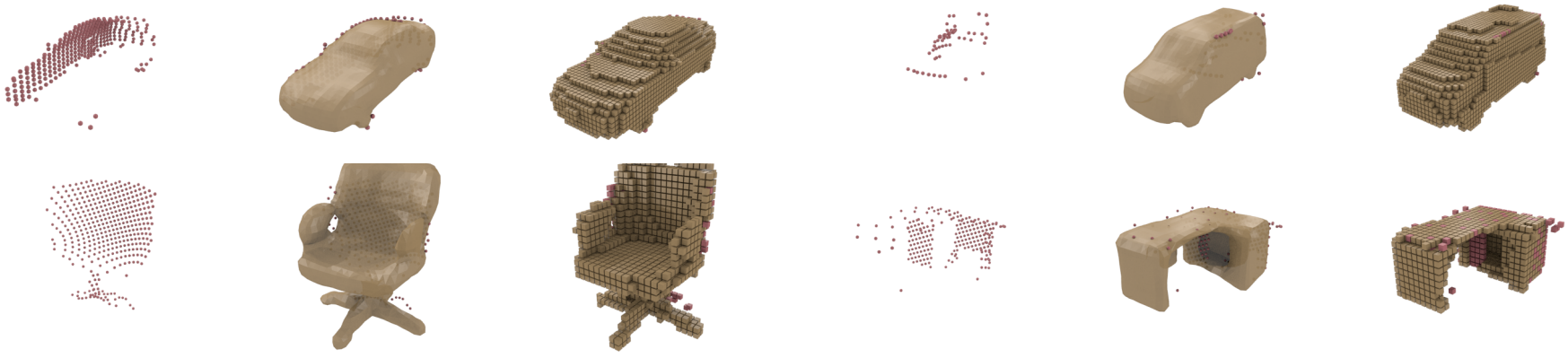
Learning Shape Completion from Bounding Boxes with CAD Shape Priors



David Stutz

David Stutz, Andreas Geiger[1]. *Learning 3D Shape Completion under Weak Supervision.* (in preparation for Deep Learning for Robotic Vision, IJCV).

---

[1]Autonomous Vision Group, MPI for Intelligent Systems and University of Tübingen

Outline

- Problem
- Related Work
- Proposed Approach
  - Shape Prior
  - Shape Inference
- Experiments
- Conclusion

Problem

Problem Definition

Problem

Given observations $\mathcal{X} = \{x_1, \ldots, x_N\}$ and reference shapes $\mathcal{Y} = \{y_1, \ldots, y_M\}$, learn a mapping $x_n \mapsto y(x_n)$ such that $y(x_n)$ fits the unknown target shape y_n^*.

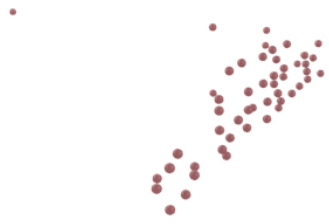
Observation x_n

Problem Definition

Problem

Given observations $\mathcal{X} = \{x_1, \ldots, x_N\}$ *and reference shapes* $\mathcal{Y} = \{y_1, \ldots, y_M\}$*, learn a mapping* $x_n \mapsto y(x_n)$ *such that* $y(x_n)$ *fits the unknown target shape* y_n^**.*

Reference Shapes y_m



Observation x_n

Problem Definition

Problem

Given observations $\mathcal{X} = \{x_1, \ldots, x_N\}$ *and reference shapes* $\mathcal{Y} = \{y_1, \ldots, y_M\}$*, learn a mapping* $x_n \mapsto y(x_n)$ *such that* $y(x_n)$ *fits the unknown target shape* y_n^**.*

Reference Shapes y_m

Observation x_n

Unknown Shape y_n^*

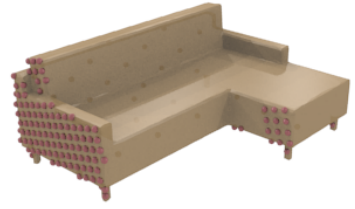
Problem Definition

Problem

Given observations $\mathcal{X} = \{x_1, \ldots, x_N\}$ *and reference shapes* $\mathcal{Y} = \{y_1, \ldots, y_M\}$*, learn a mapping* $x_n \mapsto y(x_n)$ *such that* $y(x_n)$ *fits the unknown target shape* y_n^**.*

Reference Shapes y_m

Observation x_n



Unknown Shape y_n^*

Problem Definition

Problem

Given observations $\mathcal{X} = \{x_1, \ldots, x_N\}$ and reference shapes $\mathcal{Y} = \{y_1, \ldots, y_M\}$, learn a mapping $x_n \mapsto y(x_n)$ such that $y(x_n)$ fits the unknown target shape y_n^*.

Weak supervision:

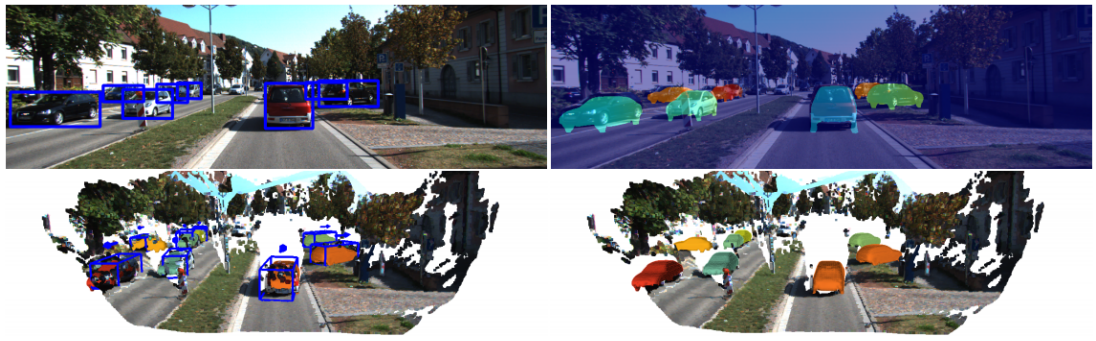
- known object categories;
- bounding boxes on real data.

Related Work

Selected Related Work

Data-driven shape completion, *e.g.*, [Engelmann et al., 2016]:

- Learn shape prior, *e.g.*, using PCA, and pose shape completion as optimization problem;
- applicable to real data, *e.g.*, KITTI, but slow inference.

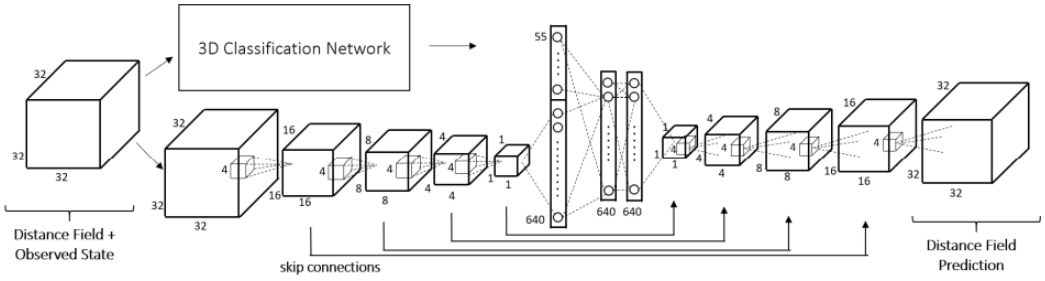


[Engelmann et al., 2016]

Selected Related Work

Learning-based shape completion, *e.g.*, [Dai et al., 2017]:

- Learn an encoder-decoder network on synthetic data and post-process if necessary;
- full supervision required, but efficient inference.



[Dai et al., 2017]

Contributions

We “unify” data-driven and learning-based approaches and show:

Hypothesis

Strong shape priors allow to learn shape completion under weak supervision.

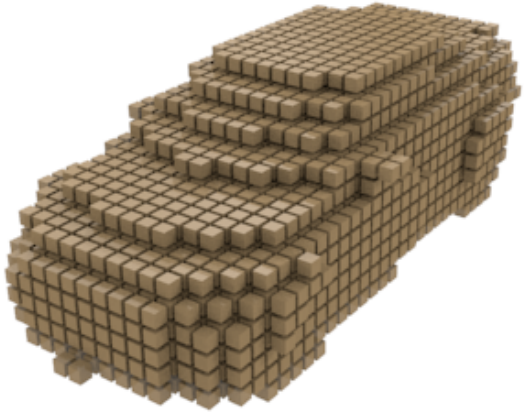
- Efficient shape completion;
- and learning on real data, *i.e.*, without supervision.

Proposed Approach

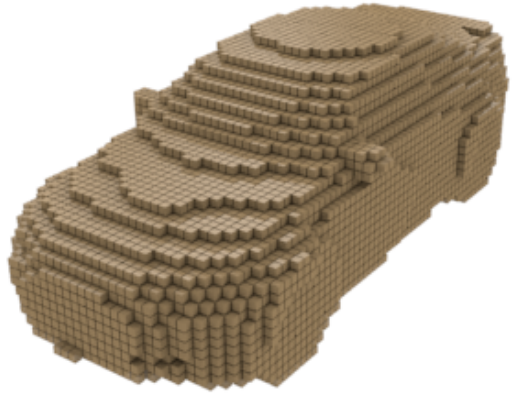
Shape Representation

Occupancy grids:

- Volumes $y \in \{0,1\}^{H \times W \times D}$ with $y_i = 1$ iff voxel i is occupied, *i.e.*, inside the shape or on its surface.



$24 \times 54 \times 24$



$48 \times 108 \times 48$

Shape Representation

Signed distance functions (SDFs):

- Volumes $y \in \mathbb{R}^{H \times W \times D}$ with y_i being the distance of voxel i's center to the surface; negative distances indicate occupied voxels.

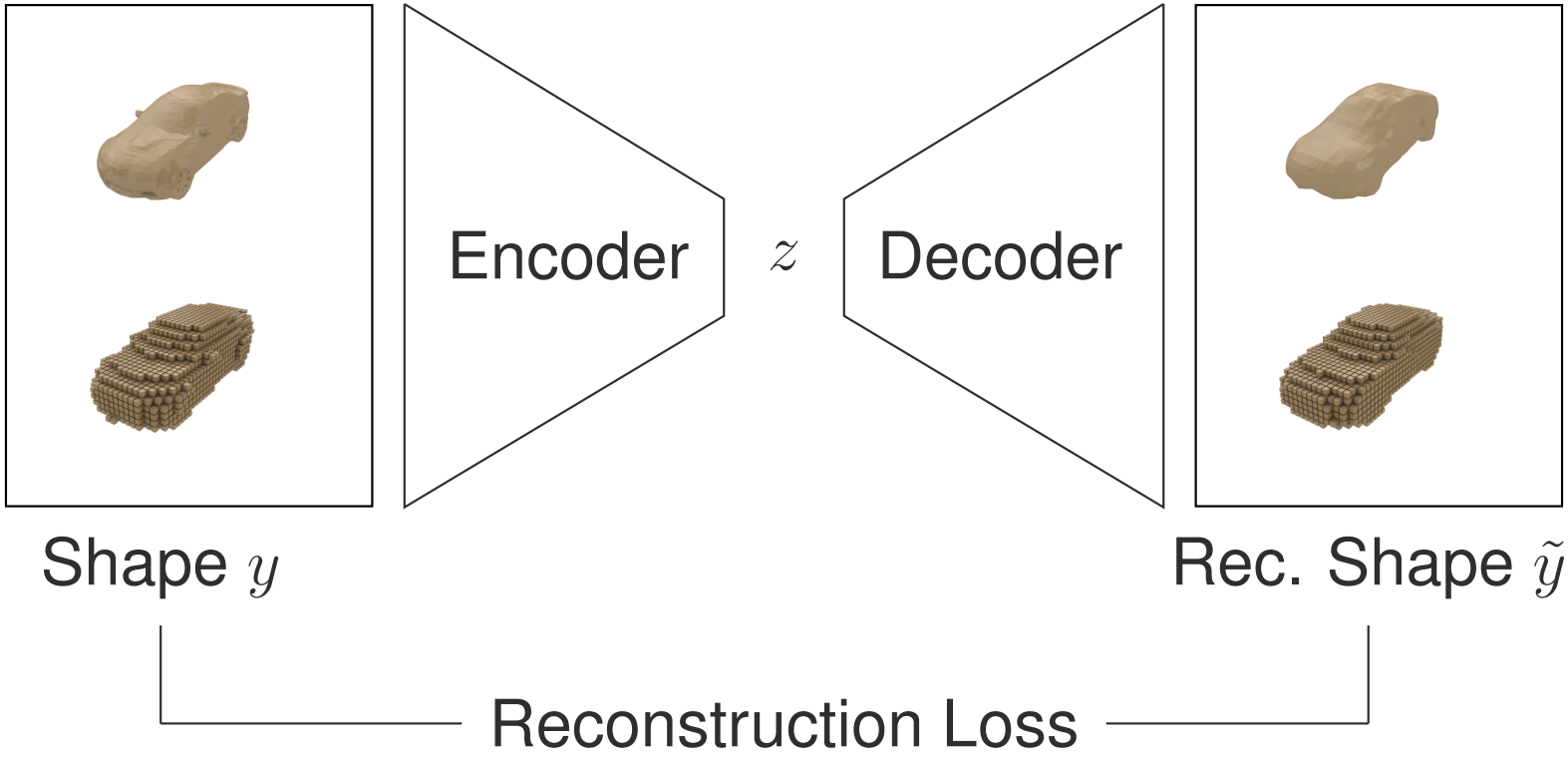
Original

$24 \times 54 \times 24$

$48 \times 108 \times 48$

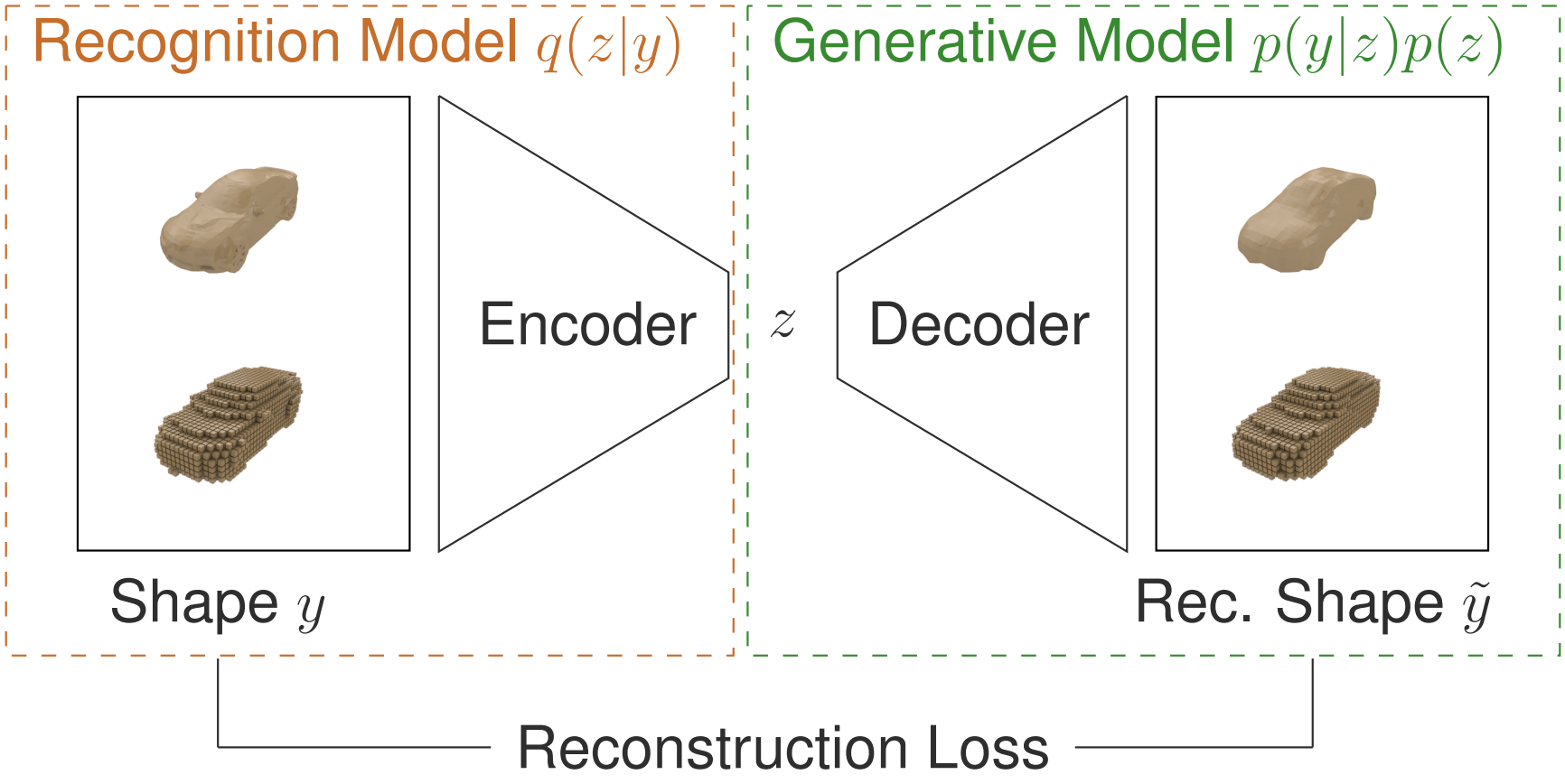
Shape Prior

Learn a variational auto-encoder:



Shape Prior

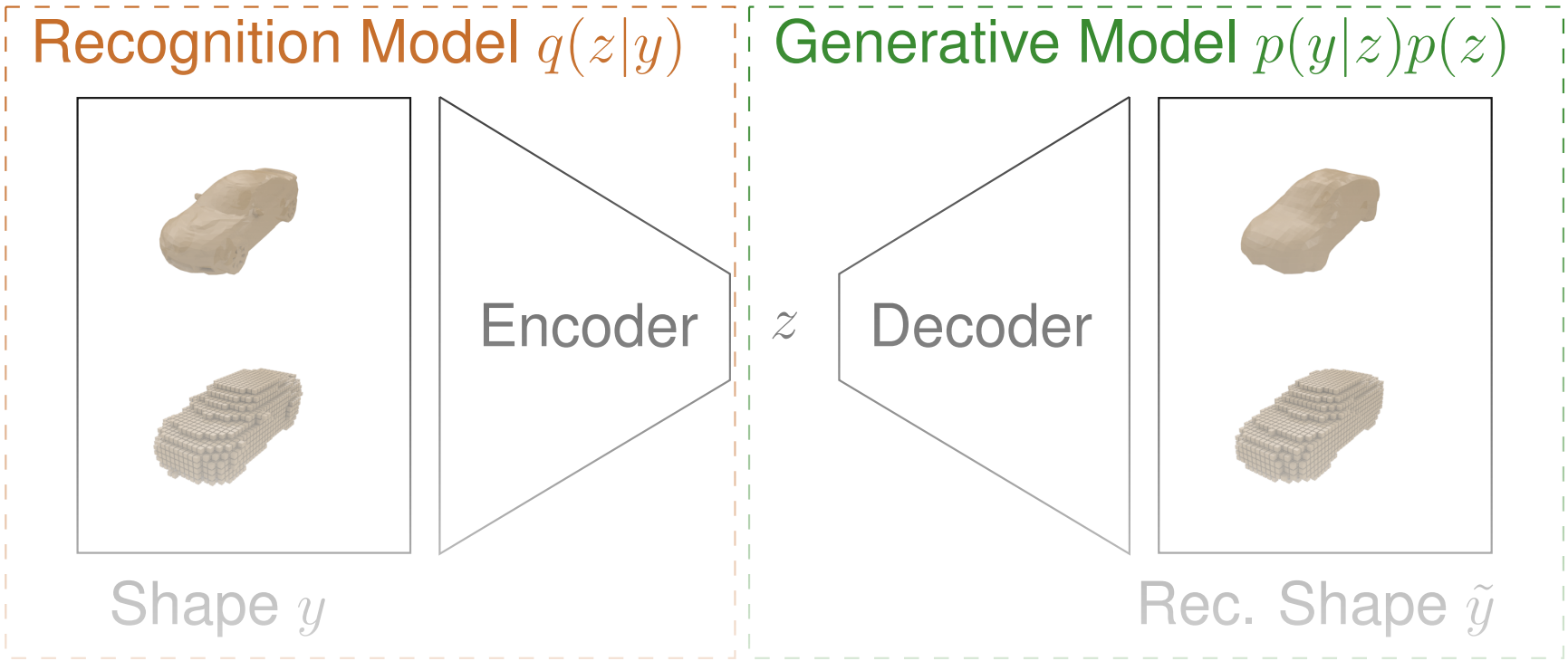
Learn a variational auto-encoder:



Shape Prior

Learn a variational auto-encoder:

$$p(z) = \mathcal{N}(z|0, I)$$

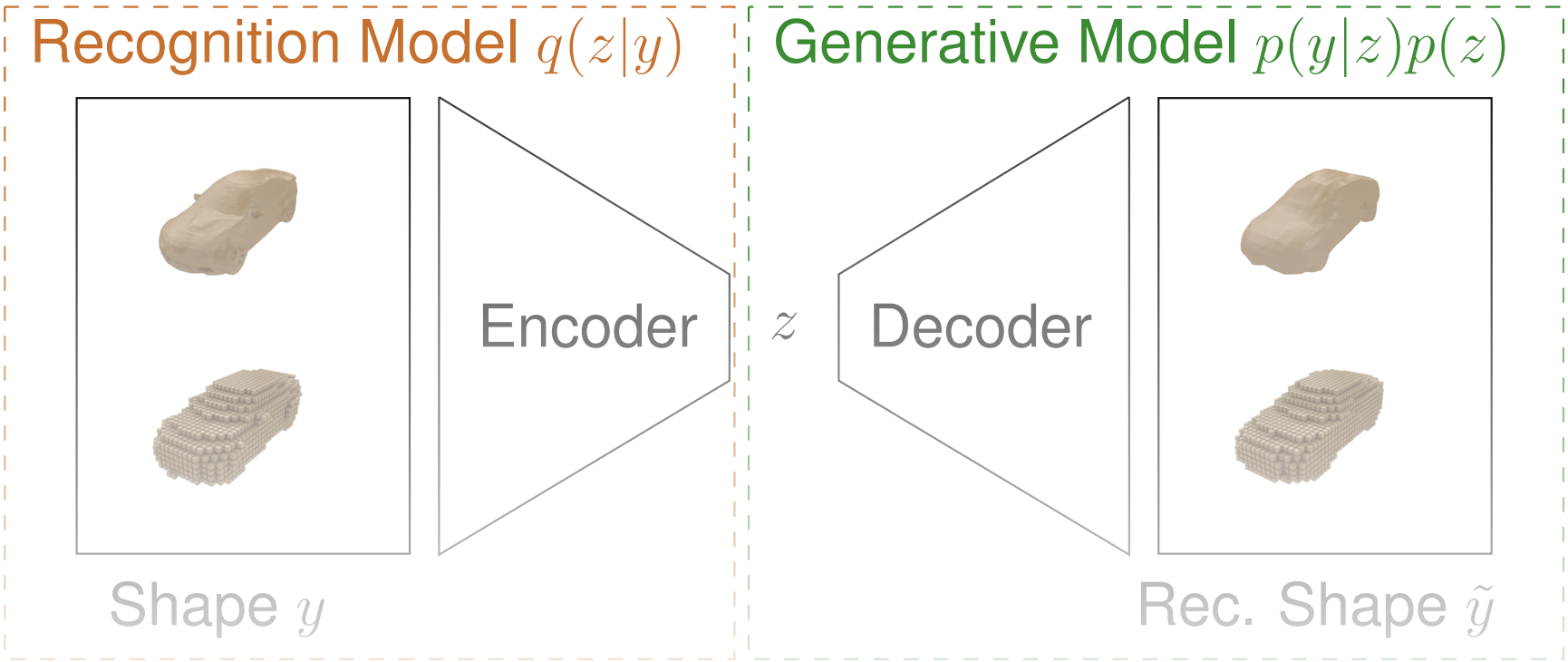


Shape Prior

Learn a variational auto-encoder:

$$q(z|y) = \mathcal{N}(z|\mu(y;w), \text{diag}(\sigma^2(y;w)))$$

encoder neural network

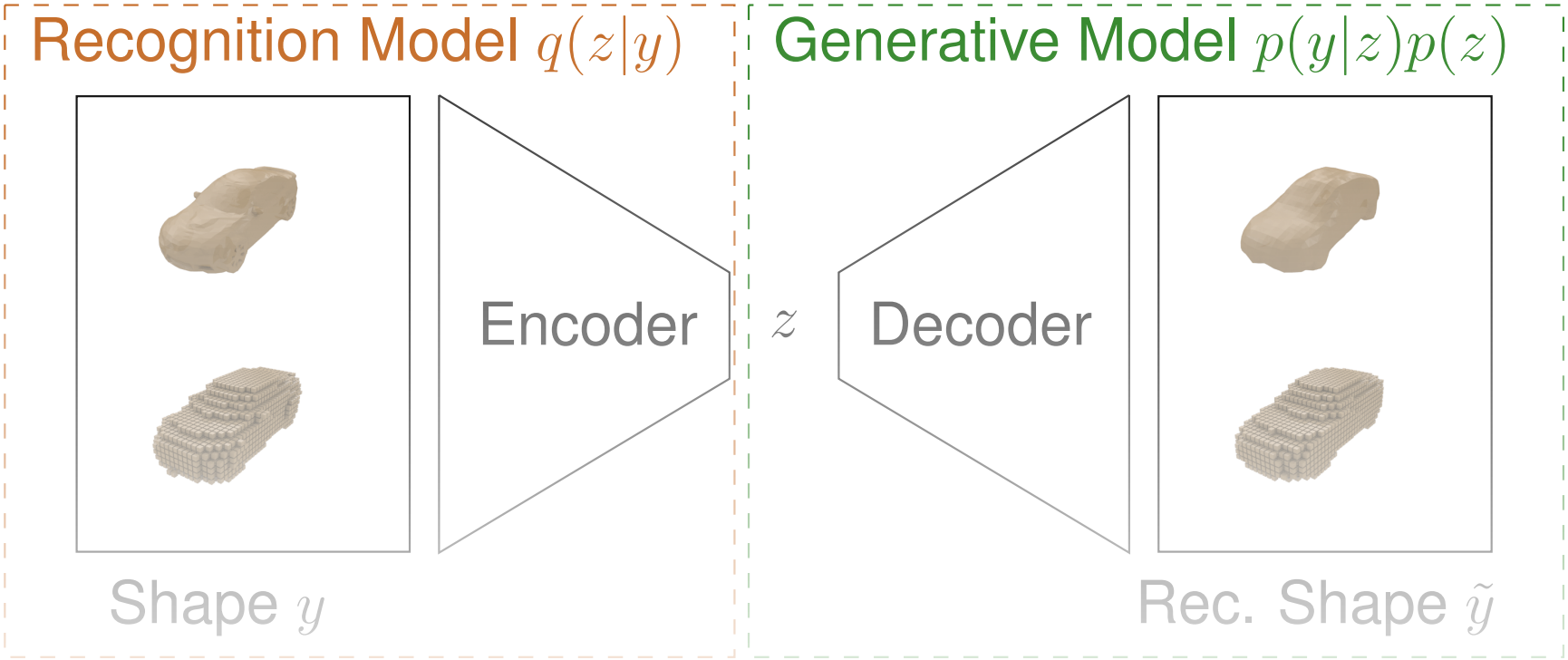


Shape Prior

Learn a variational auto-encoder:

decoder neural network

$$p(y_i|z) = \begin{cases} \text{Ber}(y_i|\theta_i(z;w)) & \text{occupancy grids} \\ \mathcal{N}(y_i|\mu_i(z;w),\sigma^2) & \text{SDFs} \end{cases}$$

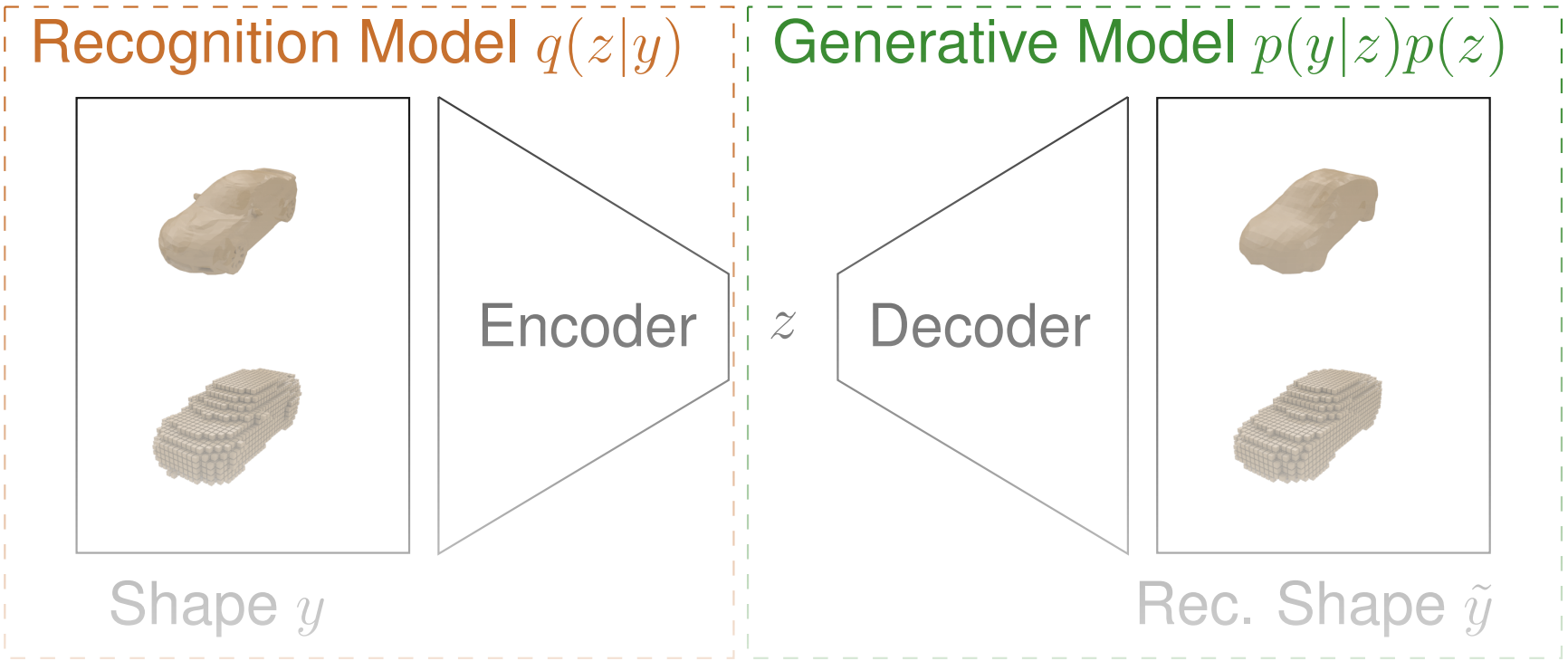


Shape Prior

Learn a variational auto-encoder:

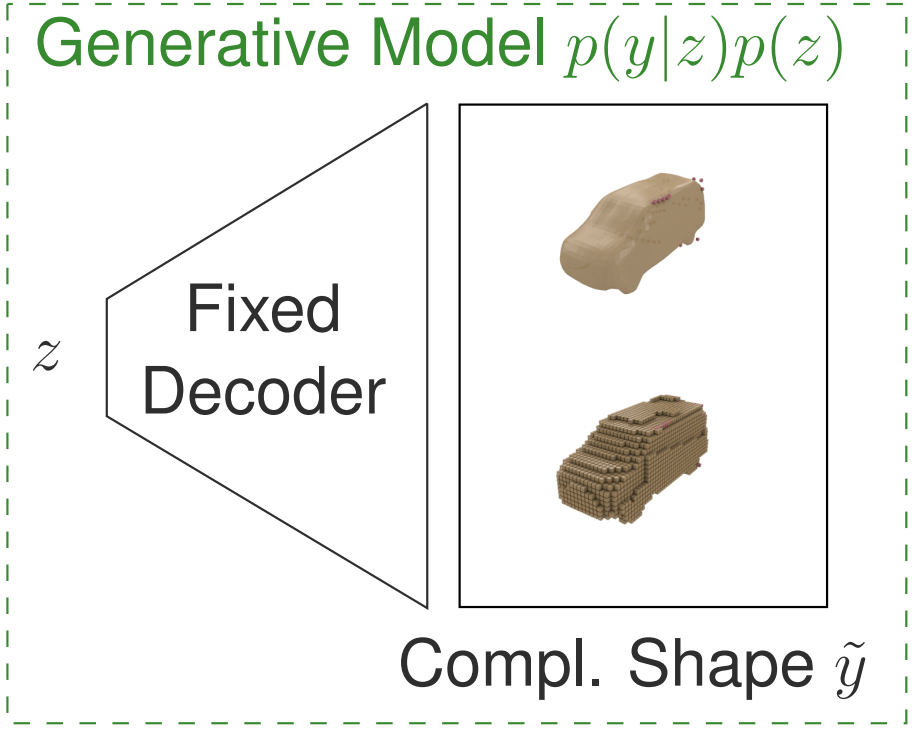
$$\mathcal{L}_{\text{VAE}}(w) = \mathbb{E}_{q(z|y)}[-\log p(y|z)] + \text{KL}(q(z|y)|p(z))$$

BCE or SSE loss



Shape Inference

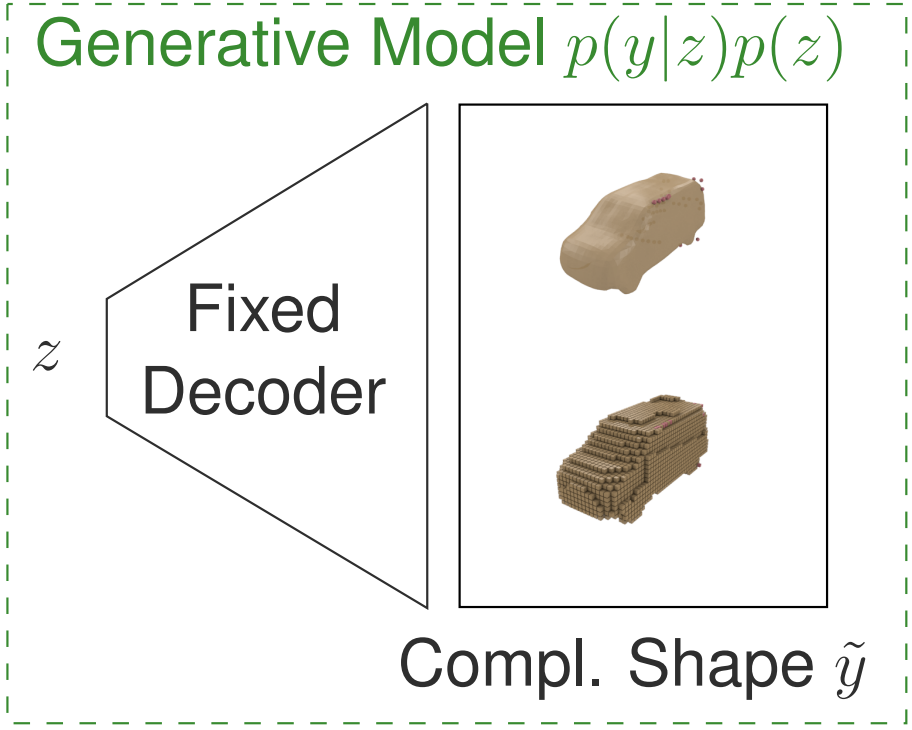
Perform maximum likelihood:



Shape Inference

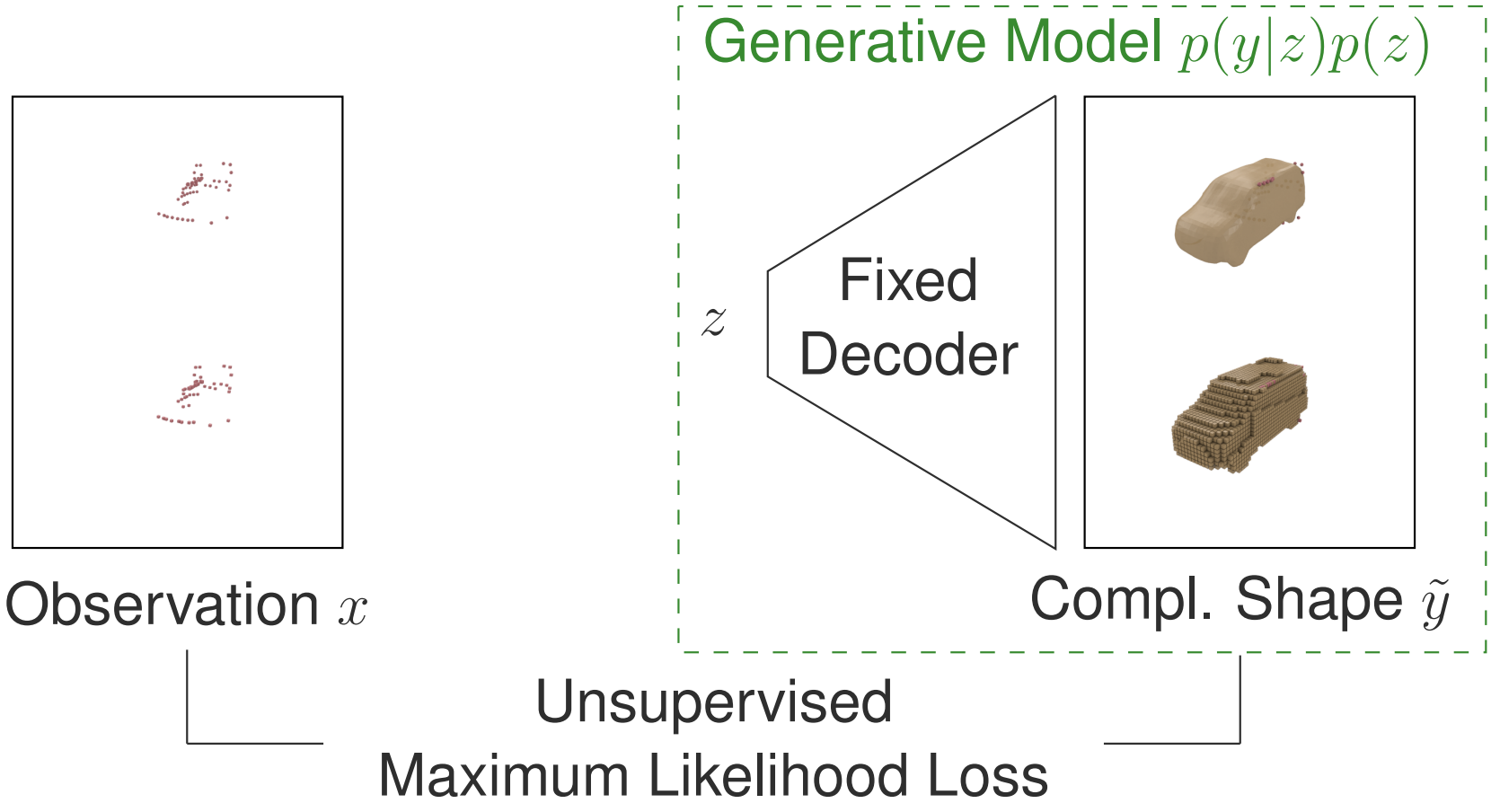
Perform maximum likelihood:

Observation x



Shape Inference

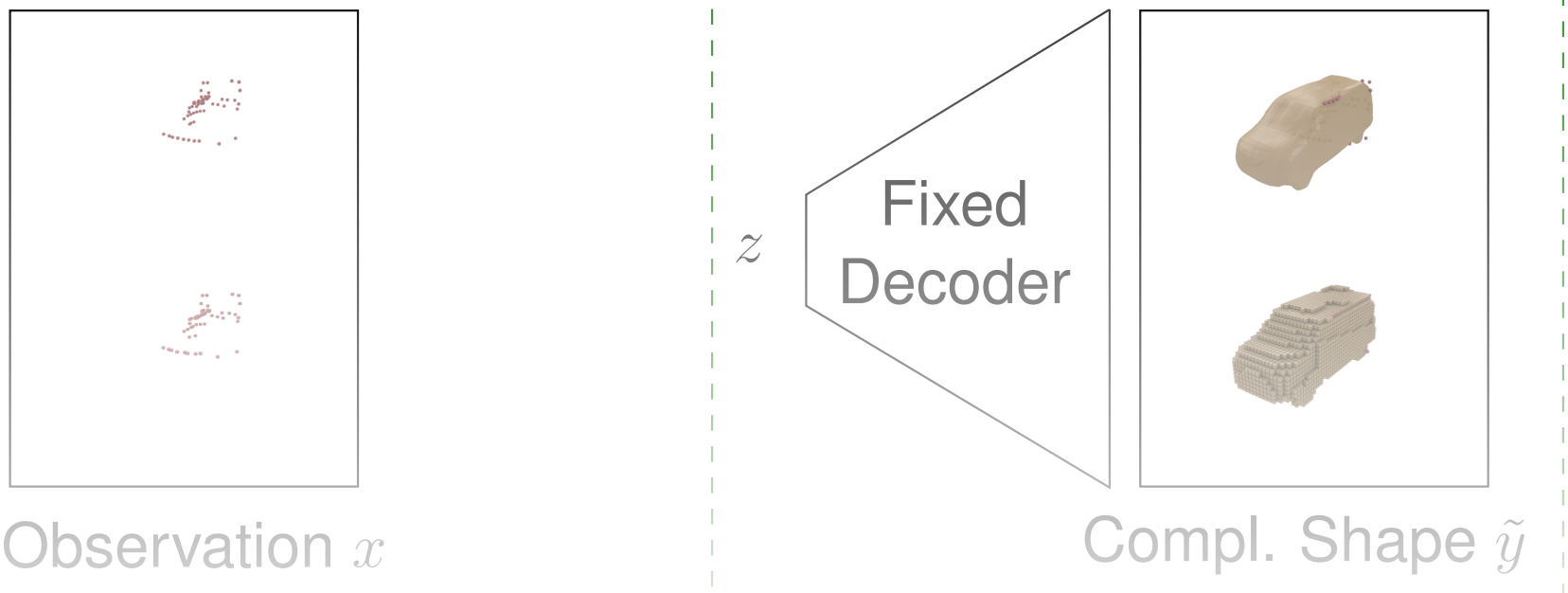
Perform maximum likelihood:



Shape Inference

Perform maximum likelihood:

$$\mathcal{L}_{\text{ML}}(z) = -\log p(y = x, z)$$



Shape Inference

Perform maximum likelihood:

$$\mathcal{L}_{\text{ML}}(z) = \sum_i -\log p(y_i = x_i|z) - \log p(z)$$

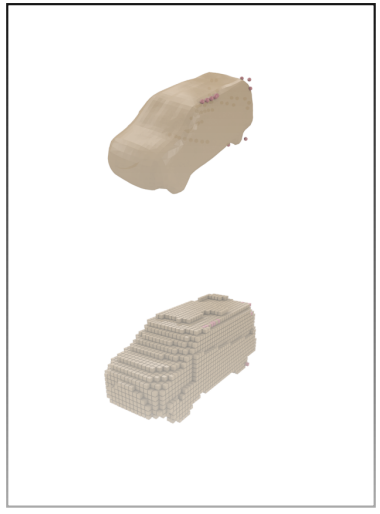
sum over voxels

Observation x

Generative Model $p(y|z)p(z)$

z

Fixed Decoder



Compl. Shape $\tilde{y}$

Shape Inference

Perform maximum likelihood:

$$\mathcal{L}_{\text{ML}}(z) = \sum_i -\log p(y_i = x_i | z) - \log p(z)$$

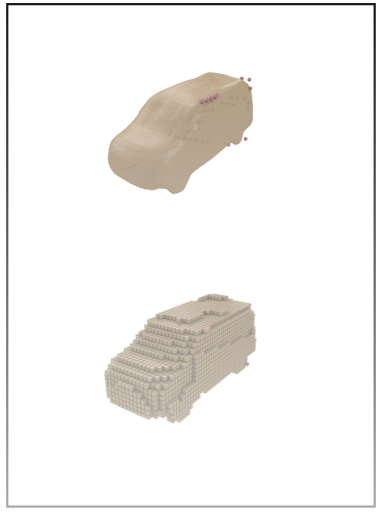
sum over voxels

Observation x

Generative Model $p(y|z)p(z)$

z

Fixed Decoder



Compl. Shape $\tilde{y}$

Shape Inference

What observations do we have?

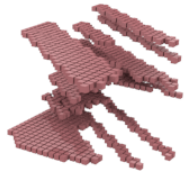
- Occupied voxels $x_i = 1$ (from observed points);

Occupied
Voxels $x_i = 1$

Shape Inference

What observations do we have?

- Occupied voxels $x_i = 1$ (from observed points);
- unoccupied voxels $x_i = 0$ (from free space);



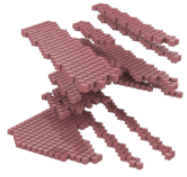
Free Space
Voxels $x_i = 0$

Occupied
Voxels $x_i = 1$

Shape Inference

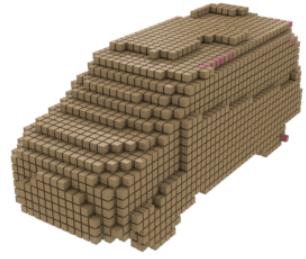
What observations do we have?

- Occupied voxels $x_i = 1$ (from observed points);
- unoccupied voxels $x_i = 0$ (from free space);
- and unknown voxels "$x_i = \perp$".



Free Space
Voxels $x_i = 0$

Occupied
Voxels $x_i = 1$



Unknown Shape y^*

Shape Inference

Perform maximum likelihood:

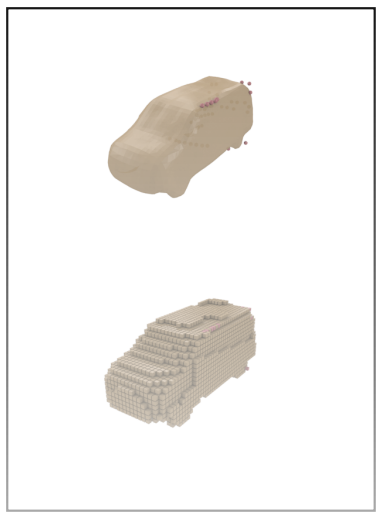
$$\mathcal{L}_{\text{ML}}(z) = \sum_i -\log p(y_i = x_i|z) - \log p(z)$$

Observation x

Generative Model $p(y|z)p(z)$

z

Fixed Decoder



Compl. Shape $\tilde{y}$

Shape Inference

Perform maximum likelihood:

$$\mathcal{L}_{\text{ML}}(z) = \sum_{x_i \neq \perp} -\log p(y_i = x_i | z) - \log p(z)$$

Observation x

Generative Model $p(y|z)p(z)$

z

Fixed Decoder

Compl. Shape $\tilde{y}$

Shape Inference

Perform maximum likelihood:

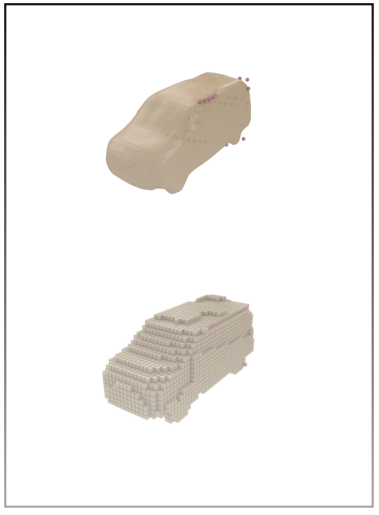
$$\mathcal{L}_{\text{ML}}(z) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} - \log p(z)$$

Observation x

Generative Model $p(y|z)p(z)$

z

Fixed Decoder



Compl. Shape $\tilde{y}$

Shape Inference

Perform maximum likelihood:

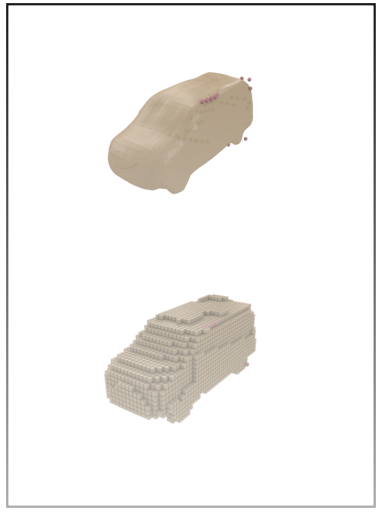
$$\mathcal{L}_{\text{ML}}(z) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \|z\|_2^2$$

Observation x

Generative Model $p(y|z)p(z)$

z

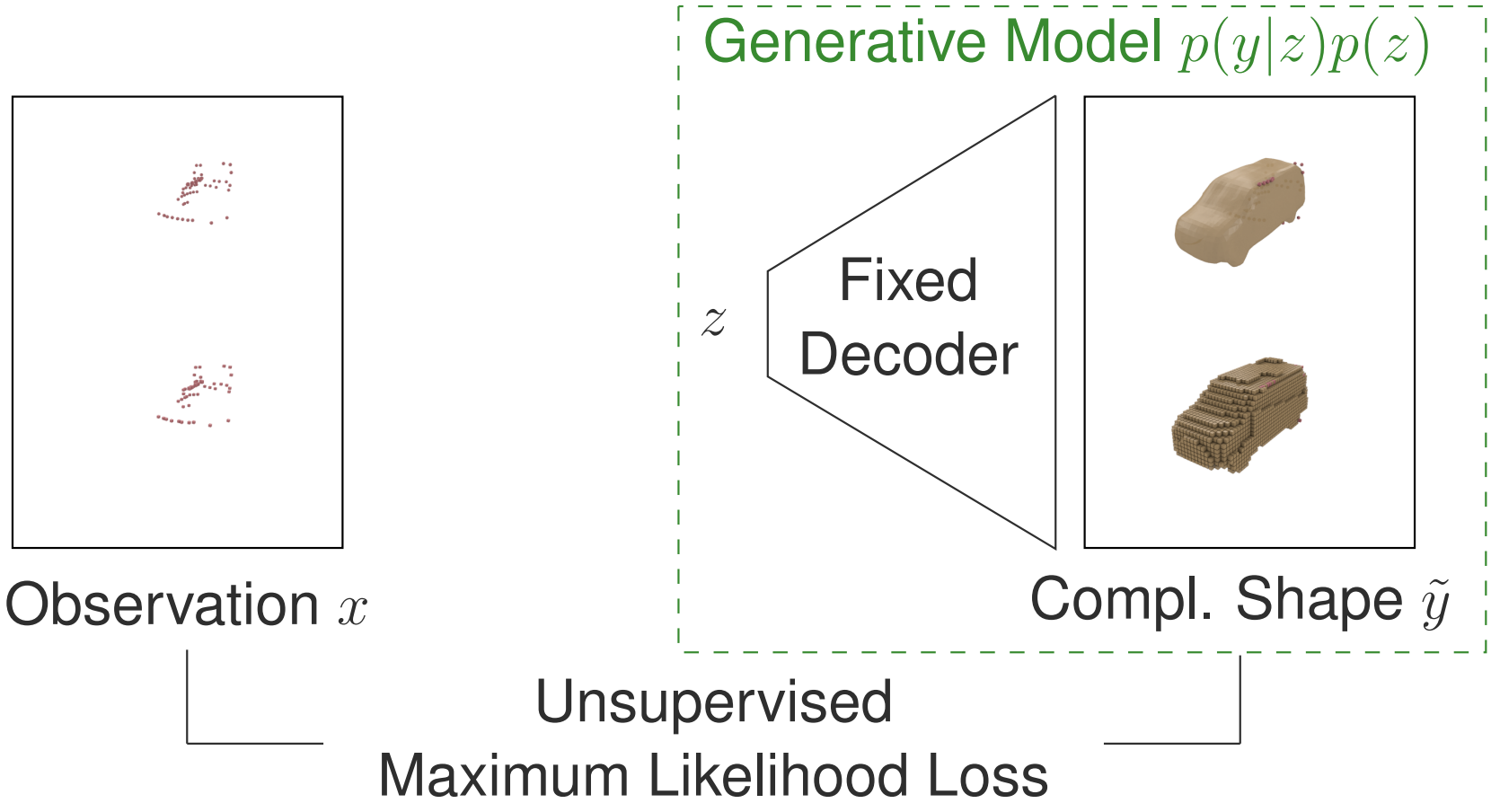
Fixed Decoder



Compl. Shape $\tilde{y}$

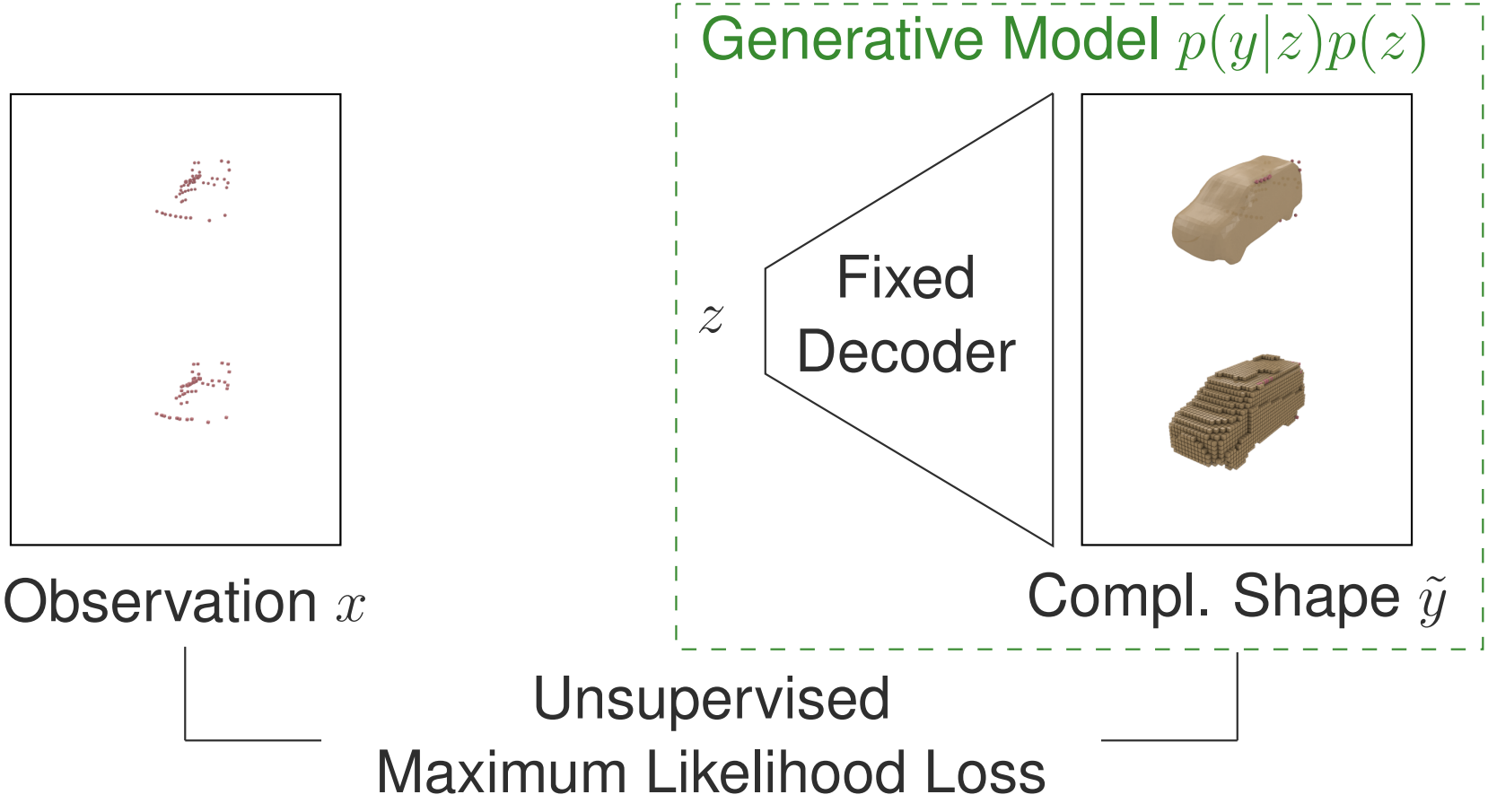
Shape Inference

Perform maximum likelihood:



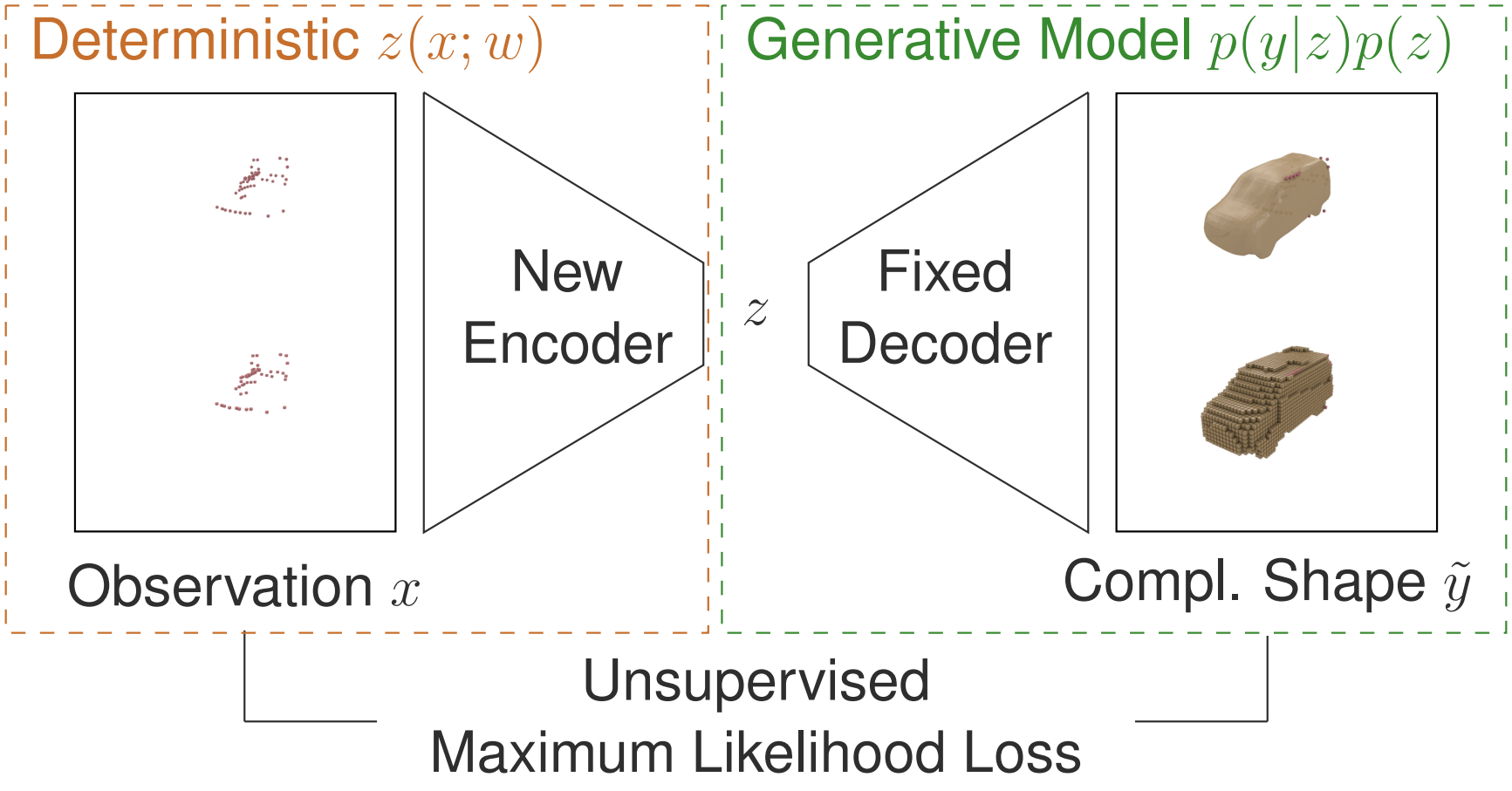
Shape Inference

~~Perform~~ Learn, *i.e.*, *amortize*, maximum likelihood:



Shape Inference

~~Perform~~ Learn, *i.e.*, *amortize*, maximum likelihood:

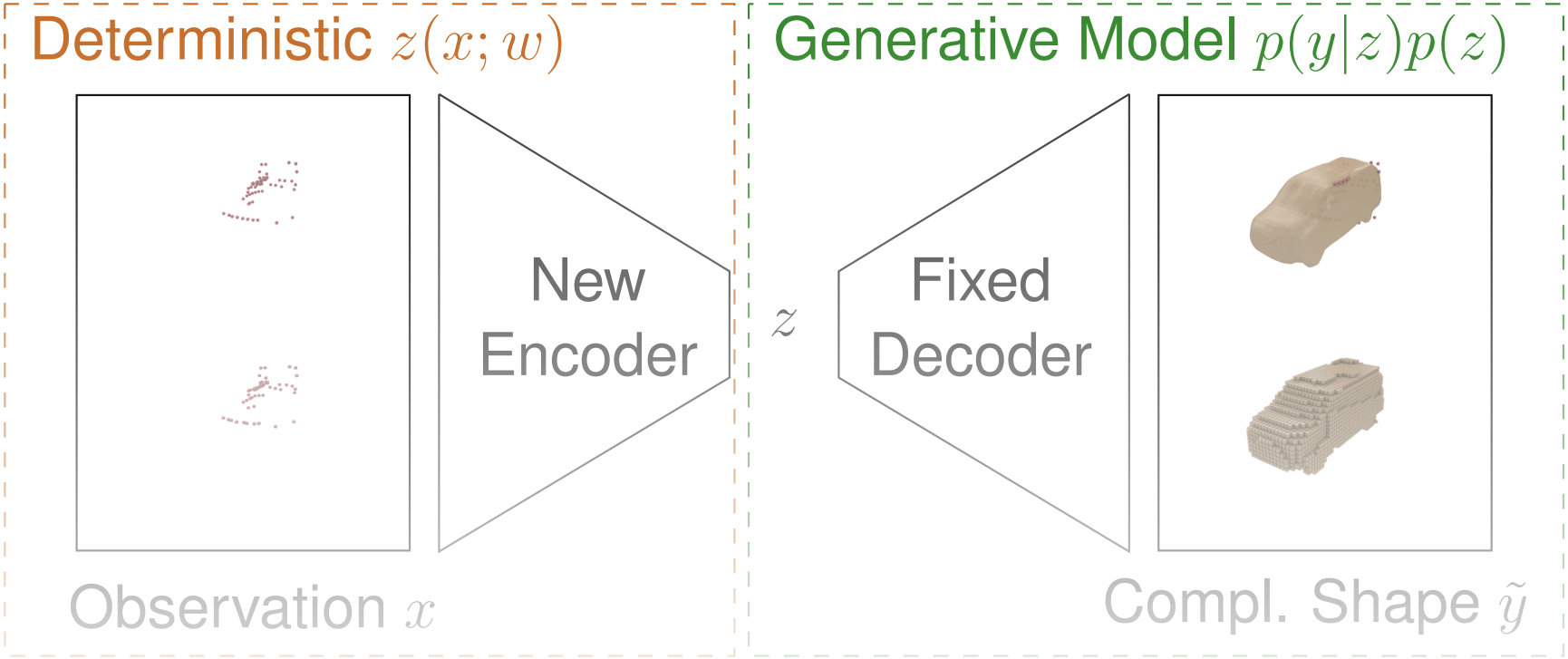


Shape Inference

~~Perform~~ Learn, *i.e.*, *amortize*, maximum likelihood:

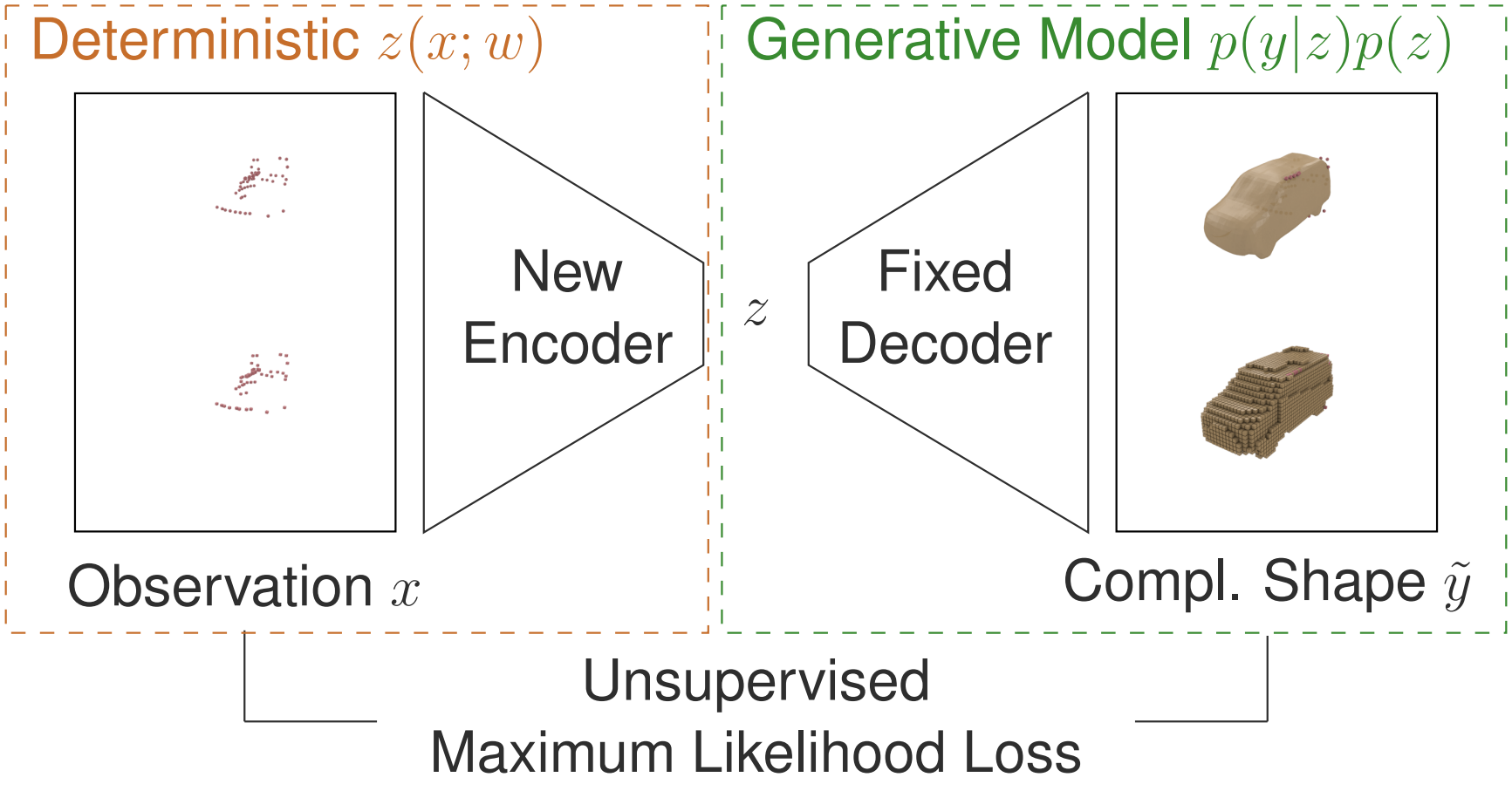
$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \|z(x; w)\|_2^2$$

new encoder neural network



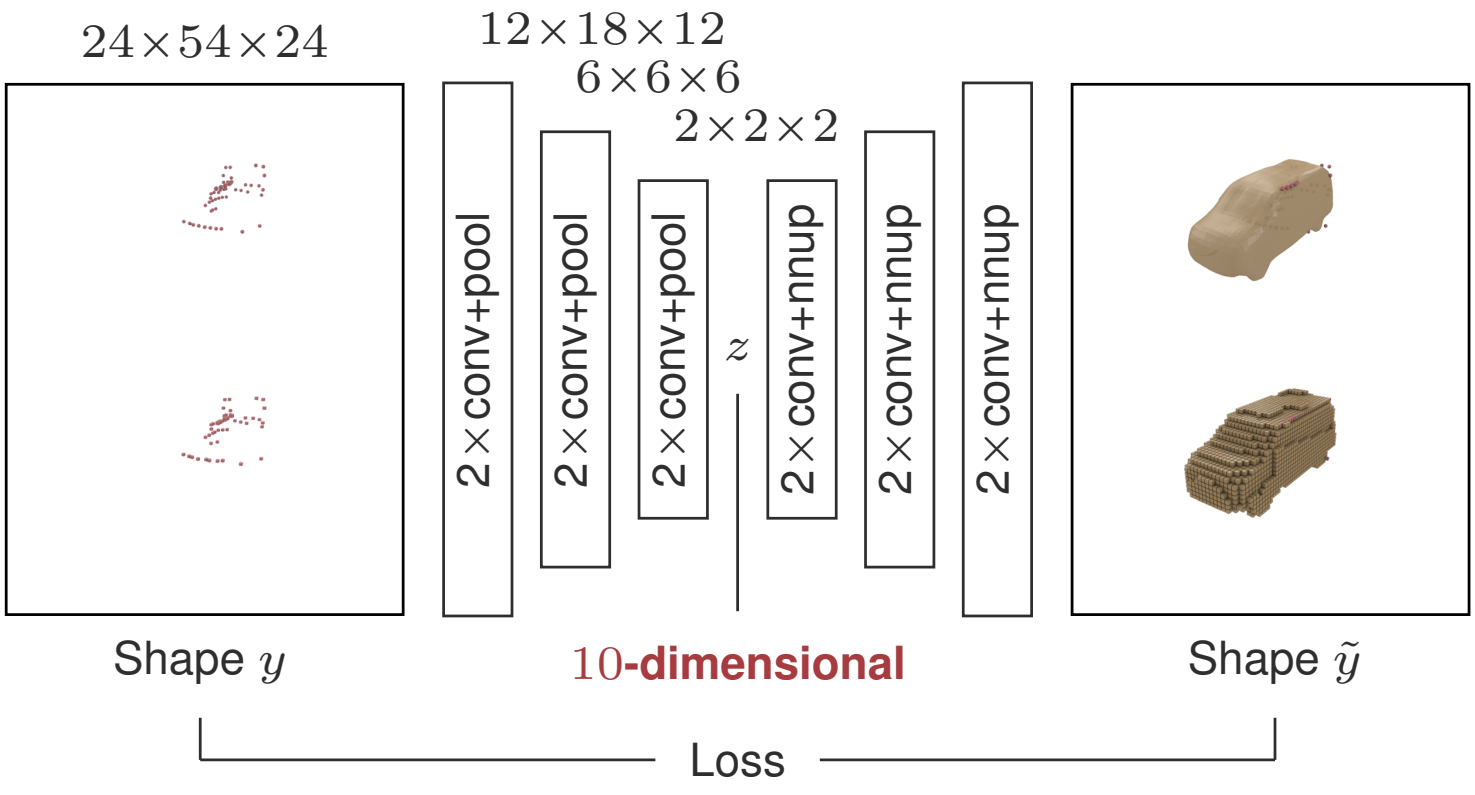
Shape Inference

~~Perform~~ Learn, *i.e.*, *amortize*, maximum likelihood:



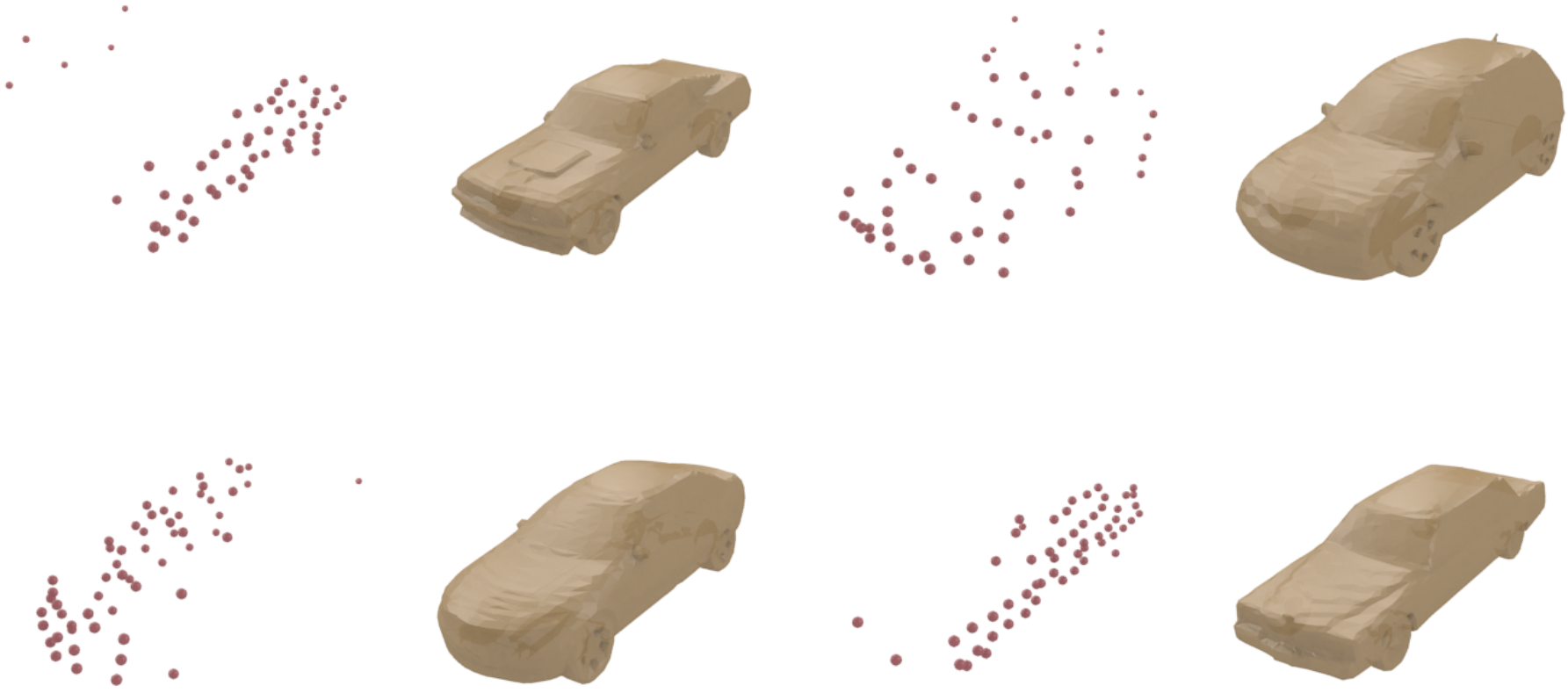
Experiments

Architecture



ShapeNet

Obs GT Obs GT



Quantitative Results on ShapeNet

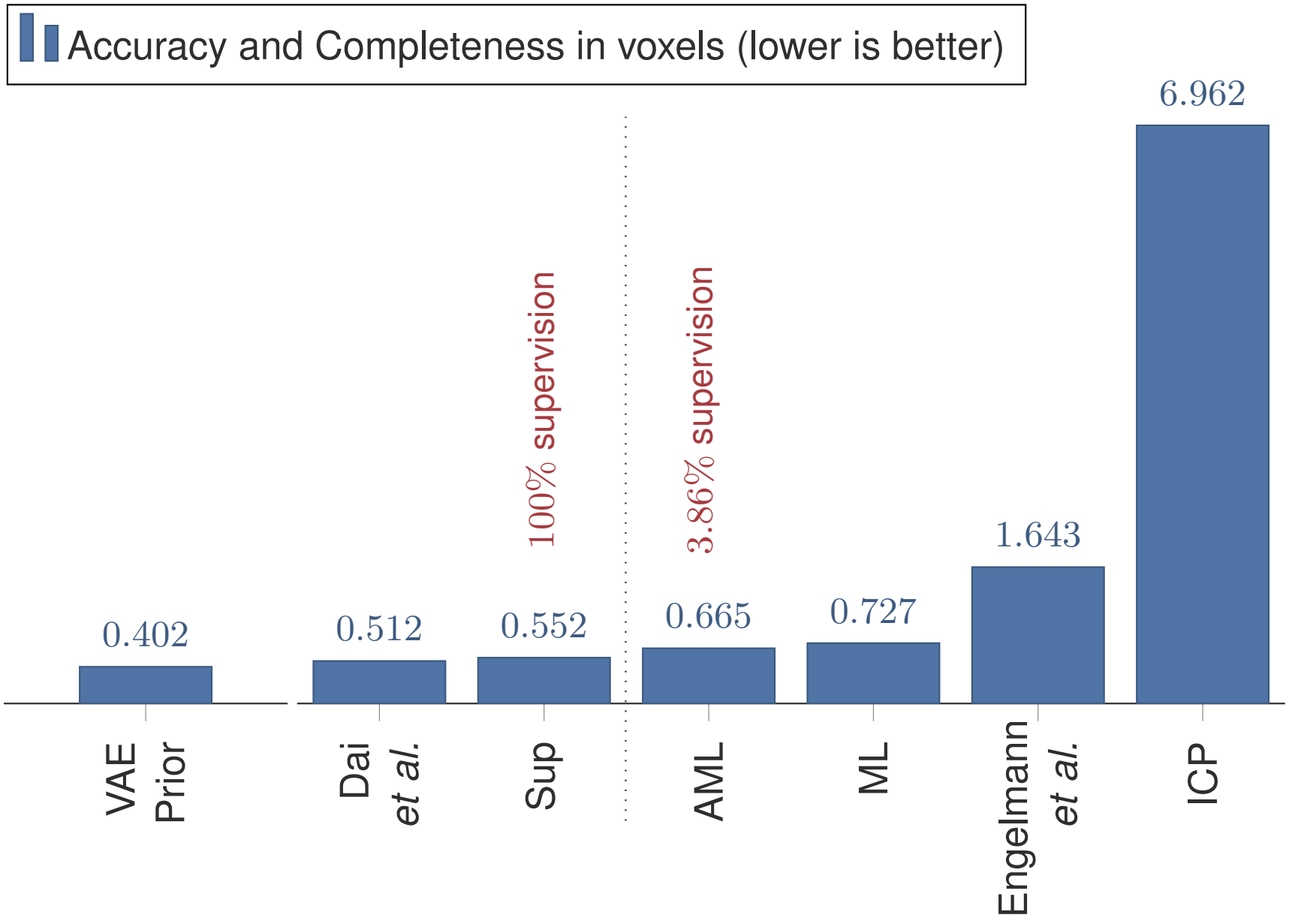
Completeness = distance of points on y^* to $\tilde{y}$

True Shape y^*

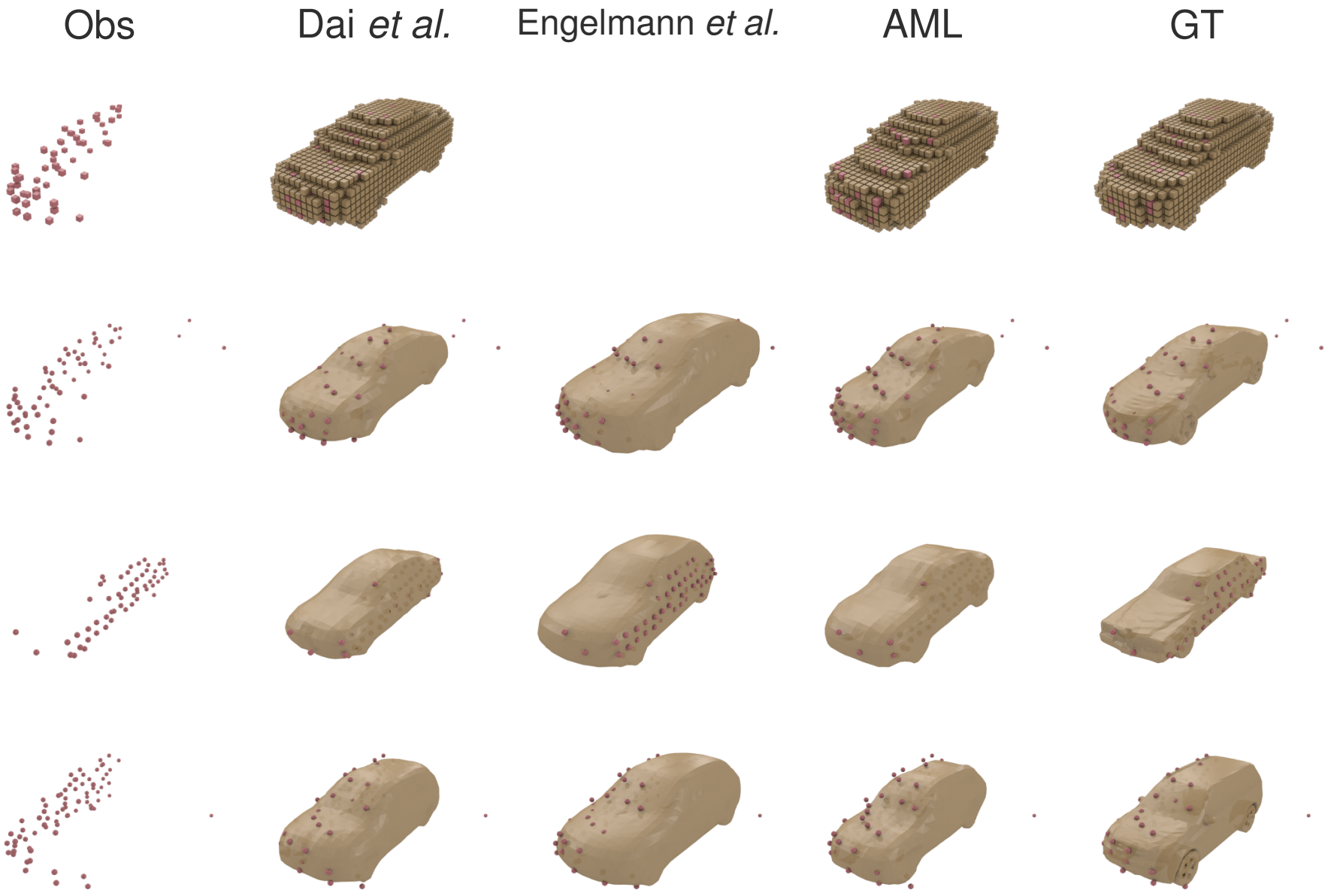
Rec. Shape $\tilde{y}$

Accuracy = distance of points on $\tilde{y}$ to y^*

Quantitative Results on ShapeNet

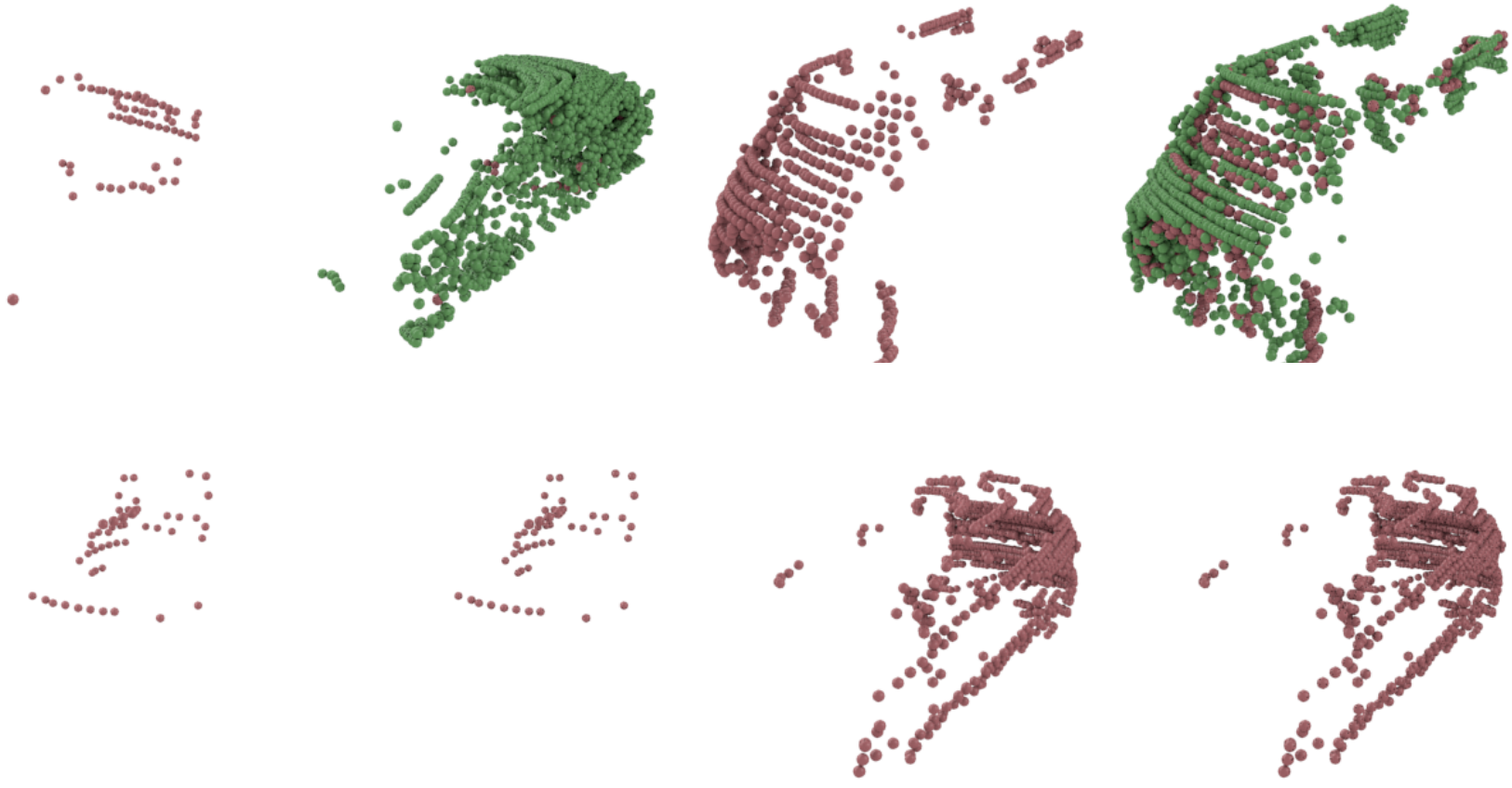


Qualitative Results on ShapeNet (noisy)

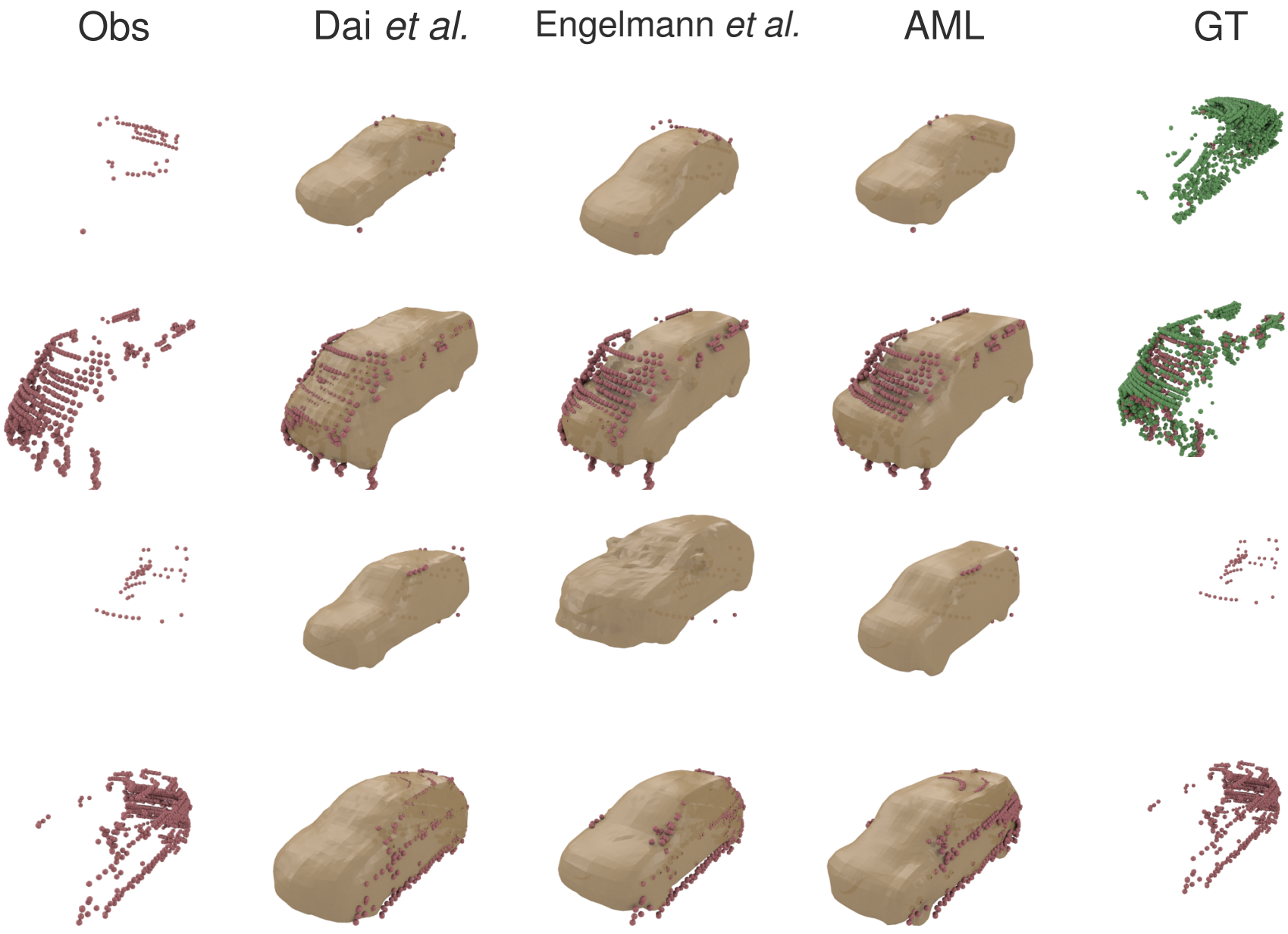


KITTI

Obs GT Obs GT

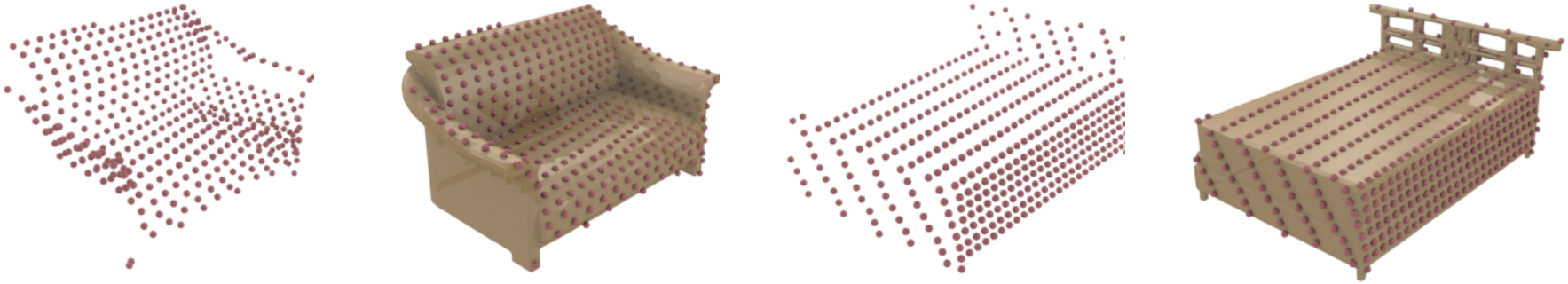


Qualitative Results on KITTI

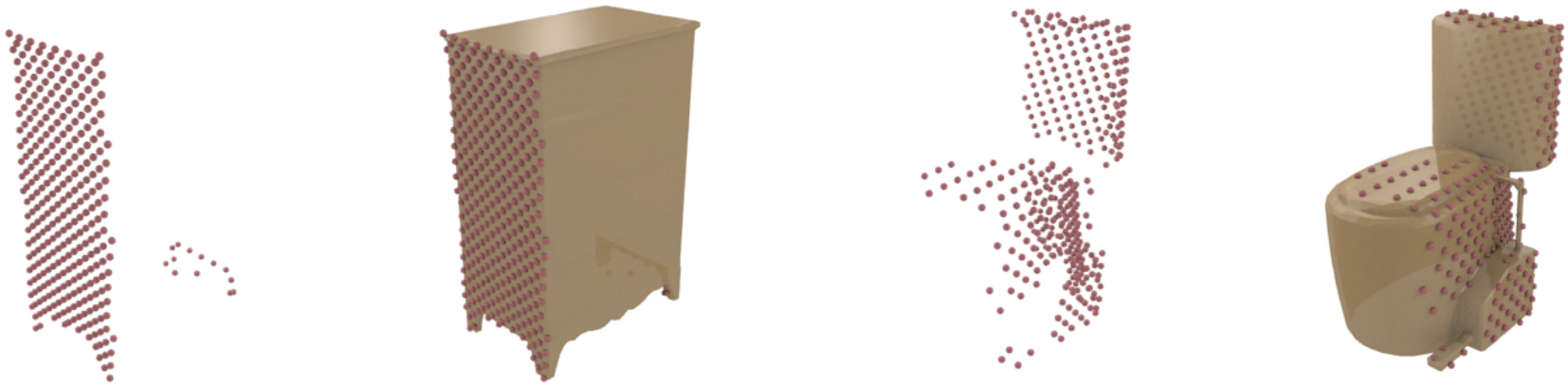


ModelNet10

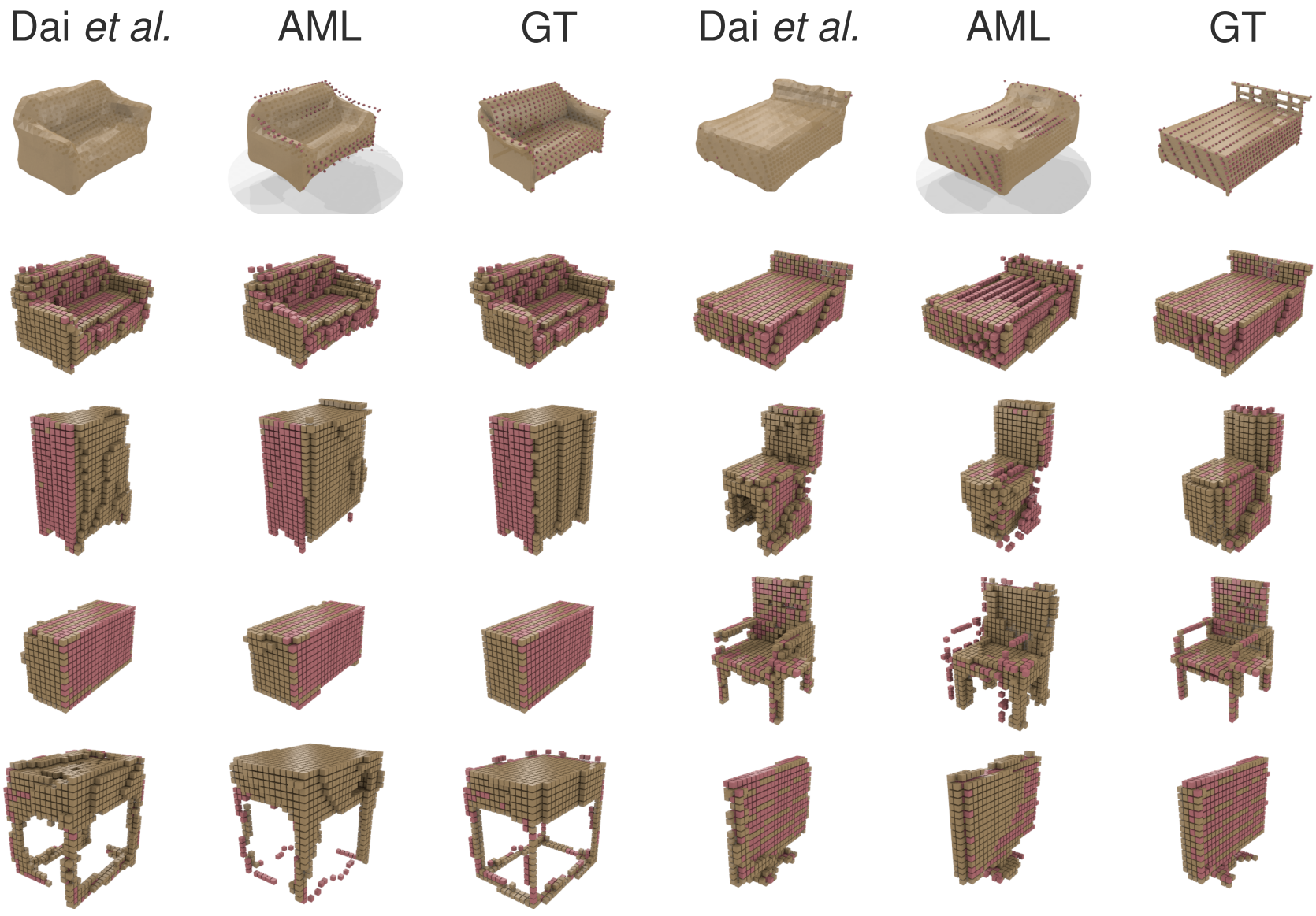
Obs GT Obs GT



Obs GT Obs GT



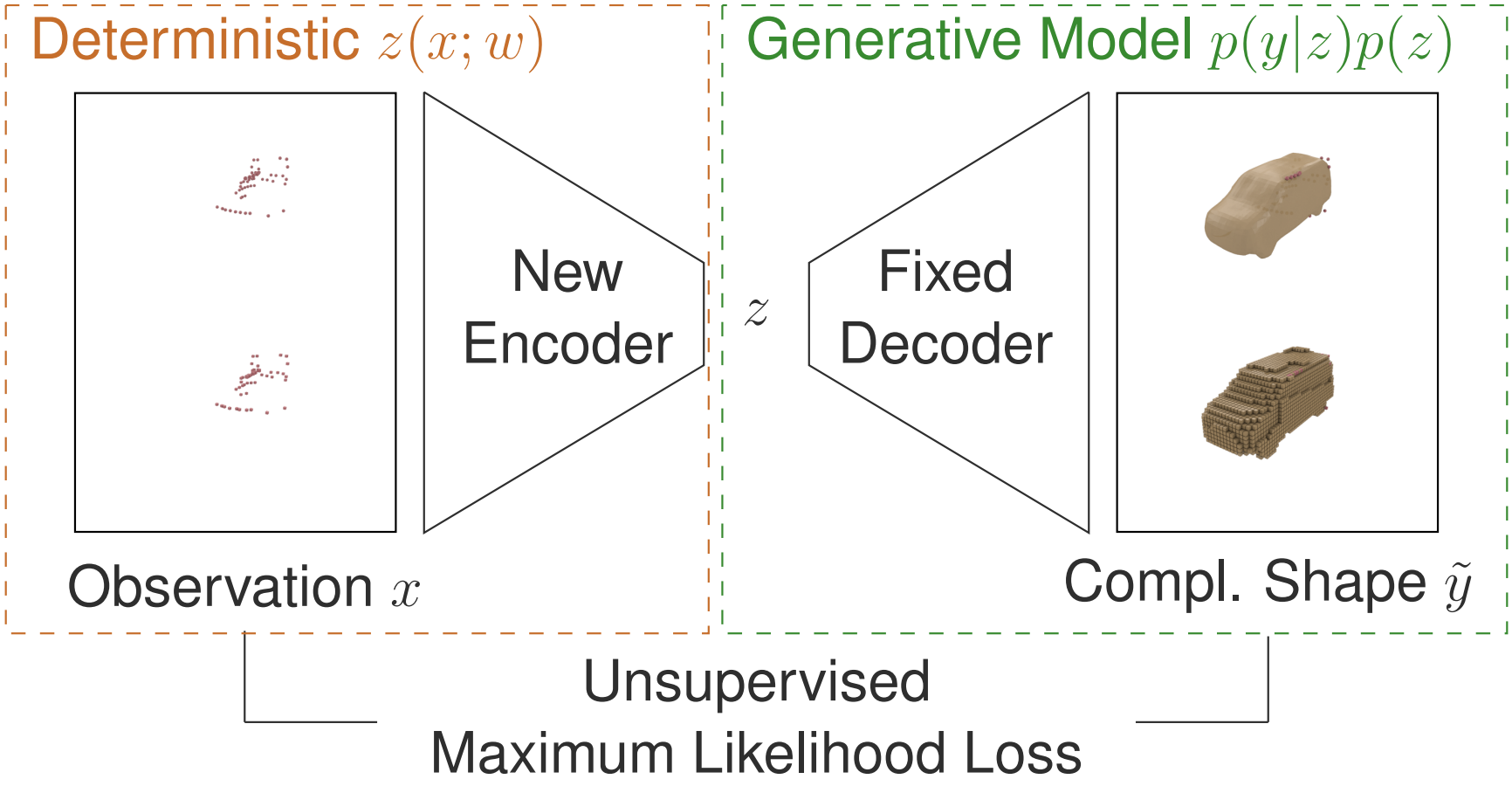
Qualitative Results on ModelNet10



Conclusion

Summary

We proposed an amortized maximum likelihood framework for **learning** 3D shape completion in a **weakly-supervised** setting.



Conclusion

Our experiments show that:

Hypothesis

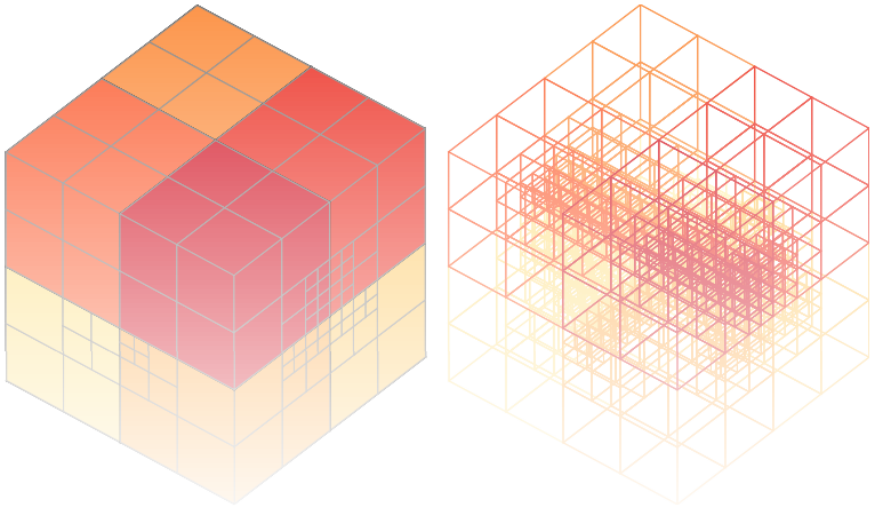
Strong shape priors allow to learn shape completion under weak supervision.

- Additionally, inference is efficient;
- and can be learned on real data.

Future Work

Possible improvements:

- Expressive shape models, *e.g.*, VAEs, GANs ...
- Efficient training in high resolutions ...

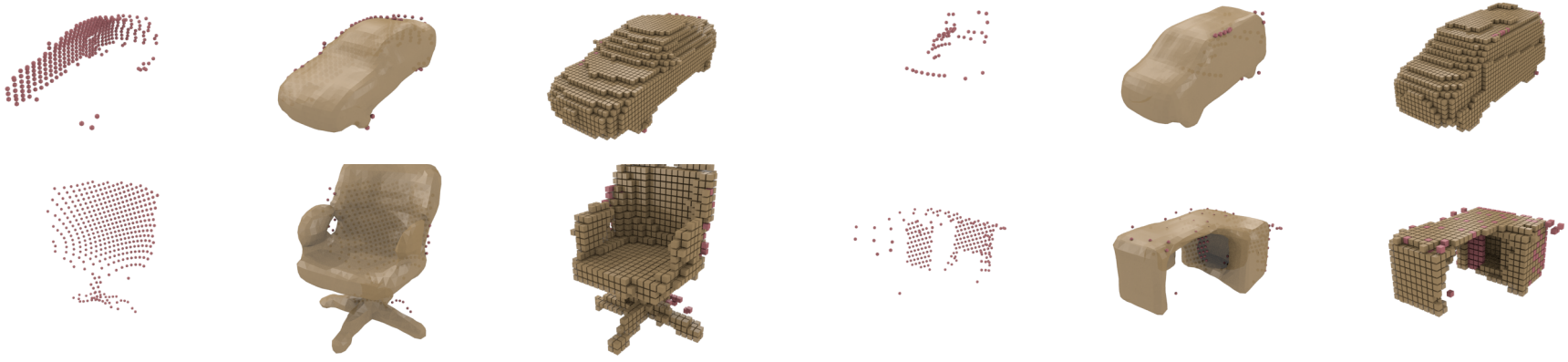


[Riegler et al., 2017b]

[Smith and Meger, 2017]

Questions?

Learning Shape Completion from Bounding Boxes with CAD Shape Priors



David Stutz

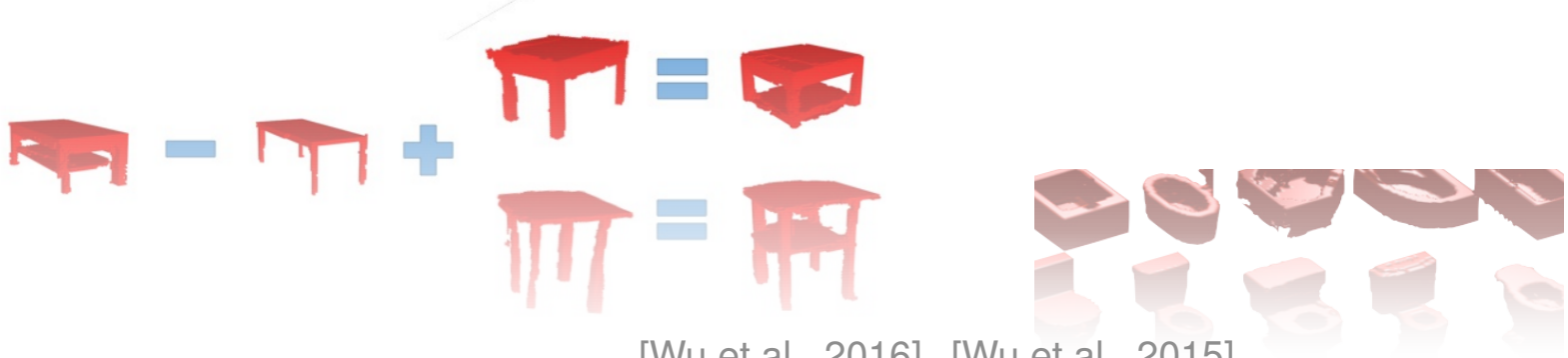
davidstutz.de

Appendix

Appendix – Selected Related Work

Generative shape modeling, *e.g.*
[Wu et al., 2015, Wu et al., 2016]:

- Learn a generative model of shapes;
- and use for shape classification, manipulation and generation.

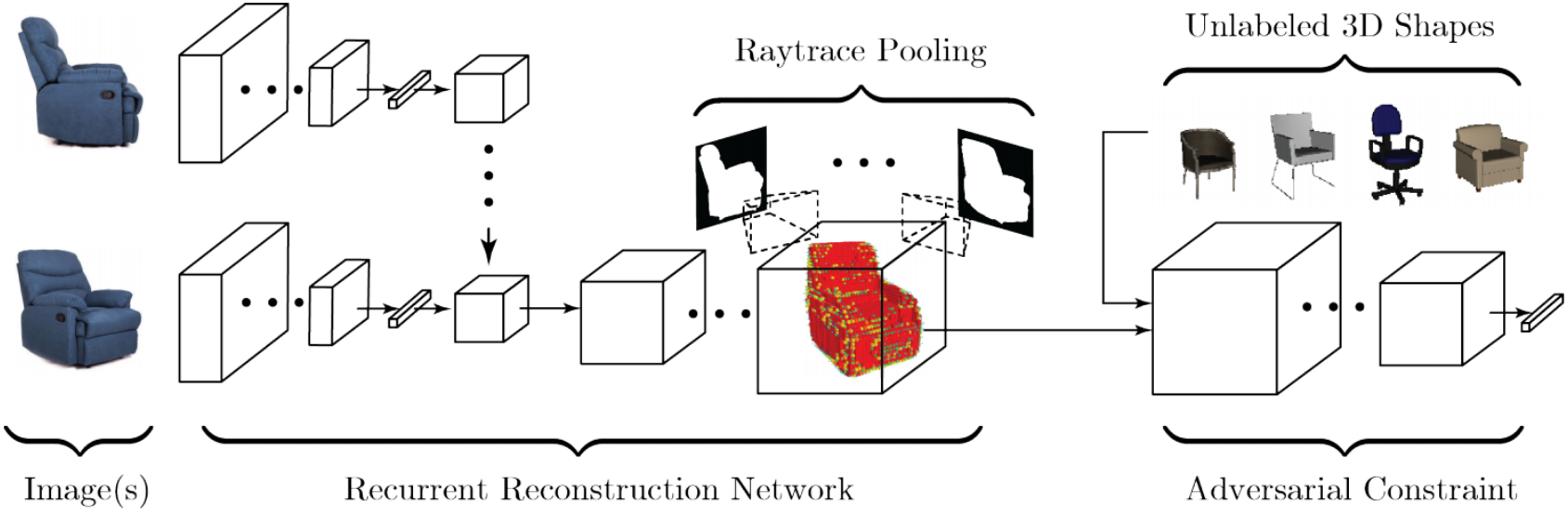


[Wu et al., 2016] [Wu et al., 2015]

Appendix – Selected Related Work

Weakly-supervised 3D reconstruction, specifically [Gwak et al., 2017].

- Enforce implicit shape prior through adversarial training.



[Gwak et al., 2017]

Appendix – Related Work

Symmetry-based approaches:

- [Thrun and Wegbreit, 2005, Pauly et al., 2008, Zheng et al., 2010, Kroemer et al., 2012, Law and Aliaga, 2011]

Appendix – Related Work

Data-driven approaches:

- Explicit retrieval
  [Pauly et al., 2005, Gupta et al., 2015];
- learned shape prior [Haene et al., 2014,
  Li et al., 2015, Engelmann et al., 2016,
  Engelmann et al., 2017, Bao et al., 2013,
  Dame et al., 2013, Nguyen et al., 2016] ...
  - *e.g.*, PCA, GP-LVM or mean shape + anchor points.

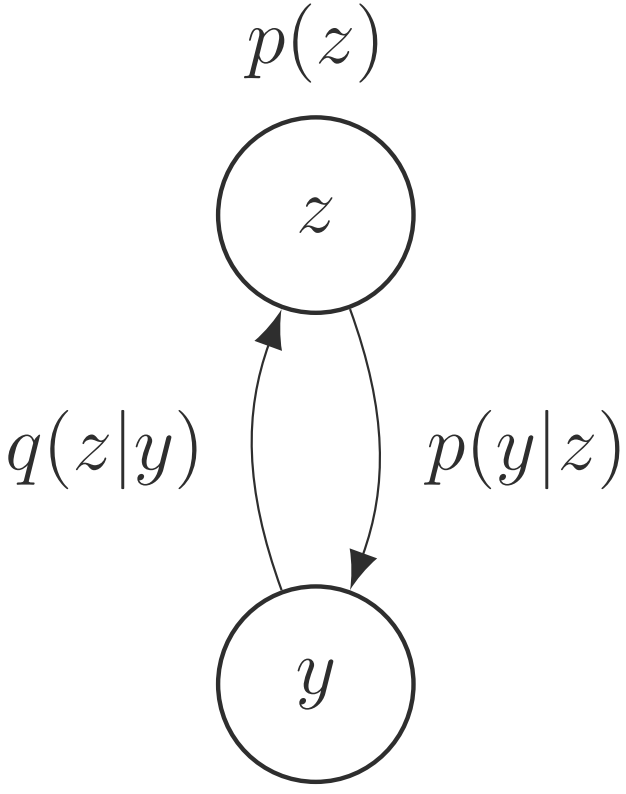
Appendix – Related Work

Learning-based approaches:

- [Firman et al., 2016, Smith and Meger, 2017, Dai et al., 2016, Sharma et al., 2016, Rezende et al., 2016, Fan et al., 2016, Riegler et al., 2017a, Han et al., 2017, Yang et al., 2017, Yang et al., 2018];
  - fully-supervised, U-net like encoder-decoder architecture;
  - predicting point clouds [Fan et al., 2016] or octrees [Riegler et al., 2017a].

Appendix – Shape Prior

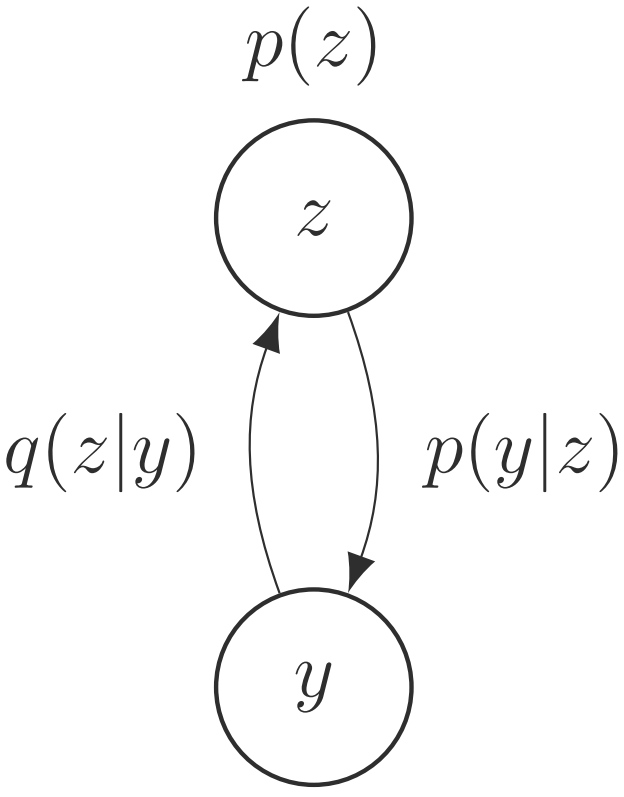
Learn a variational auto-encoder:



Prior $p(z) = \mathcal{N}(z|0, I)$

Appendix – Shape Prior

Learn a variational auto-encoder:



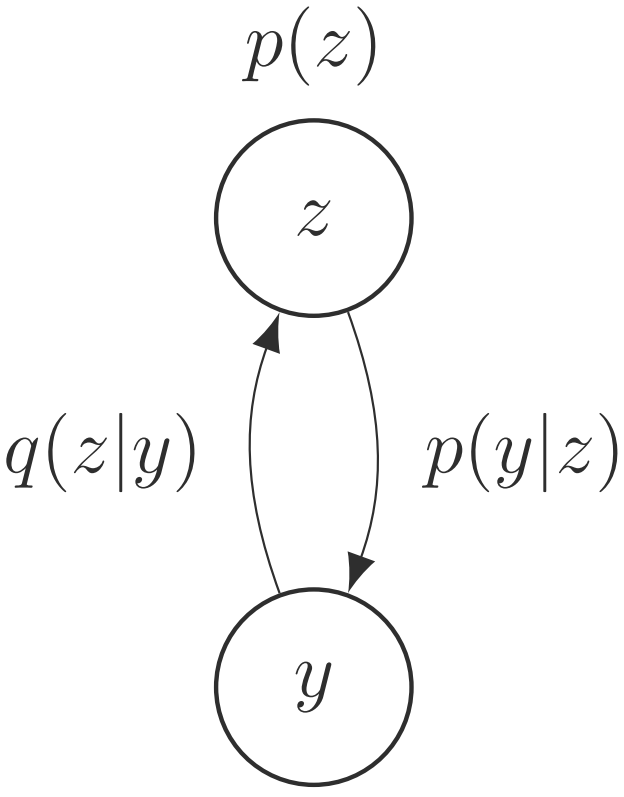
Prior $p(z) = \mathcal{N}(z|0, I)$

Decoder/Generative Model

$$p(y|z) = \prod_i \begin{Bmatrix} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{Bmatrix}$$

Appendix – Shape Prior

Learn a variational auto-encoder:



Prior $p(z) = \mathcal{N}(z|0, I)$

Decoder/Generative Model

$$p(y|z) = \prod_i \begin{Bmatrix} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{Bmatrix}$$

Encoder/Recognition Model

$$q(z|y) = \mathcal{N}(z|\mu(y), \sigma^2(y))$$

Appendix – Shape Prior

Maximum likelihood leads to:

$$\mathcal{L}_{\text{ELBO}} = -\mathbb{E}_{q(z|y)}[\log p(y|z)] + \text{KL}(q(z|y)|p(z))$$

- Encoder: $q(z|y) = \mathcal{N}(z|\mu(y), \sigma^2(y))$;
- and decoder: $p(y|z) = \prod_i \left\{ \begin{matrix} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{matrix} \right\}$.

Appendix – Shape Prior

Maximum likelihood leads to:

$$\mathcal{L}_{\text{ELBO}} = -\mathbb{E}_{q(z|y)}[\log p(y|z)] + \mathcal{P}_{\text{Reg}}(\mu(y), \sigma^2(y))$$

- Encoder: $q(z|y) = \mathcal{N}(z|\mu(y), \sigma^2(y))$;
- and decoder: $p(y|z) = \prod_i \begin{Bmatrix} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{Bmatrix}$.

Appendix – Shape Prior

Maximum likelihood leads to:

$$\mathcal{L}_{\text{ELBO}} = \sum_i \mathcal{L}_{\text{BCE}}(\tilde{y}_i, y_i) + \mathcal{P}_{\text{Reg}}(\mu(y), \sigma^2(y))$$

- Encoder: $q(z|y) = \mathcal{N}(z|\mu(y), \sigma^2(y))$;
- and decoder: $p(y|z) = \prod_i \left\{ \begin{array}{c} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{array} \right\}$.

Appendix – Shape Prior

Maximum likelihood leads to:

$$\mathcal{L}_{\text{VAE}}(\tilde{y}, y) = \sum_i \mathcal{L}_{\text{BCE}}(\tilde{y}_i, y_i) + \mathcal{P}_{\text{Reg}}(\mu(y), \sigma^2(y))$$

- Encoder: $q(z|y) = \mathcal{N}(z|\mu(y), \sigma^2(y))$;
- and decoder: $p(y|z) = \prod_i \begin{Bmatrix} \text{Ber}(y_i|\theta_i(z)) \\ \mathcal{N}(y_i|\mu_i(z), \sigma^2) \end{Bmatrix}$.

Appendix – Shape Prior

Reconstructions.

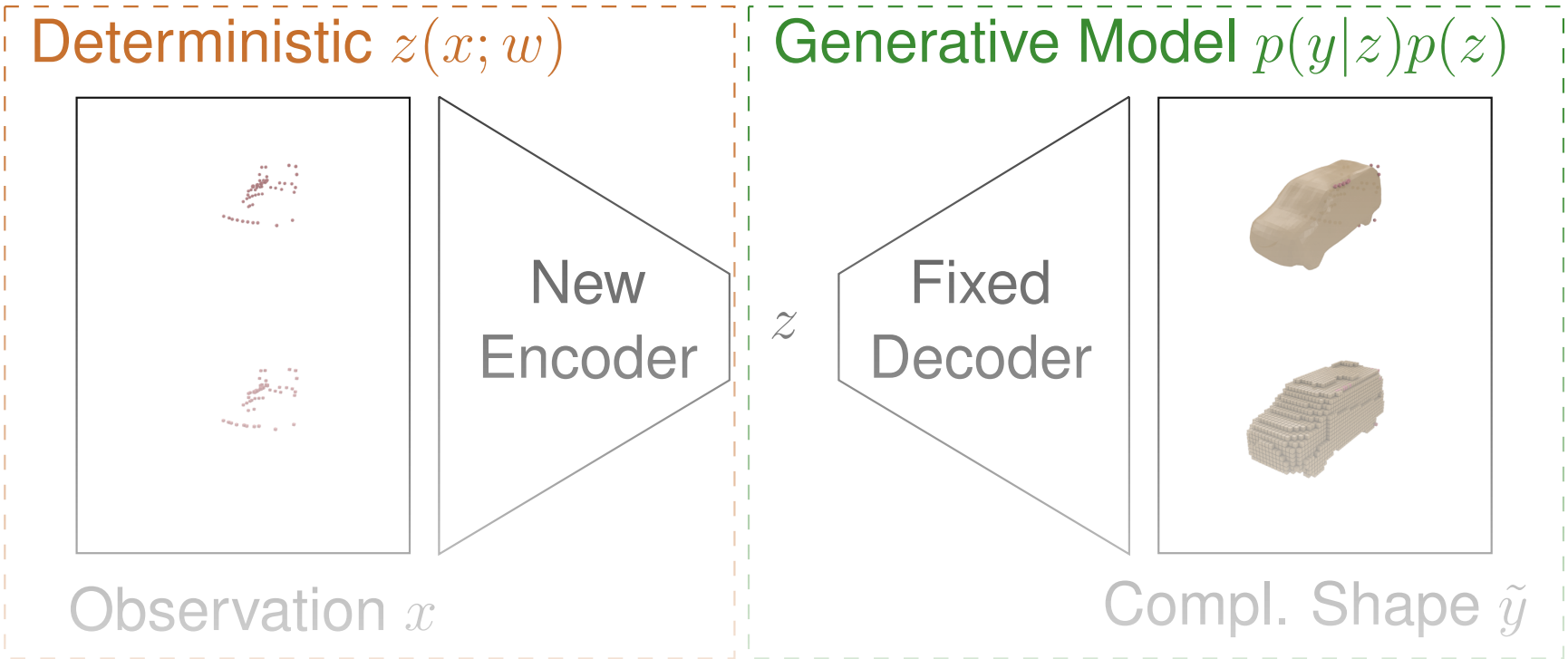
Random samples.

Appendix – Shape Inference

Enforcing variety:

$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \|z(x; w)\|_2^2$$

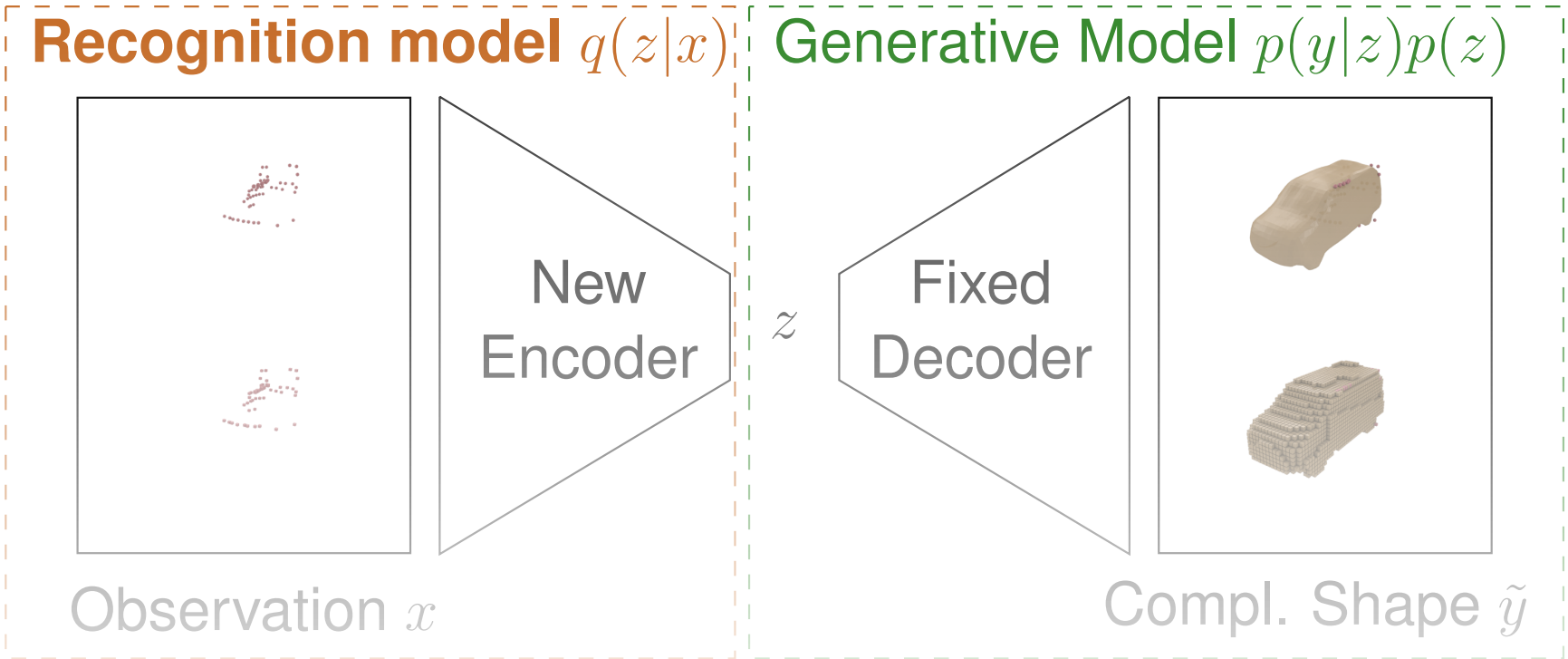
scales with #observations



Appendix – Shape Inference

Enforcing variety:

$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \text{KL}(q(z|x)|p(z))$$

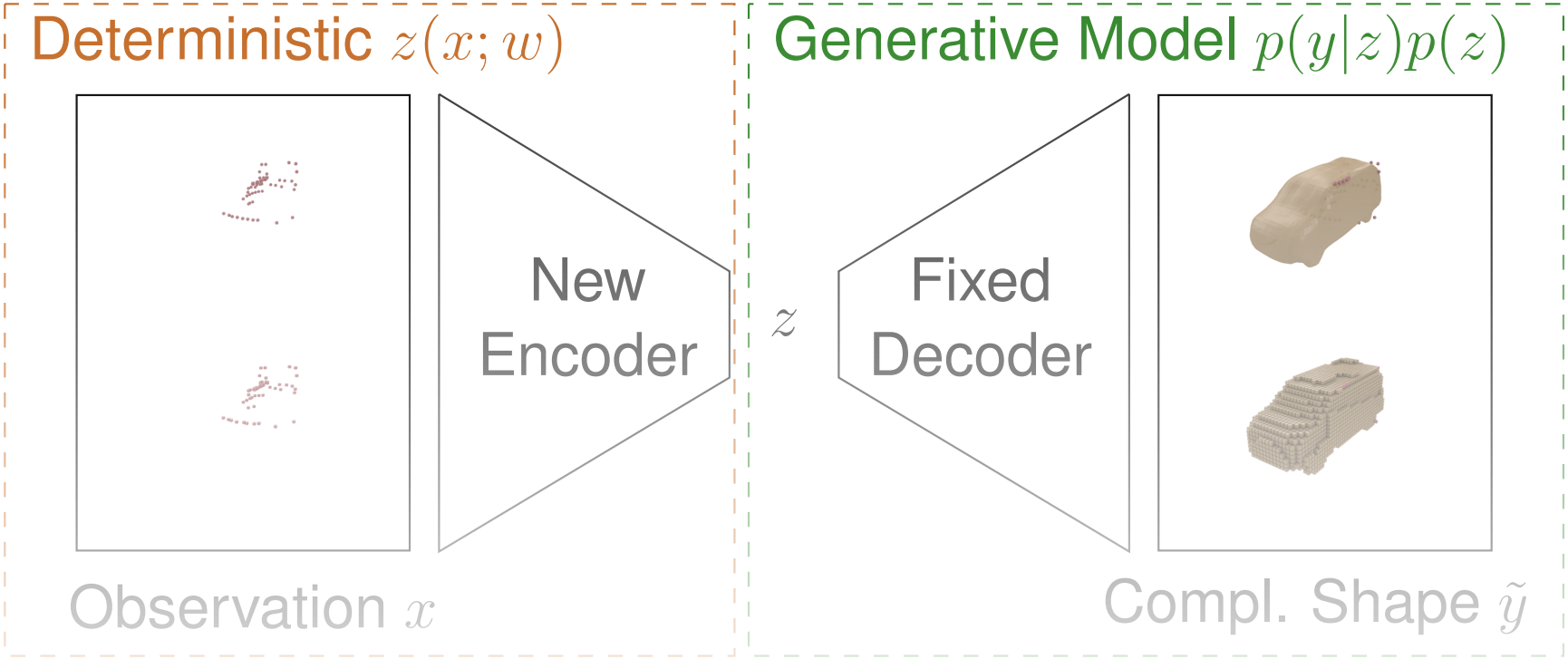


Appendix – Shape Inference

~~Perform~~ Learn, *i.e.*, *amortize*, maximum likelihood:

$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \|z(x; w)\|_2^2$$

new encoder neural network



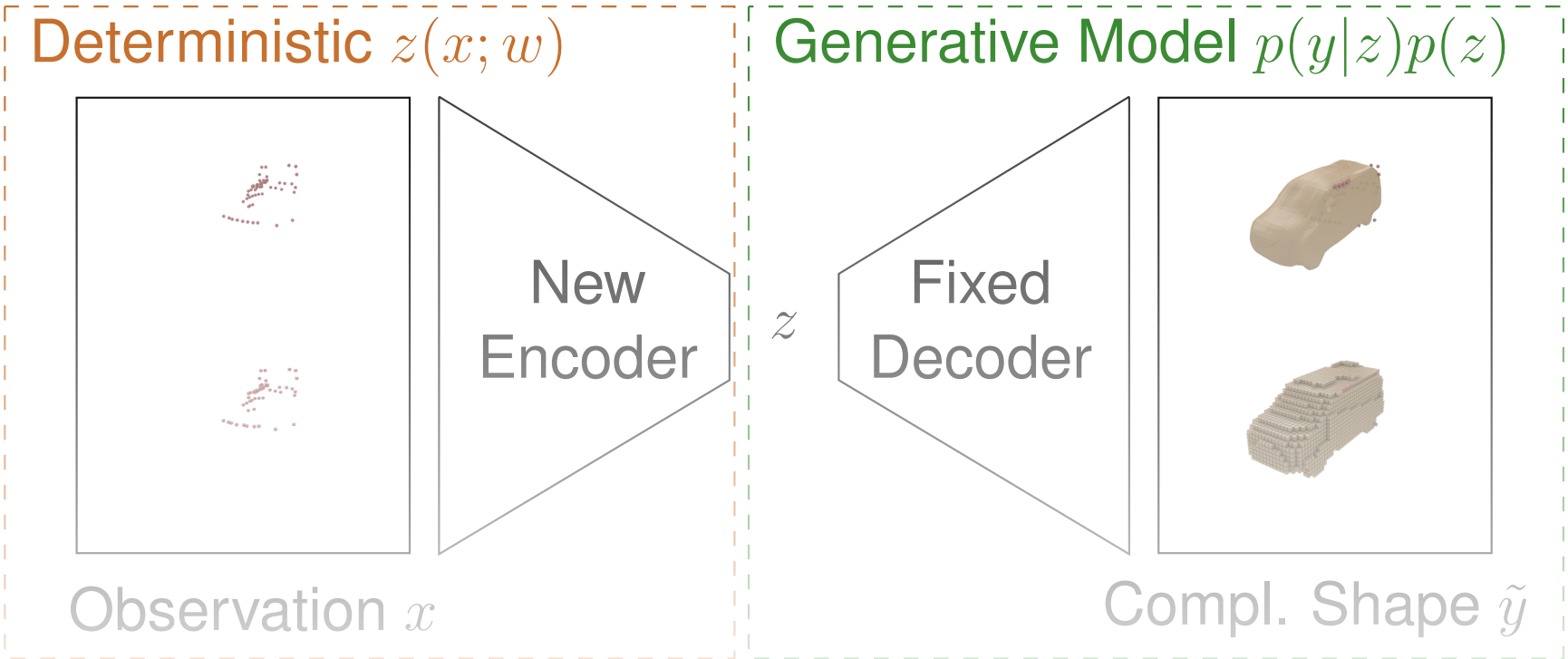
Appendix – Shape Inference

Handling SDFs:

$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{matrix} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{SSE}}(\tilde{y}_i, x_i) \end{matrix} \right\} + \|z(x; w)\|_2^2$$

continuous distance

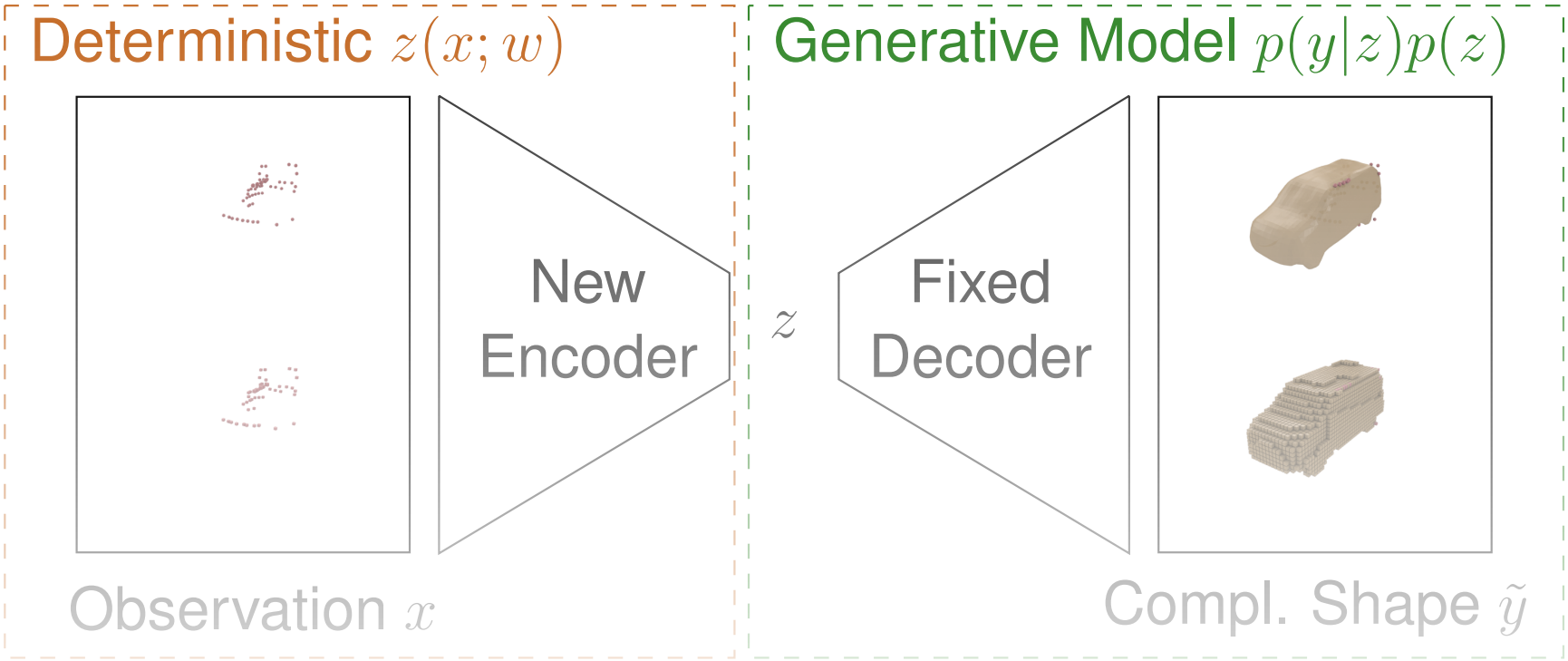
binary observation



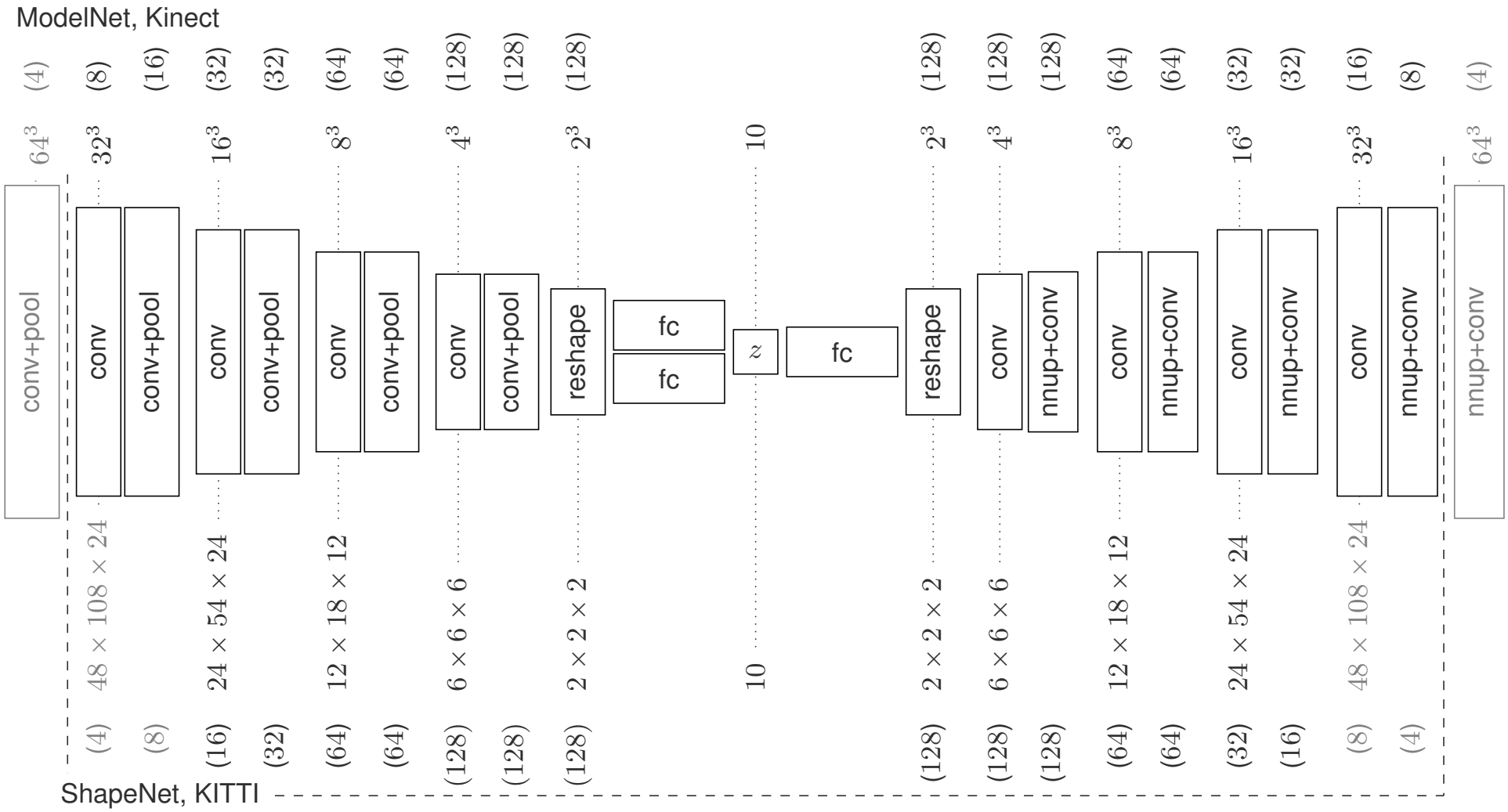
Appendix – Shape Inference

Handling SDFs:

$$\mathcal{L}_{\text{AML}}(w) = \sum_{x_i \neq \perp} \left\{ \begin{array}{c} \mathcal{L}_{\text{BCE}}(\tilde{y}_i, x_i) \\ \mathcal{L}_{\text{BCE}}(\tilde{\theta}(y_i), x_i) \end{array} \right\} + \|z(x; w)\|_2^2$$

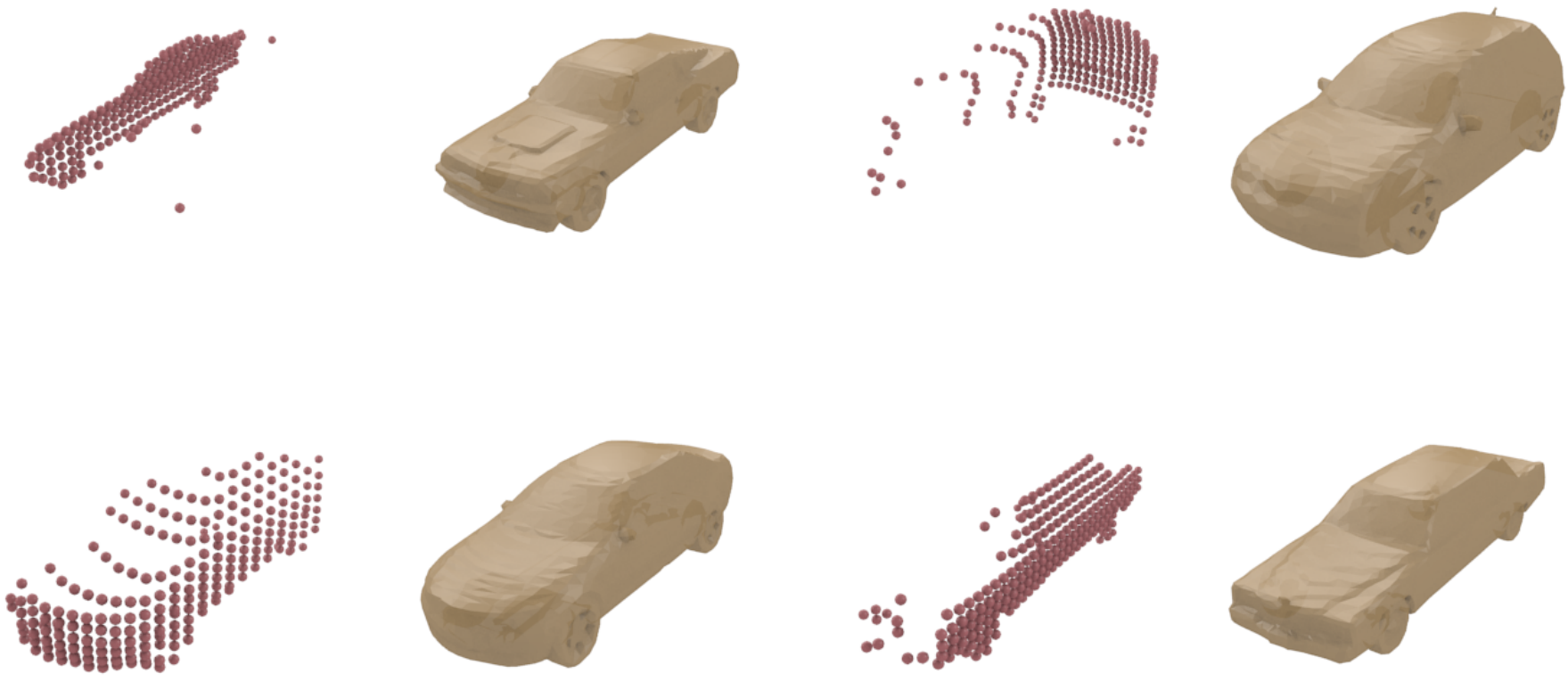


Appendix – Architectures

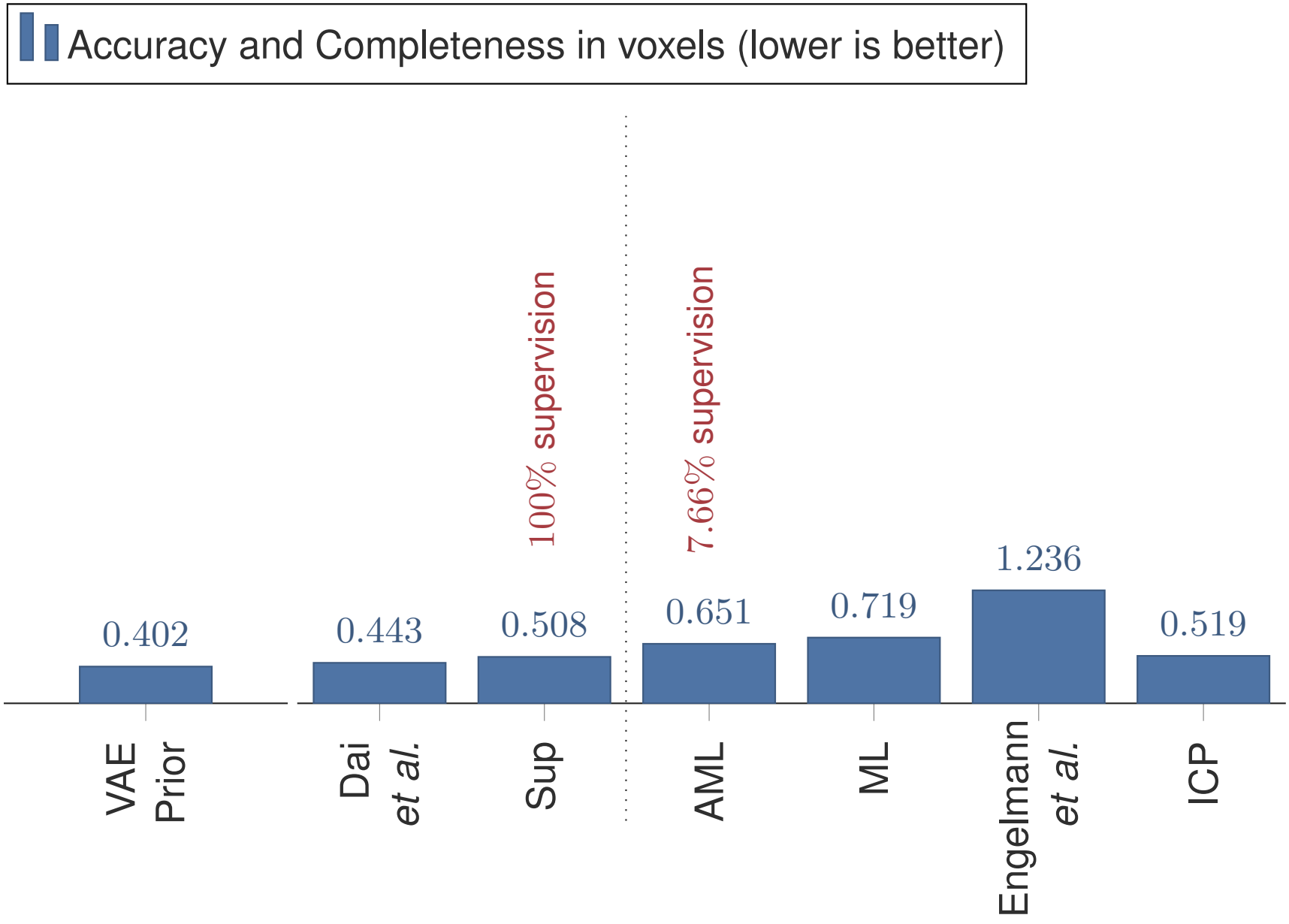


Appendix – ShapeNet (clean)

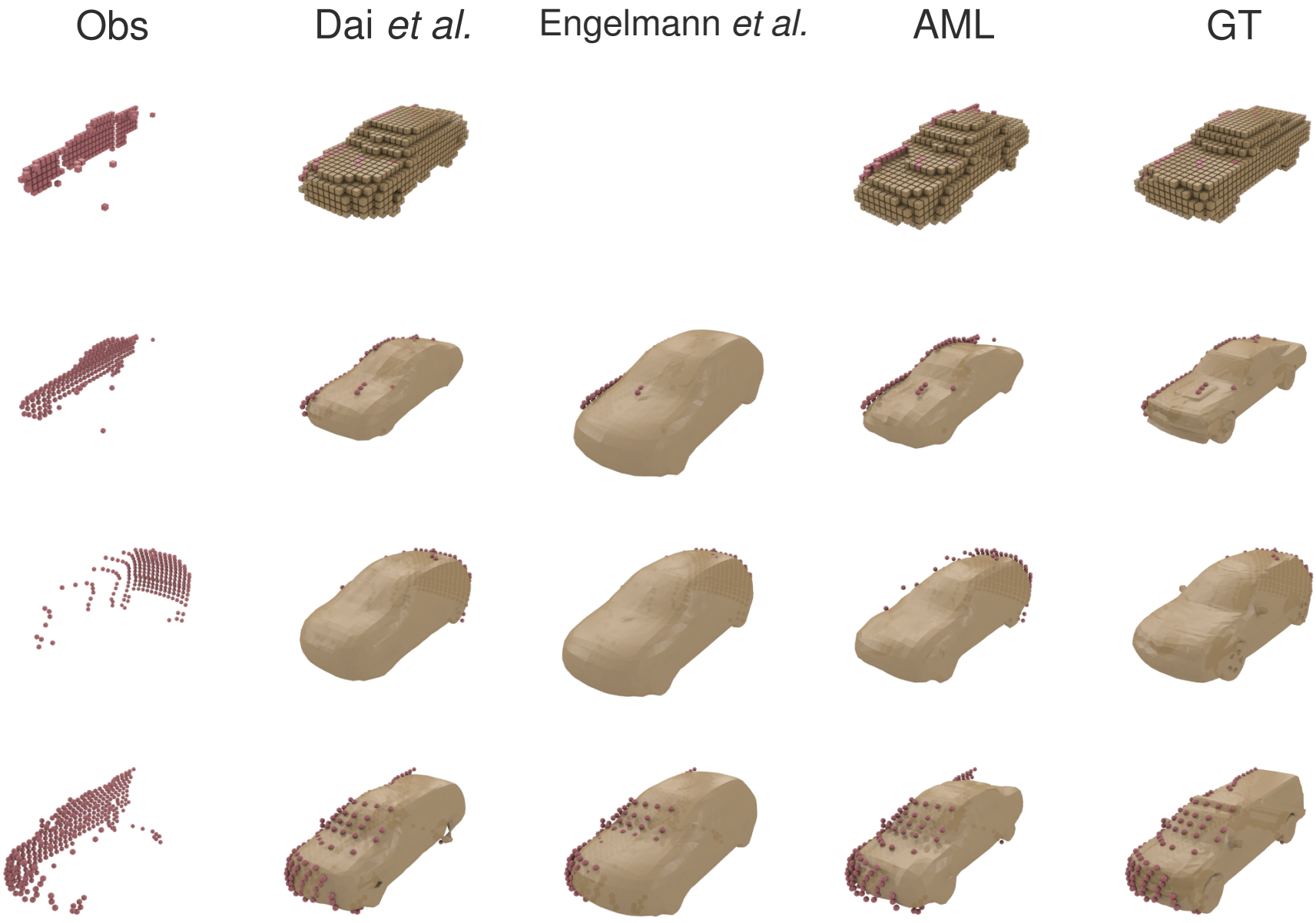
Obs GT Obs GT



Appendix – Results on ShapeNet (clean)

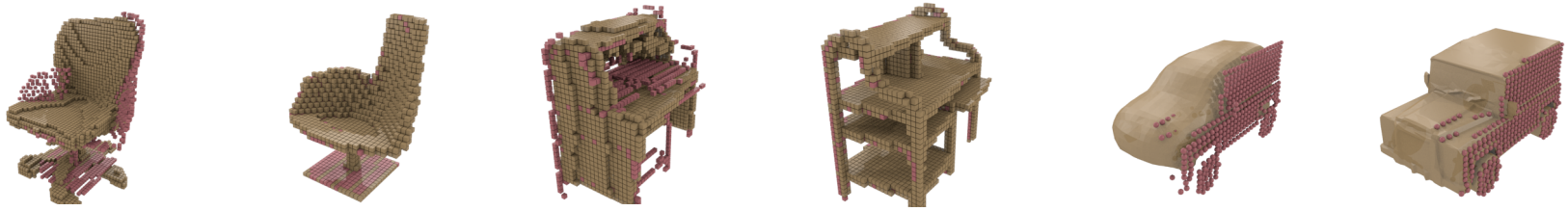


Appendix – Results on ShapeNet (clean)



Appendix – Failure Cases

AML GT AML GT AML GT



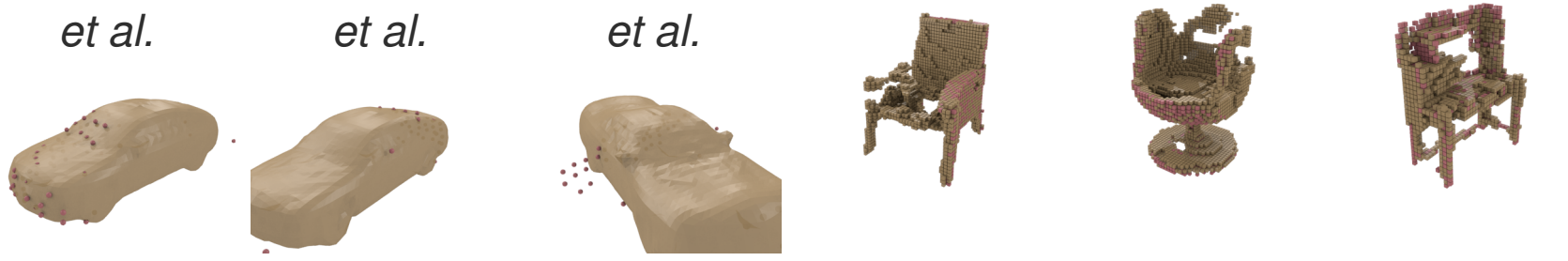
Exotic shapes, not part of the latent shape space.

AML GT AML GT AML GT

Fine structures and details.

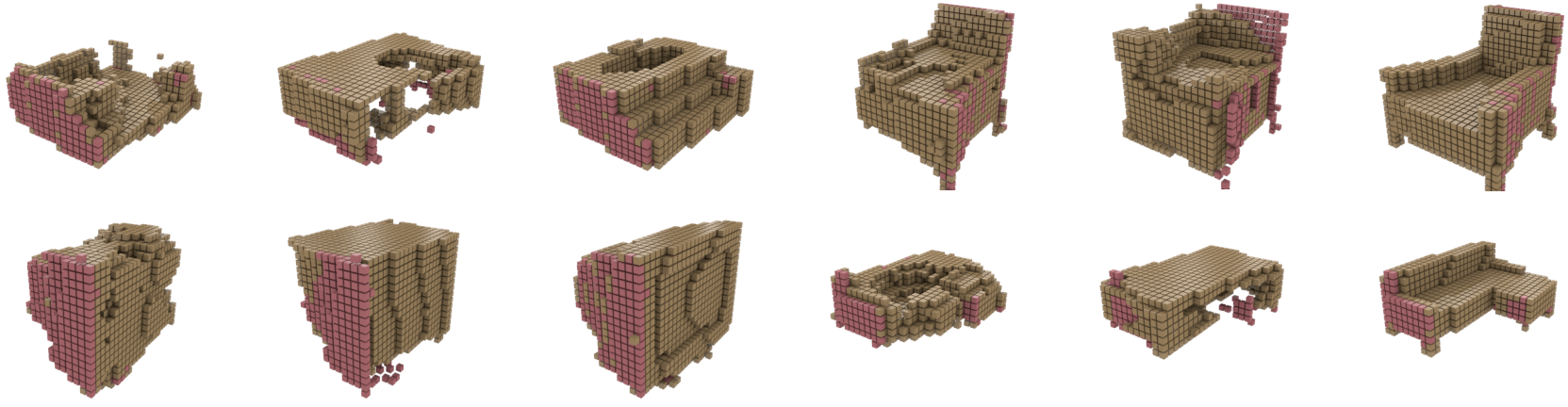
Appendix – Failure Cases

Engelmann *et al.* | Engelmann *et al.* | Engelmann *et al.* | Dai *et al.* | Dai *et al.* | Dai *et al.*



[Engelmann et al., 2016] over-estimates size and overfits noise; [Dai et al., 2017] has difficulties with details.

Dai *et al.* | AML | GT | Dai *et al.* | AML | GT

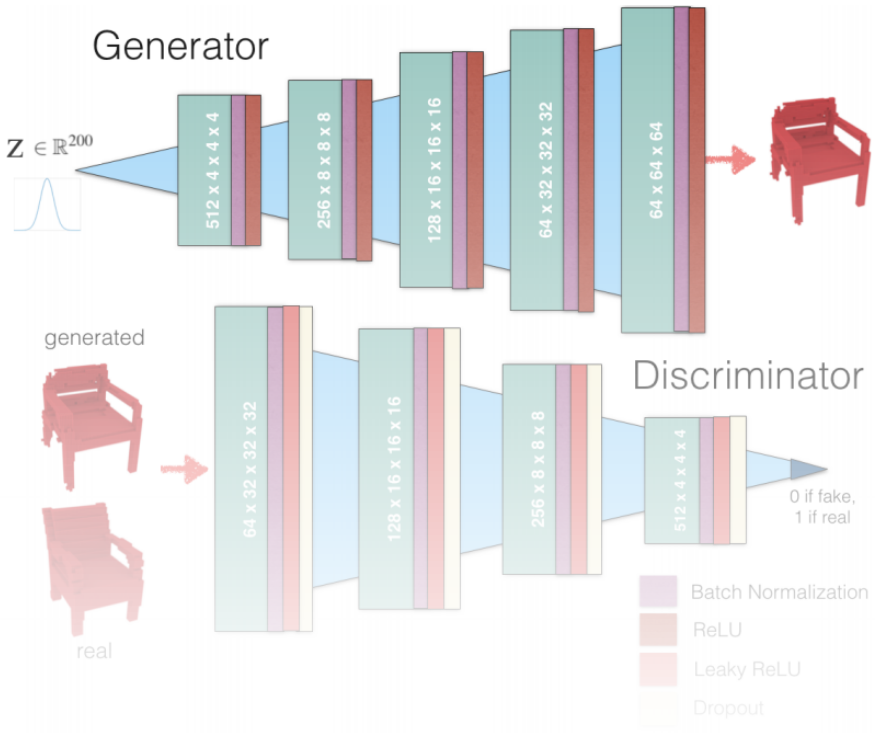


Identifying object categories from sparse observations.

Appendix – Future Work

Learn better shape models/priors.

- Efficiently learn more expressive, generative models, *e.g.*, VAEs, GANs ...

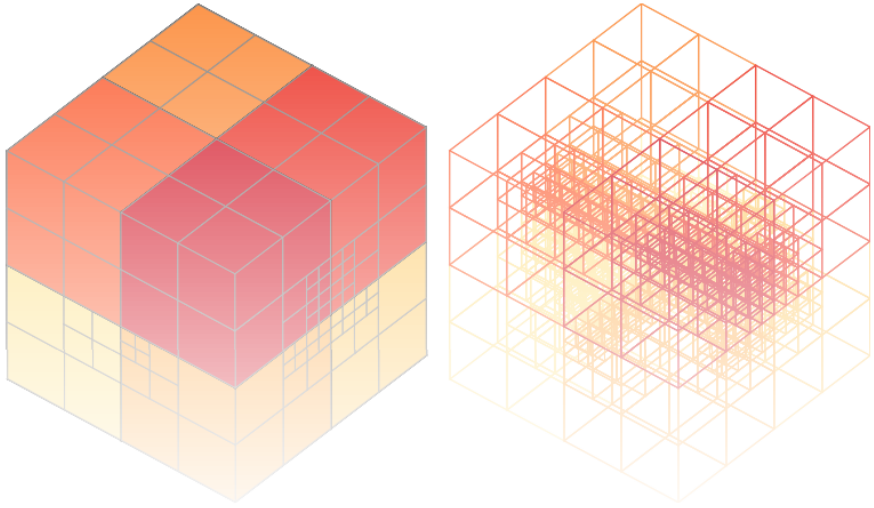


[Liu et al., 2017] [Smith and Meger, 2017]

Appendix – Future Work

Use efficient shape representations allowing higher resolution.

- Efficient training in high resolutions is essential ...



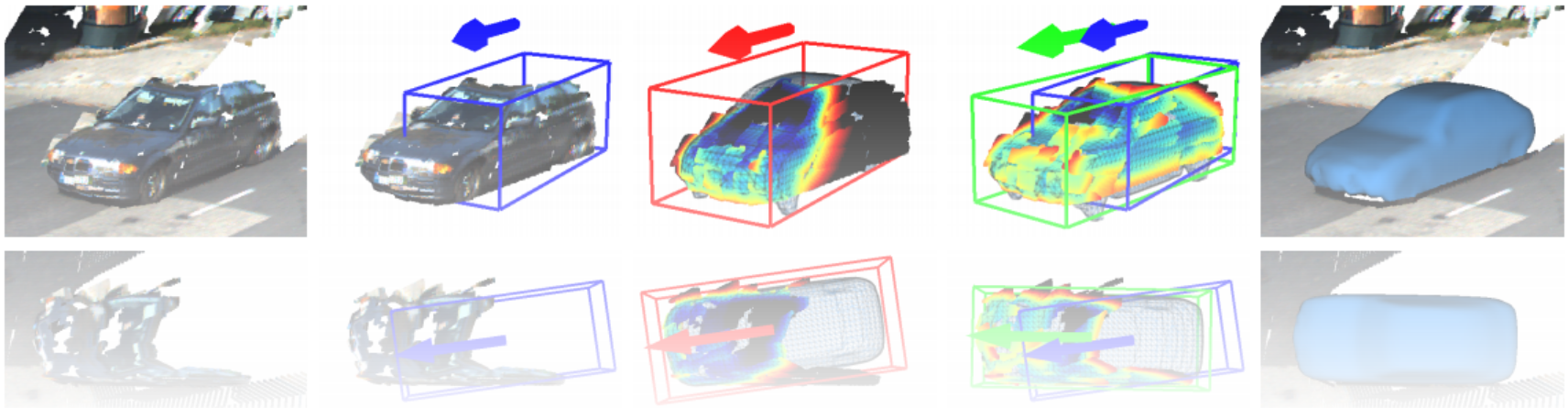
[Riegler et al., 2017b]

[Fan et al., 2016]

Appendix – Future Work

Integrate pose estimation/refinement, *e.g.*, as in [Engelmann et al., 2016].

- Currently, reference shapes need to be roughly aligned to real observations ...



[Engelmann et al., 2016]

References I

Bao, S., Chandraker, M., Lin, Y., and Savarese, S. (2013).

Dense object reconstruction with semantic priors.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Dai, A., Qi, C. R., and Nießner, M. (2016).

Shape completion using 3d-encoder-predictor cnns and shape synthesis.

arXiv.org, abs/1612.00101.

Dai, A., Qi, C. R., and Nießner, M. (2017).

Shape completion using 3d-encoder-predictor cnns and shape synthesis.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Dame, A., Prisacariu, V., Ren, C., and Reid, I. (2013).

Dense reconstruction using 3D object shape priors.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

References II

Engelmann, F., Stückler, J., and Leibe, B. (2016).

Joint object pose estimation and shape reconstruction in urban street scenes using 3D shape priors.

In *Proc. of the German Conference on Pattern Recognition (GCPR)*.

Engelmann, F., Stückler, J., and Leibe, B. (2017).

SAMP: shape and motion priors for 4d vehicle reconstruction.

In *Proc. of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 400–408.

Fan, H., Su, H., and Guibas, L. J. (2016).

A point set generation network for 3d object reconstruction from a single image.

arXiv.org, abs/1612.00603.

References III

Firman, M., Mac Aodha, O., Julier, S., and Brostow, G. J. (2016).

Structured prediction of unobserved voxels from a single depth image.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Gupta, S., Arbeláez, P. A., Girshick, R. B., and Malik, J. (2015).

Aligning 3D models to RGB-D images of cluttered scenes.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Gwak, J., Choy, C. B., Garg, A., Chandraker, M., and Savarese, S. (2017).

Weakly supervised generative adversarial networks for 3d reconstruction.

arXiv.org, 1705.10904.

Haene, C., Savinov, N., and Pollefeys, M. (2014).

Class specific 3d object shape priors using surface normals.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

References IV

Han, X., Li, Z., Huang, H., Kalogerakis, E., and Yu, Y. (2017).

High-resolution shape completion using deep neural networks for global structure and local geometry inference.

In *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, pages 85–93.

Kroemer, O., Amor, H. B., Ewerton, M., and Peters, J. (2012).

Point cloud completion using extrusions.

In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 680–685.

Law, A. J. and Aliaga, D. G. (2011).

Single viewpoint model completion of symmetric objects for digital inspection.

Computer Vision and Image Understanding (CVIU), 115(5):603–610.

References V

Li, Y., Dai, A., Guibas, L. J., and Nießner, M. (2015).

Database-assisted object retrieval for real-time 3d reconstruction.

Computer Graphics Forum, 34(2):435–446.

Liu, J., Yu, F., and Funkhouser, T. A. (2017).

Interactive 3d modeling with a generative adversarial network.

arXiv.org, abs/1706.05170.

Nguyen, D. T., Hua, B., Tran, M., Pham, Q., and Yeung, S. (2016).

A field model for repairing 3d shapes.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 5676–5684.

References VI

Pauly, M., Mitra, N. J., Giesen, J., Gross, M. H., and Guibas, L. J. (2005).

Example-based 3d scan completion.

In *Eurographics Symposium on Geometry Processing (SGP)*.

Pauly, M., Mitra, N. J., Wallner, J., Pottmann, H., and Guibas, L. J. (2008).

Discovering structural regularity in 3d geometry.

ACM Trans. on Graphics, 27(3):43:1–43:11.

Rezende, D. J., Eslami, S. M. A., Mohamed, S., Battaglia, P., Jaderberg, M., and Heess, N. (2016).

Unsupervised learning of 3d structure from images.

arXiv.org, 1607.00662.

References VII

Riegler, G., Ulusoy, A. O., Bischof, H., and Geiger, A. (2017a).

OctNetFusion: Learning depth fusion from data.

In *Proc. of the International Conf. on 3D Vision (3DV)*.

Riegler, G., Ulusoy, A. O., and Geiger, A. (2017b).

Octnet: Learning deep 3d representations at high resolutions.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Sharma, A., Grau, O., and Fritz, M. (2016).

Vconv-dae: Deep volumetric shape learning without object labels.

arXiv.org, 1604.03755.

Smith, E. and Meger, D. (2017).

Improved adversarial systems for 3d object generation and reconstruction.

arXiv.org, abs/1707.09557.

References VIII

Thrun, S. and Wegbreit, B. (2005).

Shape from symmetry.

In *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, pages 1824–1831.

Wu, J., Zhang, C., Xue, T., Freeman, B., and Tenenbaum, J. (2016).

Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling.

In *Advances in Neural Information Processing Systems (NIPS)*.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J. (2015).

3d shapenets: A deep representation for volumetric shapes.

In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

References IX

Yang, B., Rosa, S., Markham, A., Trigoni, N., and Wen, H. (2018).

3d object dense reconstruction from a single depth view.

arXiv.org, abs/1802.00411.

Yang, B., Wen, H., Wang, S., Clark, R., Markham, A., and Trigoni, N. (2017).

3d object reconstruction from a single depth view with adversarial learning.

arXiv.org, 1708.07969.

Zheng, Q., Sharf, A., Wan, G., Li, Y., Mitra, N. J., Cohen-Or, D., and Chen, B. (2010).

Non-local scan consolidation for 3d urban scenes.

ACM Trans. on Graphics, 29(4):94:1–94:9.