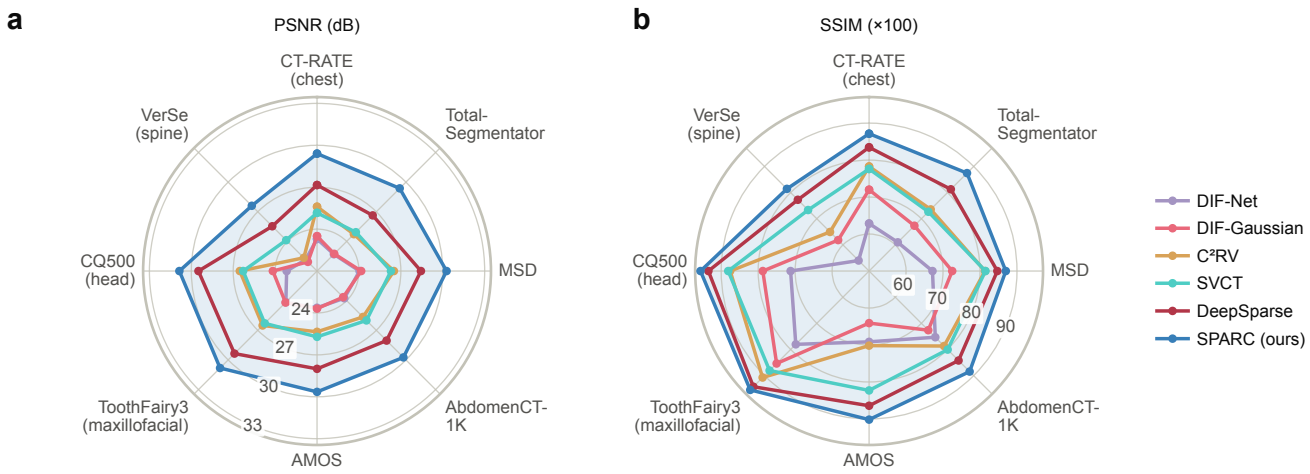
Supplementary Fig. 1 | The SPARC model. a. Stage-2 pretraining; circled letters mark the panels that show each part. **b.** Acquisition settings drawn at random for each training sample; the drawing shows the evaluation geometry, with eight sources on a circle around the patient. **c.** The 2D encoder, with the Plücker coordinates of the rays of one projection as colour maps, each component scaled for display. **d.** Fusion of the 2D features into the feature volume $\mathbf{F}$; the white box has no parameters. **e.** CT-feature supervision. The colour maps show a coronal section of each 16^3 grid: the first three principal components of the frozen CT encoder's tokens inside the body, the predicted tokens projected onto the same components, and the first three principal components of $\mathbf{F}$ rotated onto them. **f.** The point decoder, which predicts the attenuation $\mu(\mathbf{q})$ at a query point $\mathbf{q}$. **g.** Stage-1 pretraining of the CT encoder. The key at the bottom right applies to all panels. Projections, colour maps and volumes are from one CT-RATE test volume with eight projections. Settings are listed in Supplementary Tab. 32 and 33.



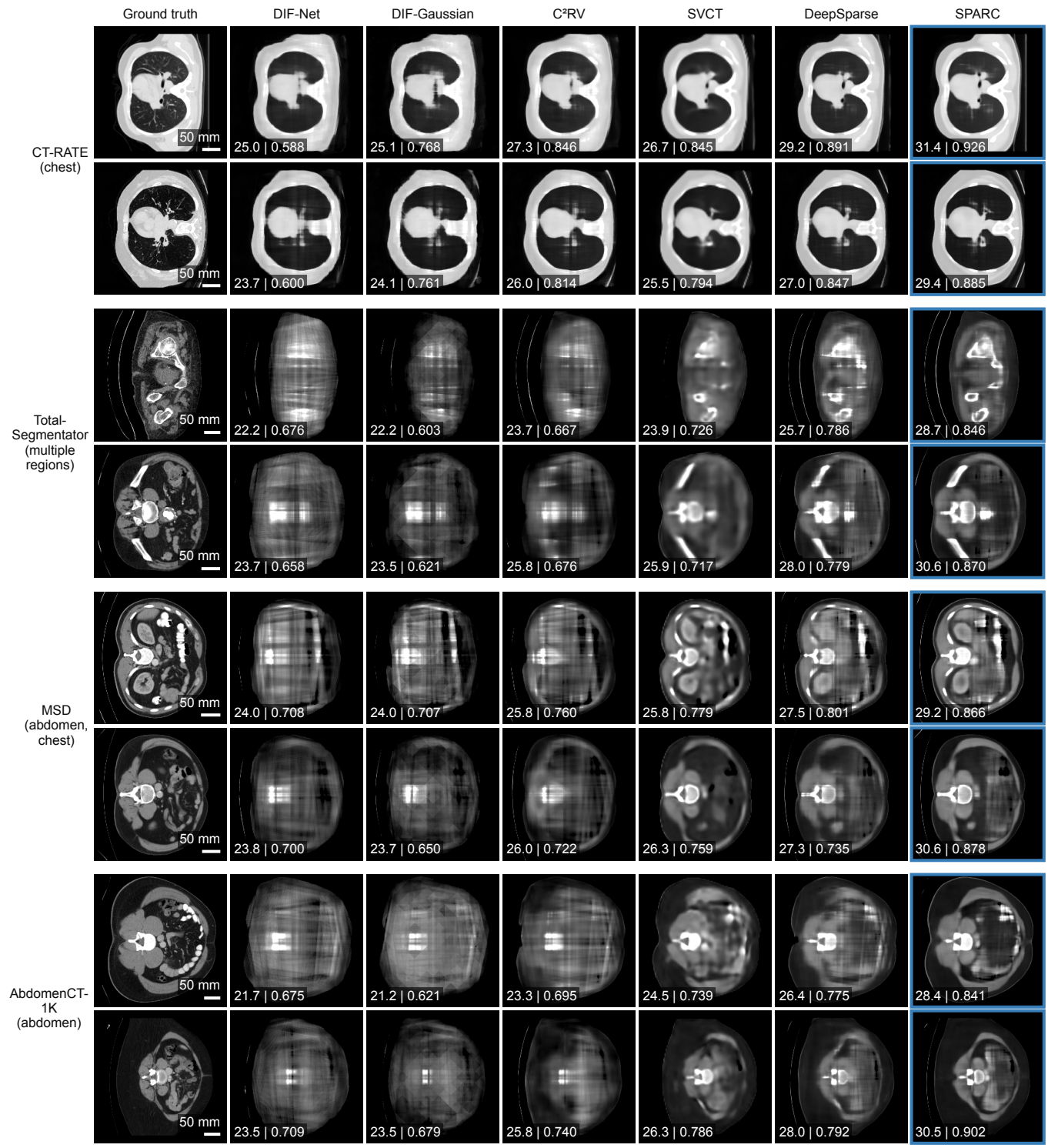
Supplementary Fig. 2 | Using the pretrained model for the clinical tasks. a. Three ways to use the pretrained SPARC: reconstruct, then give the reconstruction to a classifier or segmentation model trained on full CT (SPARC, the approach used throughout, as in Fig. 5 and 6) or to a network of the same architecture trained on SPARC's reconstructions (SPARC*); or predict the findings or structures from the projections with a head on the encoder, frozen or finetuned, without reconstruction (SPARC†). Circled letters mark the panels that show each part. **b.** The classifier and the segmentation model trained on full CT. **c.** The classification head. **d.** The segmentation head, which receives the same point features as the point decoder (Supplementary Fig. 1f). Projections, **F** and the CT and reconstruction volumes are from the CT-RATE test volume of Supplementary Fig. 1; the finding vectors and label maps are schematic. The training settings of the task networks and heads are listed in Supplementary Tab. 34.



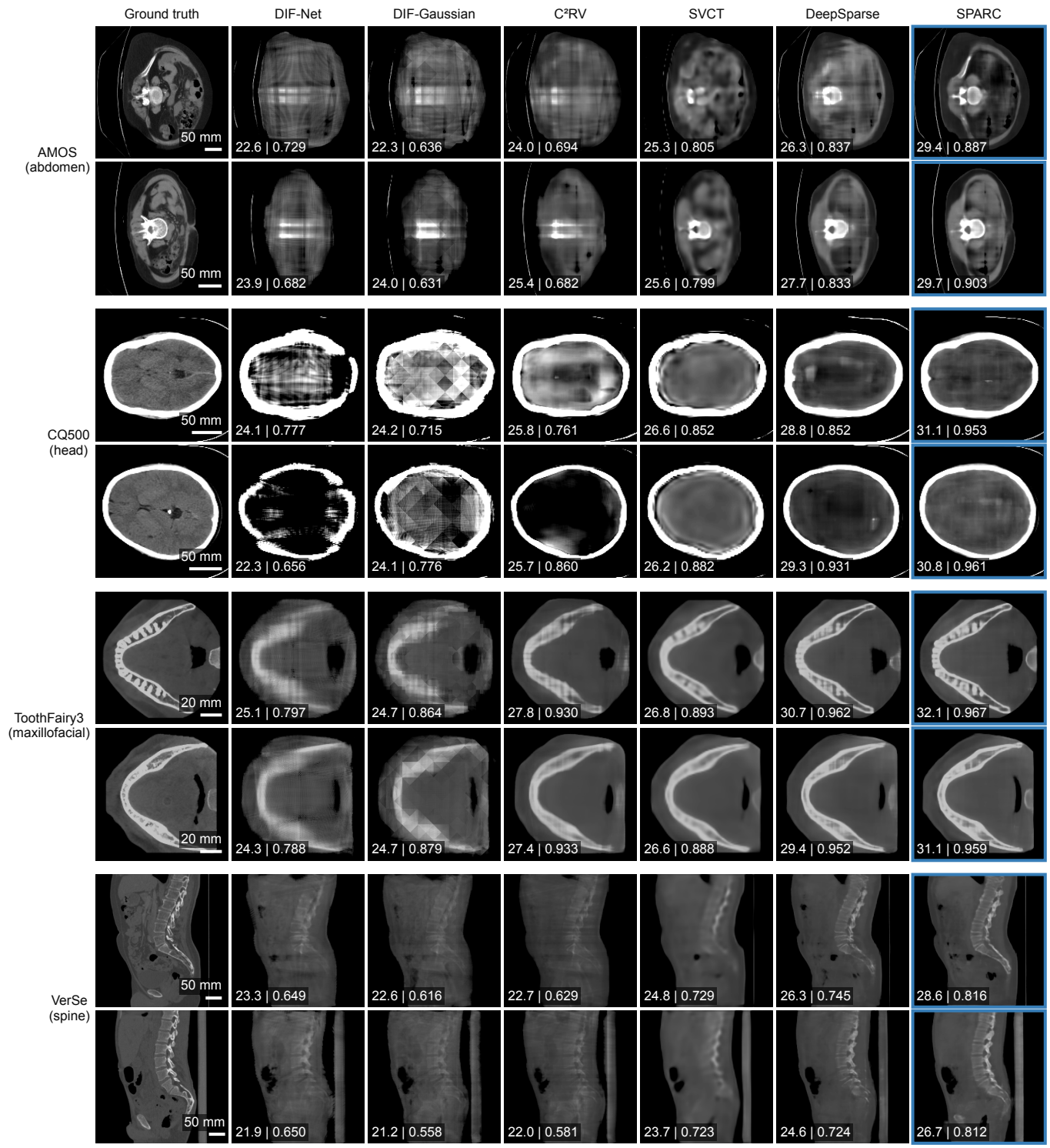
Supplementary Fig. 3 | Reconstruction from eight projections: TotalSegmentator, MSD and AbdomenCT-1K. Two test volumes from each of the three datasets not shown in Fig. 2c, reconstructed by every method with eight projections and drawn as in Fig. 2c. Display window $[-150, 250]$ HU. Labels give the volume's PSNR (dB) | SSIM; SPARC is outlined in blue, and each scale bar applies to its row. Values for the full test sets are in Supplementary Tab. 2.



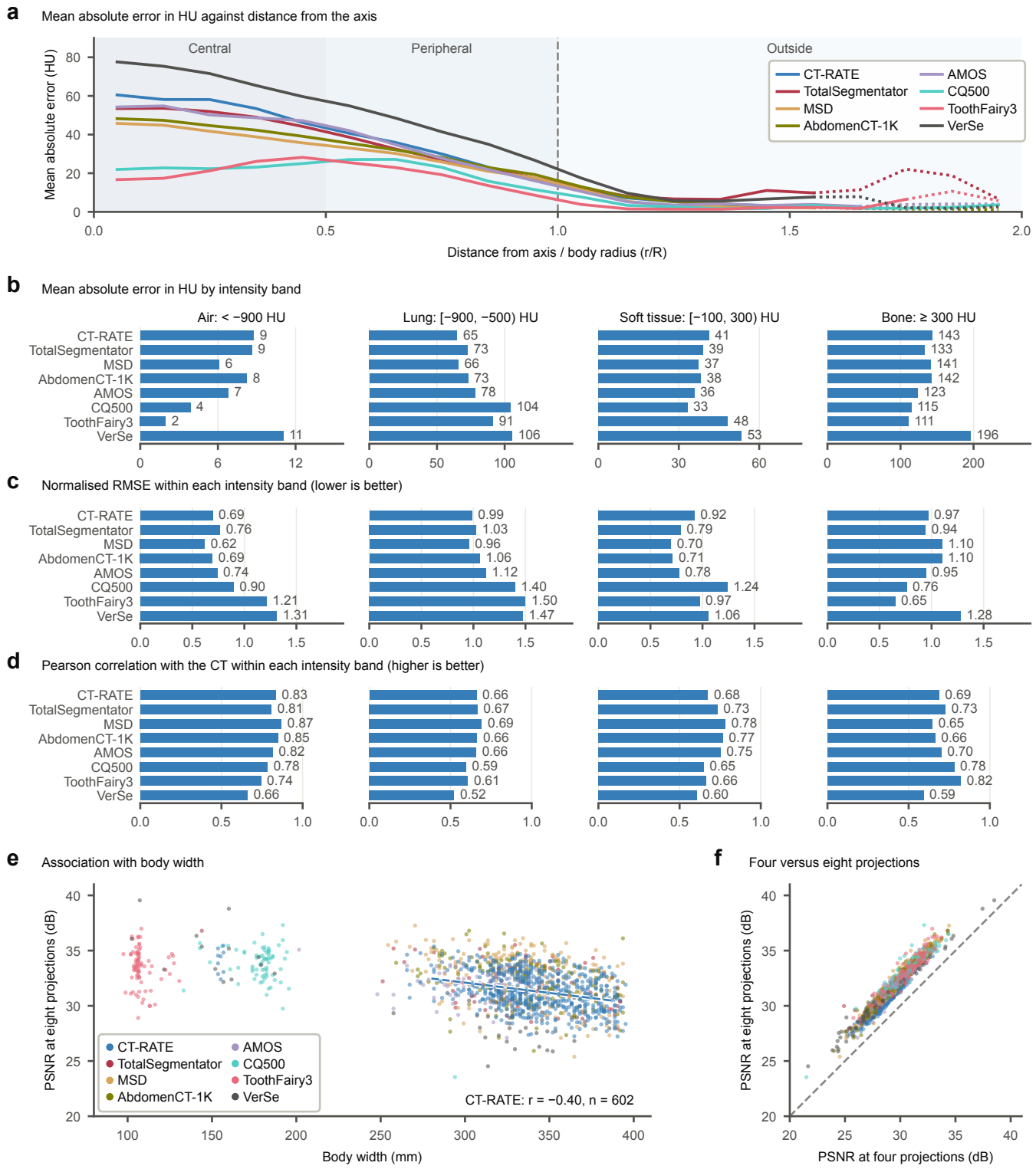
Supplementary Fig. 4 | Reconstruction fidelity with four projections. Mean PSNR (**a**) and SSIM (**b**) of each method on the eight test sets with four projections, drawn as in Fig. 2a,b and on the same radial axes, which start at 21 dB (**a**) and 50 (**b**). Values with 95% confidence intervals are in Supplementary Tab. 1, and paired differences between SPARC and the strongest competing method in Supplementary Tab. 3.



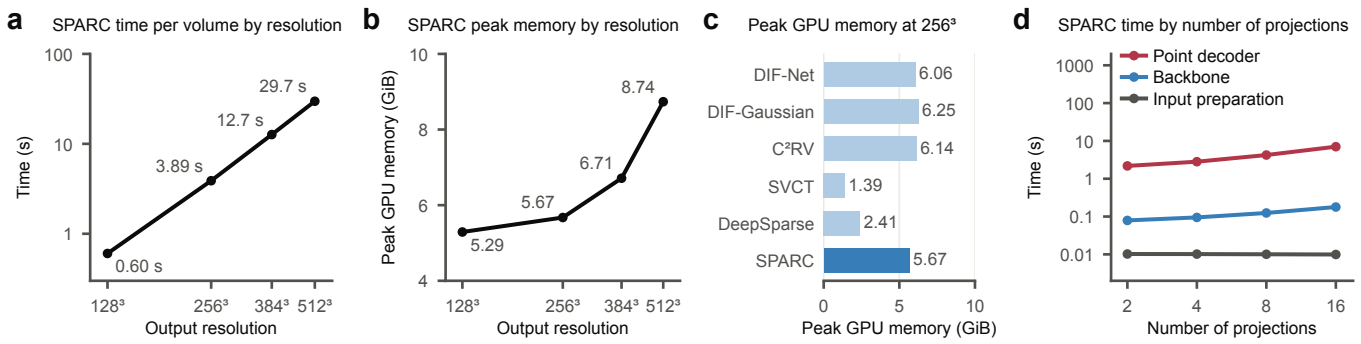
Supplementary Fig. 5 | Reconstruction from four projections: CT-RATE, TotalSegmentator, MSD and AbdomenCT-1K. Two test volumes from each dataset, reconstructed by every method with four projections and drawn as in Fig. 2c. The upper CT-RATE volume is that of Fig. 2c, and the TotalSegmentator, MSD and AbdomenCT-1K volumes are those of Supplementary Fig. 3, on the same sections. Display windows: CT-RATE $[-1{,}000, 200]$ HU; TotalSegmentator, MSD and AbdomenCT-1K $[-150, 250]$ HU. Labels give the volume's PSNR (dB) | SSIM; SPARC is outlined in blue, and each scale bar applies to its row. Values for the full test sets are in Supplementary Tab. 1.



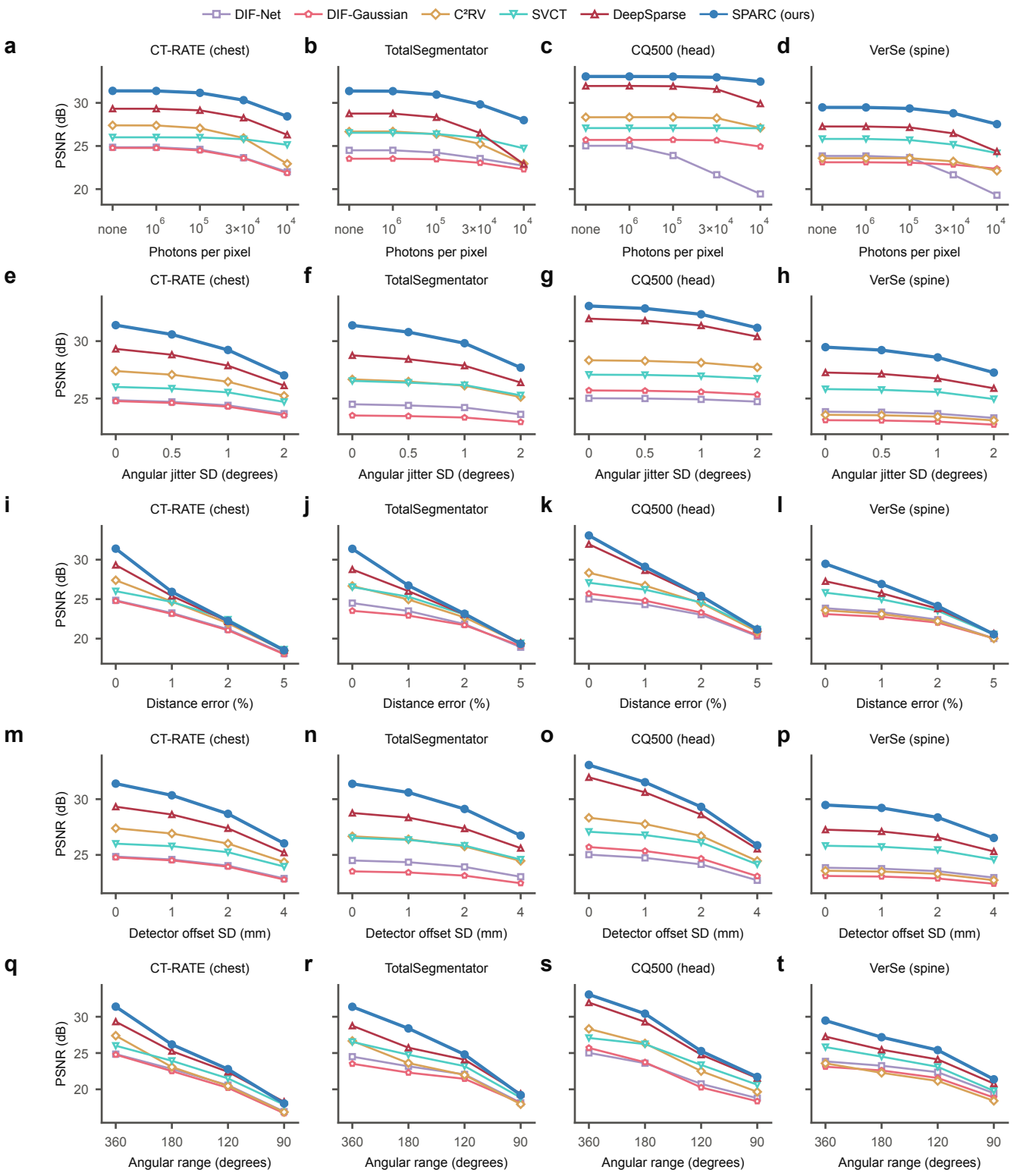
Supplementary Fig. 6 | Reconstruction from four projections: AMOS, CQ500, ToothFairy3 and VerSe. Two test volumes from each dataset, reconstructed by every method with four projections and drawn as in Fig. 2c. The upper volume of each dataset is that of Fig. 2c, on the same section. Display windows: AMOS $[-150, 250]$ HU; CQ500 $[0, 80]$ HU; ToothFairy3 $[-500, 1{,}500]$ HU; VerSe $[-500, 1{,}300]$ HU. Labels give the volume's PSNR (dB) | SSIM; SPARC is outlined in blue, and each scale bar applies to its row. Values for the full test sets are in Supplementary Tab. 1.



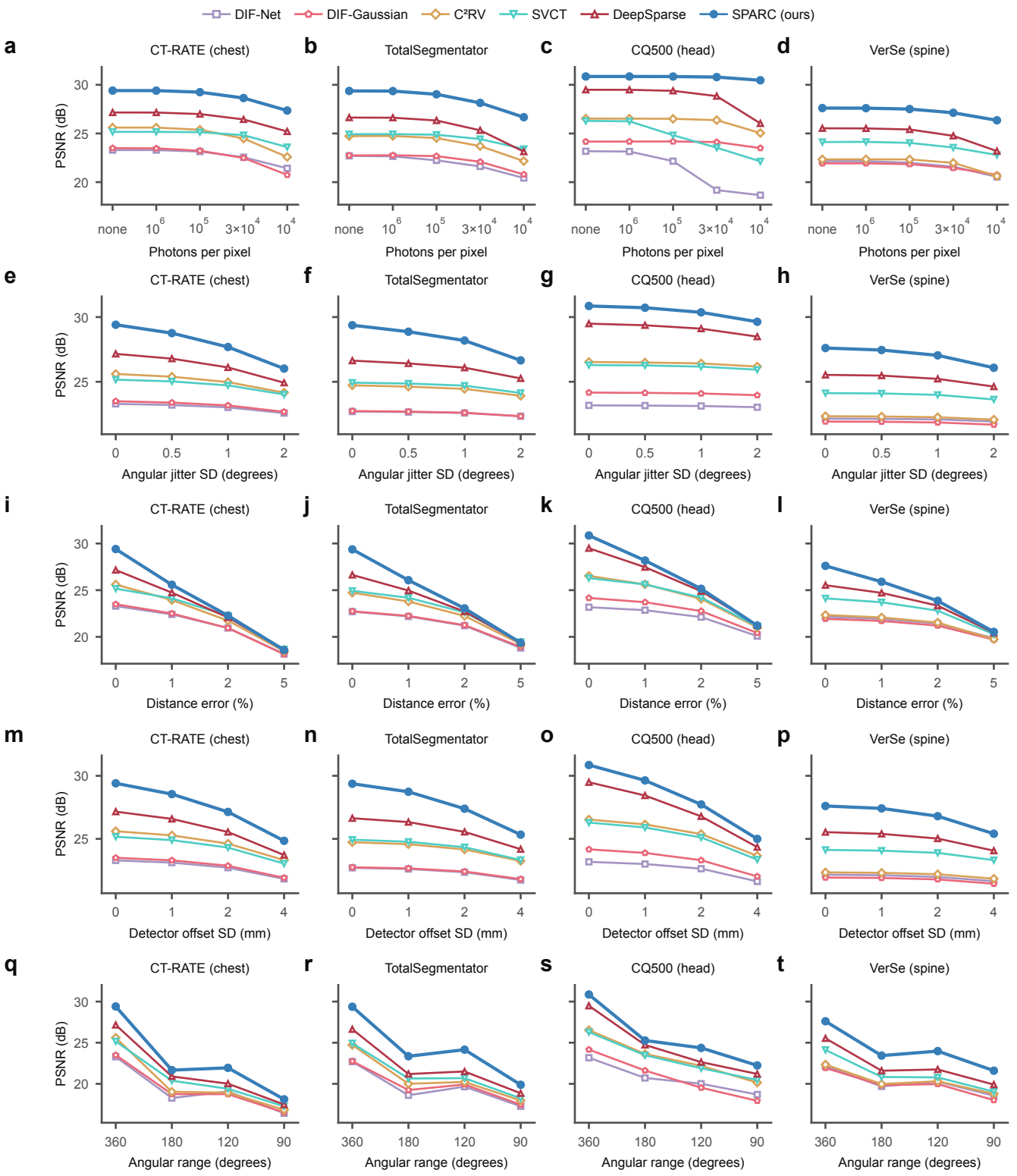
Supplementary Fig. 7 | Where SPARC's reconstruction error lies. SPARC with eight projections on the eight full test sets; **f** also uses four projections. A two-voxel rim is excluded, so PSNR values differ slightly from the main reconstruction results. **a.** Mean absolute error against the distance from the central axis of the volume, relative to the body radius (Methods); dotted segments mark radial bins reached by fewer than half of a dataset's test volumes. **b–d.** Mean absolute error (**b**), normalised RMSE (**c**) and Pearson correlation with the CT (**d**) within four intensity bands; bars show test-set means. **e.** PSNR of each volume against its body width; dashed line, linear fit to the CT-RATE volumes. **f.** PSNR of each volume with eight against four projections; dashed line, equal PSNR. Colours in **f** follow the key in **e**. Values with 95% confidence intervals are in Supplementary Tab. 6.



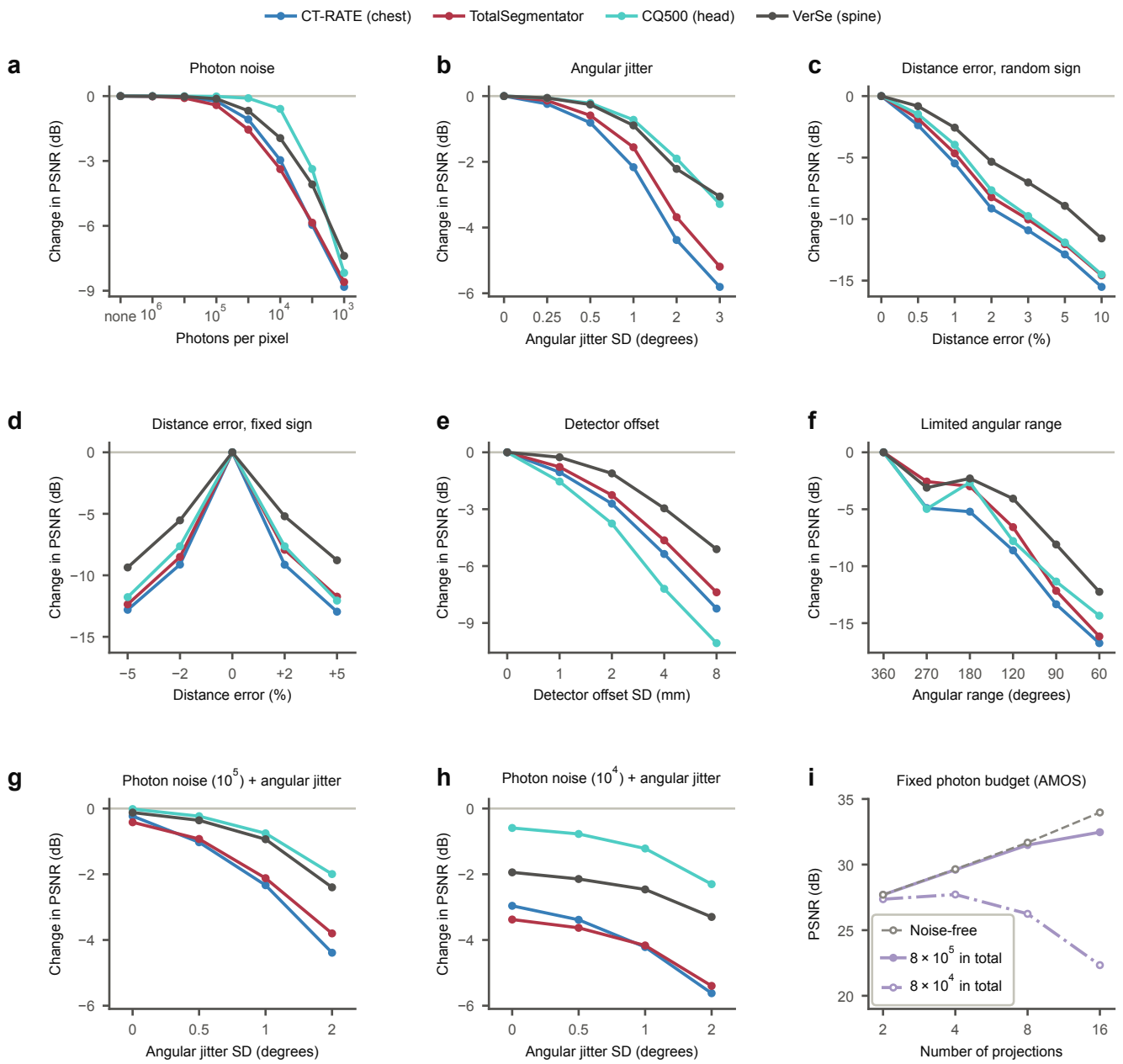
Supplementary Fig. 8 | Reconstruction time and memory. a, b. SPARC's reconstruction time per volume (**a**) and peak GPU memory (**b**) against the output resolution; the point decoder is resolution-free, so one trained model produces every grid. **c.** Peak GPU memory of every method at 256^3. **d.** Time of SPARC's three inference steps against the number of projections, for the models trained on AMOS with each number of projections. Panels **a–c** use the CT-RATE test set with eight projections; times are means over the test volumes. Accuracy against time for every method is in Fig. 3d. The values of **a–c** are in Supplementary Tab. 8, and those of **d** in Supplementary Tab. 9.



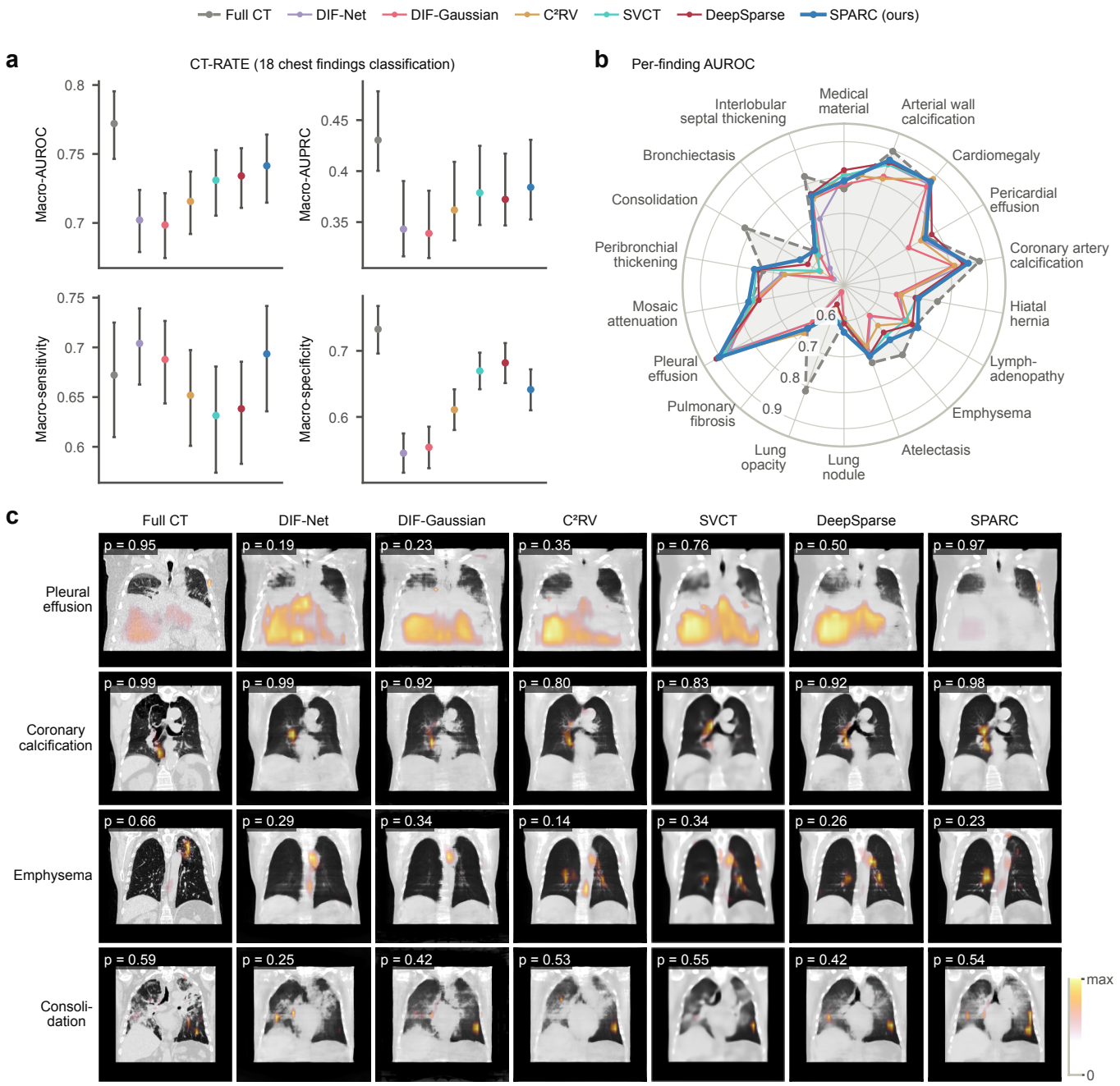
Supplementary Fig. 9 | Robustness to acquisition errors with eight projections. Mean PSNR of each method's eight-projection model on perturbed projections of the test sets, without retraining. **a–d.** Photon noise. **e–h.** Angular jitter of each projection. **i–l.** Source-to-detector distance error. **m–p.** Detector offset. **q–t.** Limited angular range, with the true angles given to the models. Panels in a row share the vertical axis; point estimates are shown. Panels **a–d** and **i–l** repeat Fig. 4a–d,i–l. Values and SPARC's paired difference from the strongest competing method under each condition are in Supplementary Tab. 11, and the SSIM in Supplementary Tab. 13.



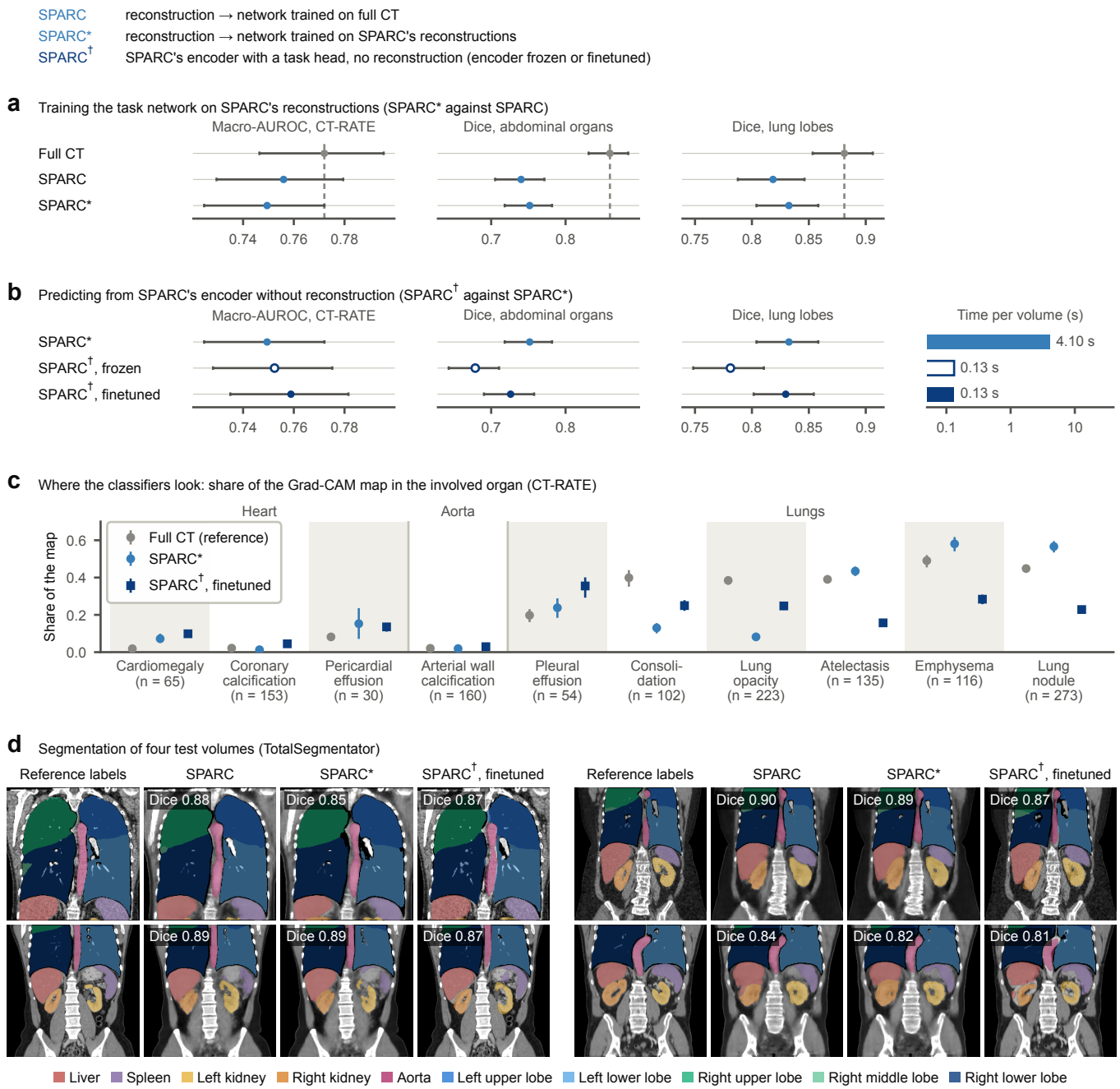
Supplementary Fig. 10 | Robustness to acquisition errors with four projections. Laid out as Supplementary Fig. 9, for the separately trained four-projection models: mean PSNR of each method on perturbed projections of the test sets, without retraining. **a–d.** Photon noise. **e–h.** Angular jitter of each projection. **i–l.** Source-to-detector distance error. **m–p.** Detector offset. **q–t.** Limited angular range, with the true angles given to the models. Panels in a row share the vertical axis; point estimates are shown. Values and SPARC's paired difference from the strongest competing method under each condition are in Supplementary Tab. 12, and the SSIM in Supplementary Tab. 14.



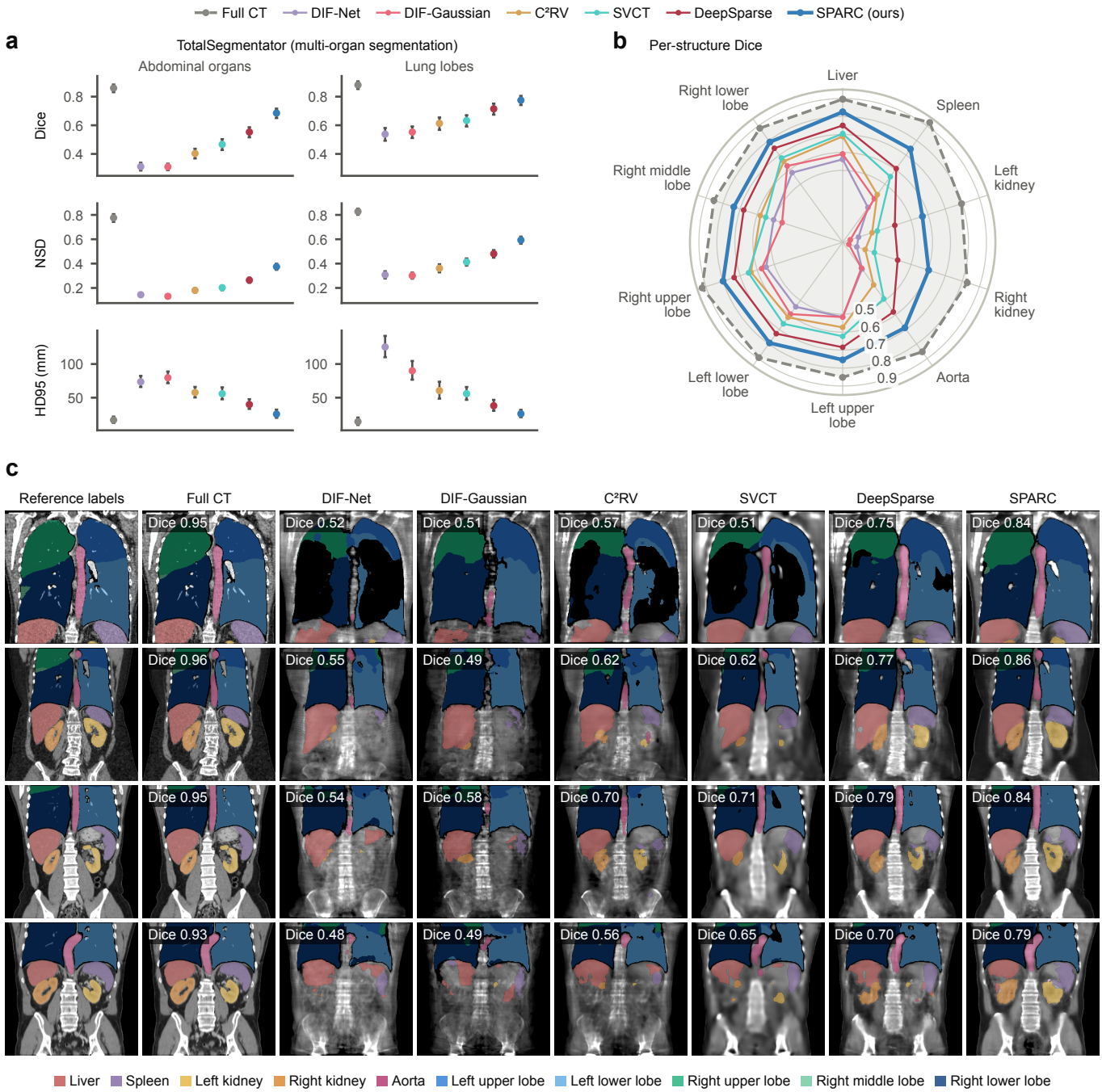
Supplementary Fig. 11 | SPARC under acquisition errors. SPARC alone on perturbed projections of the full test sets, without retraining; colours denote test sets, not methods. **a–h.** Change in PSNR from the unperturbed evaluation of the same eight-projection model on the same volumes, under one perturbation (**a–f**) or photon noise combined with angular jitter (**g, h**). In **d**, a positive error means that the true source-to-detector distance is longer than the one given to the model; in **f**, the model is given the true angles. **i.** PSNR on AMOS when a fixed total photon budget per detector pixel is divided among the projections; each number of projections has its own trained model, and the budgets are simulation parameters, not measured patient dose. Values with 95% confidence intervals are in Supplementary Tab. 15–17.



Supplementary Fig. 12 | Abnormality classification with four projections. Laid out as Fig. 5, for the separately trained four-projection models; each method's reconstructions were given unchanged to the same classifier trained on full CT. **a.** Macro-averaged metrics over the 18 findings on the CT-RATE test set. Error bars denote 95% confidence intervals. **b.** AUROC for each finding; dashed line, full CT; the radial axis starts at 0.5. **c.** Grad-CAM maps of one training seed of the classifier on full CT and on each reconstruction, for the volumes and sections of Fig. 5c; each map is scaled to its own maximum; p, the predicted probability of the finding. Values are in Supplementary Tab. 18–20, and the agreement of each map with the map on full CT in Supplementary Tab. 21.



Supplementary Fig. 13 | Other ways of using SPARC for the clinical tasks. Eight projections. SPARC, SPARC's reconstruction given to the network trained on full CT, as in Fig. 5 and 6; SPARC*, given to a network of the same architecture trained on SPARC's reconstructions; SPARC†, a task head on SPARC's pretrained encoder, frozen or finetuned, without reconstruction (Supplementary Fig. 2a). **a.** SPARC* against SPARC: macro-AUROC over the 18 findings on the CT-RATE test set (602 volumes) and mean Dice over the five abdominal organs and the five lung lobes on the TotalSegmentator test set (79 volumes); dashed line, the network trained on full CT applied to the full CT. **b.** SPARC† against SPARC*, on the axes of **a**, and the time of SPARC's model per volume for classification with a 256^3 output: with reconstruction (input preparation, backbone and point decoder) or with the encoder alone; the classifier applied after reconstruction is not timed. **c.** Share of the Grad-CAM map that lies in the involved organ, over the positive test volumes of each of the ten findings with an anatomical location (n, volumes), for the classifier trained on full CT on the CT, SPARC* and SPARC† with the finetuned encoder. **d.** Reference labels and the predictions of one training seed for the four volumes of Fig. 6c, with each volume's mean Dice; SPARC†'s prediction, with the finetuned encoder, is drawn on the CT. Points denote means and error bars 95% confidence intervals. Values are in Supplementary Tab. 9, 21 and 29–31.



Supplementary Fig. 14 | Multi-organ segmentation with four projections. Laid out as Fig. 6, for the separately trained four-projection models; each method's reconstructions were given unchanged to the same segmentation model trained on full CT. **a.** Dice, normalised surface Dice (NSD) and 95th-percentile Hausdorff distance (HD95) over the five abdominal organs and the five lung lobes. Error bars denote 95% confidence intervals. **b.** Dice for each structure; dashed line, full CT; the radial axis starts at 0.1 (0.2 in Fig. 6b). **c.** Reference labels and the predictions of one training seed of the segmentation model for the volumes and sections of Fig. 6c, with each volume's mean Dice over the ten structures. Values are in Supplementary Tab. 22, 23 and 25.