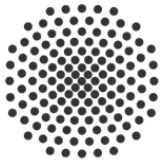
SIMPLE: A Gradient Estimator for *k*-Subset Sampling

Kareem Ahmed*[Δ], Zhe Zeng*[Δ], Mathias Niepert[†], Guy Van den Broeck[Δ]

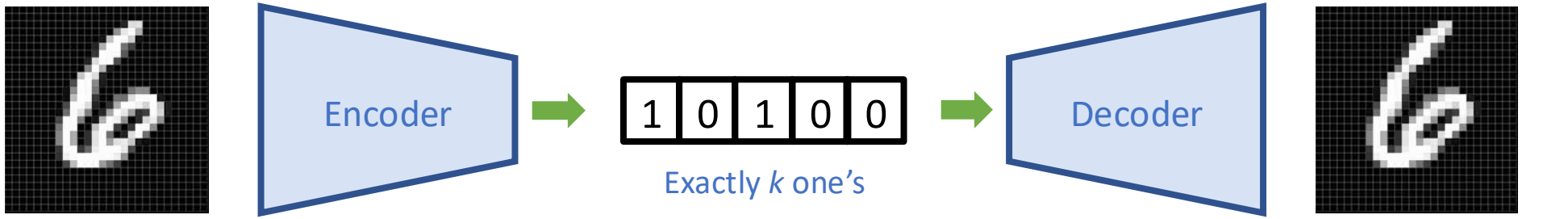
[Δ]UCLA [†]Stuttgart University



University of Stuttgart

Why k-Subset Distributions?

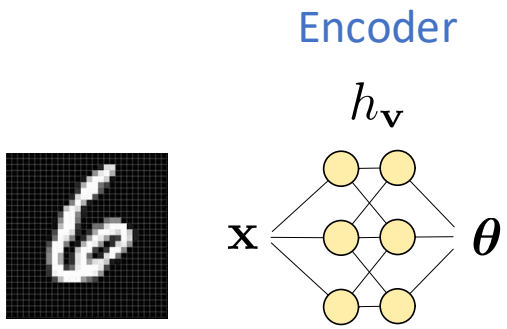
Example. Discrete VAE



Example. Learn to Explain (L2X)

Key Words (*k* = 10)	**Taste Score**
a lite bodied beer with a pleasant taste. was like a reddish color. a little like wood and caramel with a hop finish. has a sort of fruity flavor like grapes or cherry that is sort of buried in there. mouth feel was lite, sort of bubbly. not hard to down, though a bit harder then one would expect given the taste.	0.7

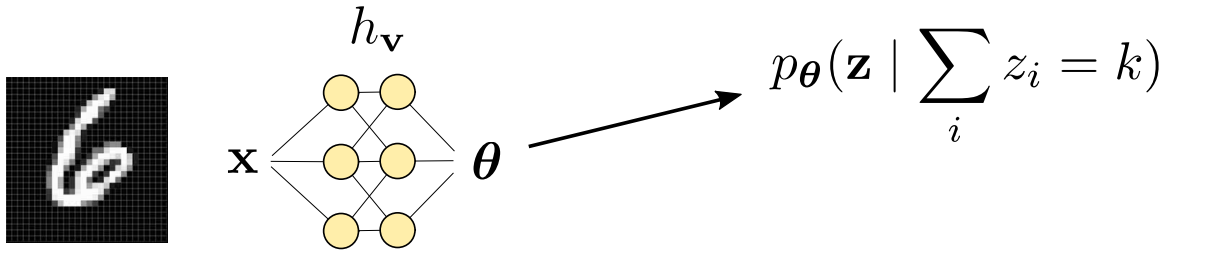
SIMPLE: Gradient Estimator for *k*-Subset Sampling



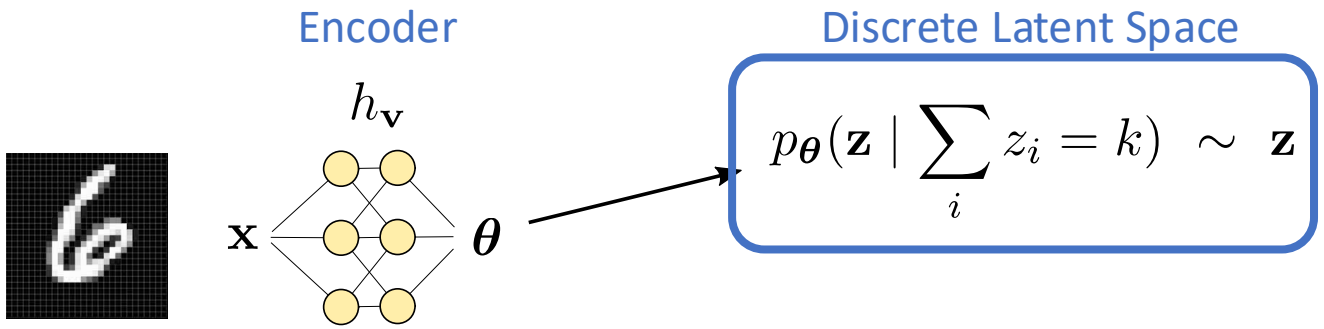
SIMPLE: Gradient Estimator for *k*-Subset Sampling

Encoder

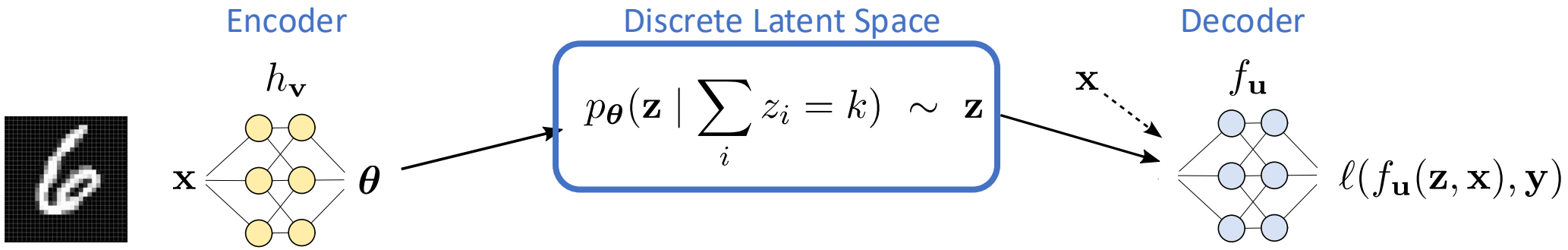
Discrete Latent Space



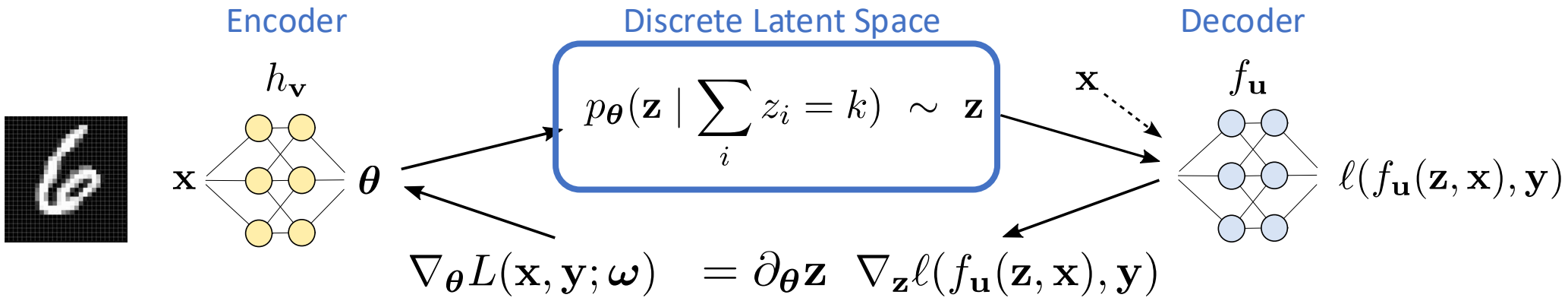
SIMPLE: Gradient Estimator for *k*-Subset Sampling



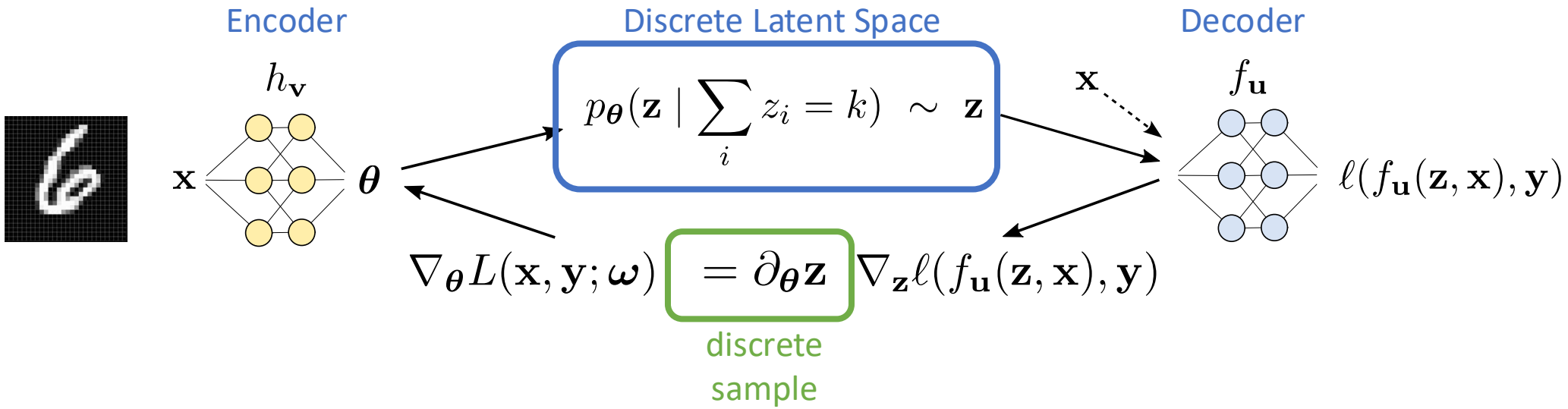
SIMPLE: Gradient Estimator for *k*-Subset Sampling



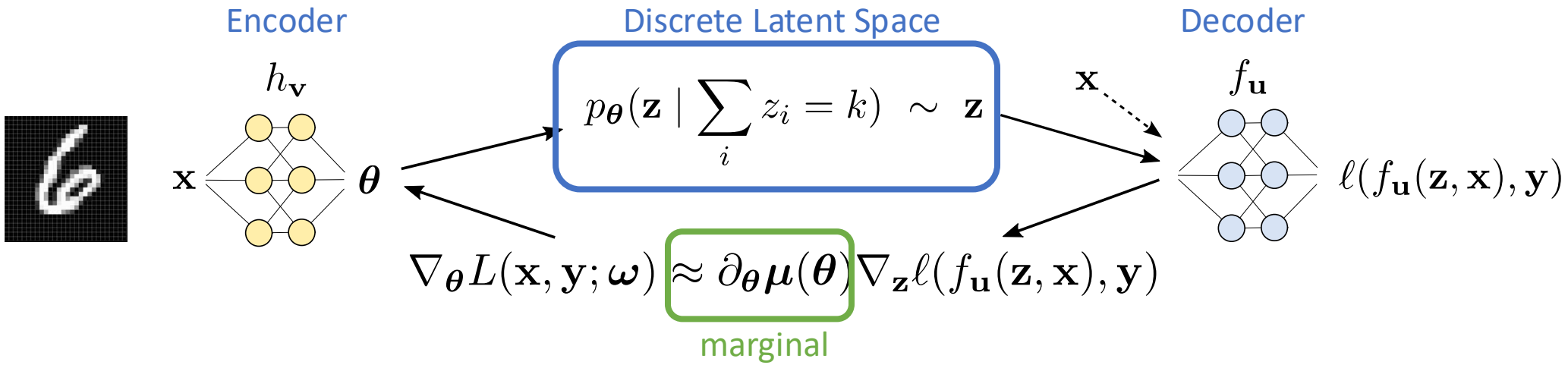
SIMPLE: Gradient Estimator for *k*-Subset Sampling



SIMPLE: Gradient Estimator for *k*-Subset Sampling

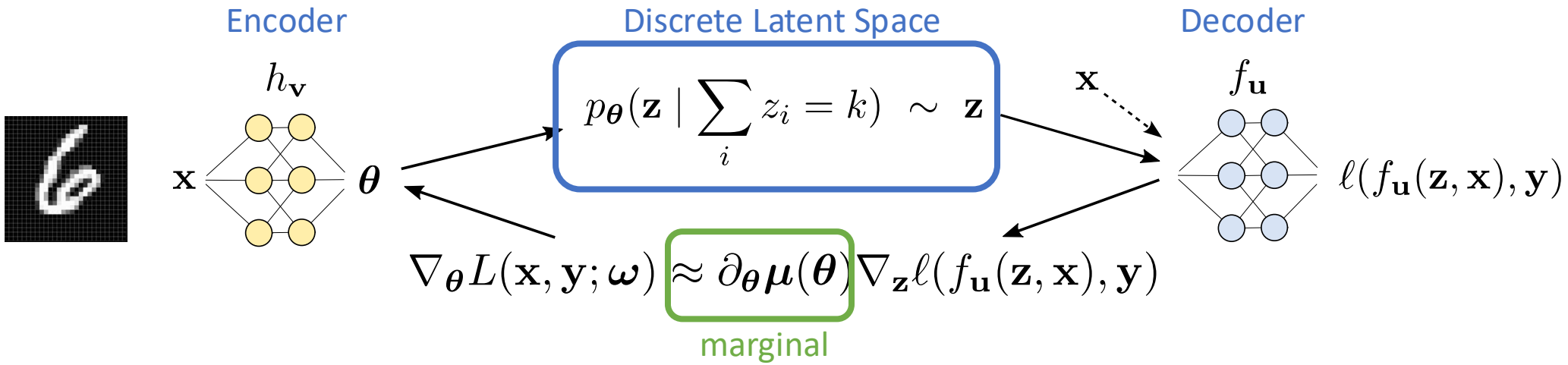


SIMPLE: Gradient Estimator for *k*-Subset Sampling

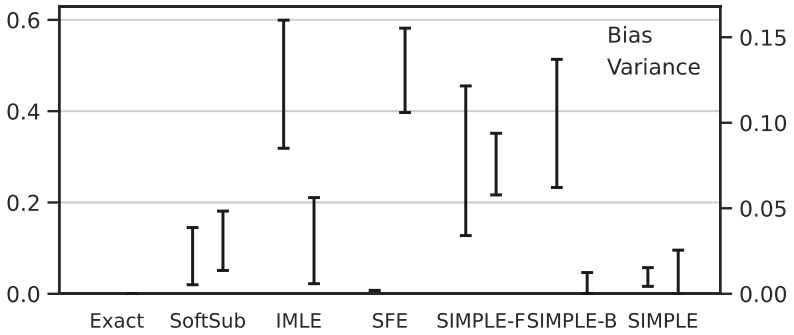


Prop. conditional marginals can be obtained by $p_{\boldsymbol{\theta}}\left(z_i \mid \sum_j z_j = k\right) = \dfrac{\partial}{\partial \theta_i} \log p_{\boldsymbol{\theta}}(\sum_j z_j = k).$

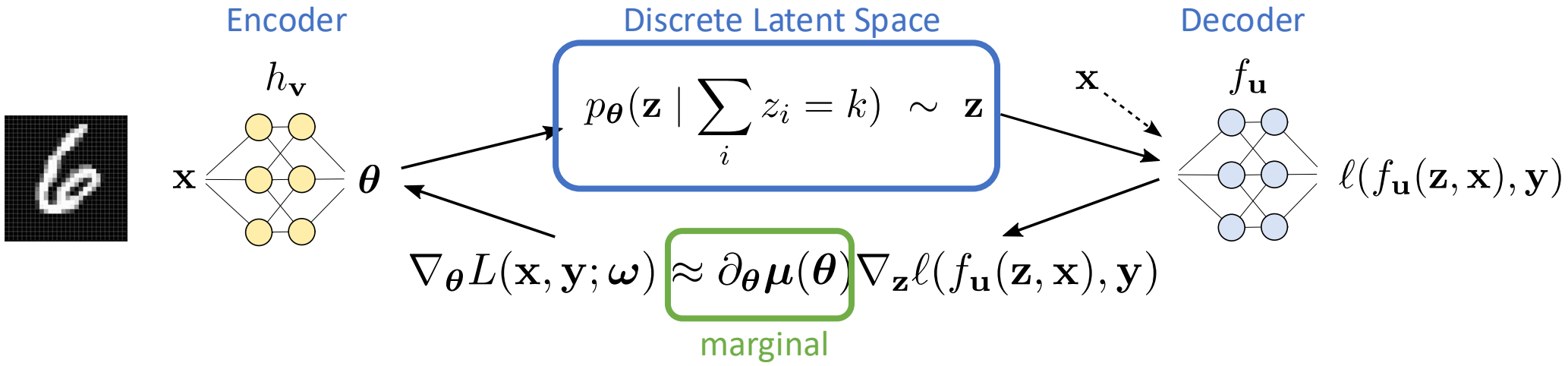
SIMPLE: Gradient Estimator for *k*-Subset Sampling



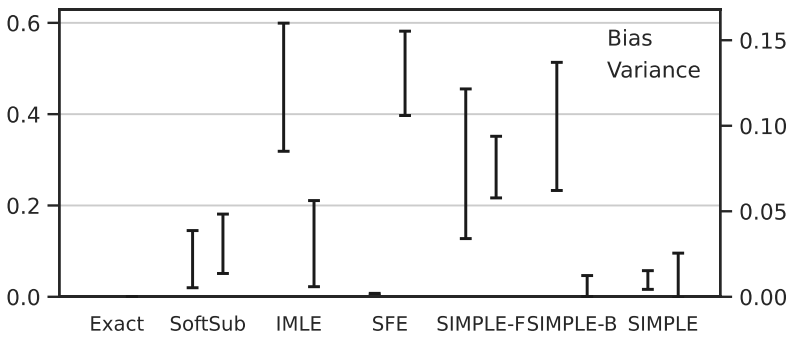
We achieve ***lower bias and variance*** by exact, discrete samples and exact derivative of conditional marginals.

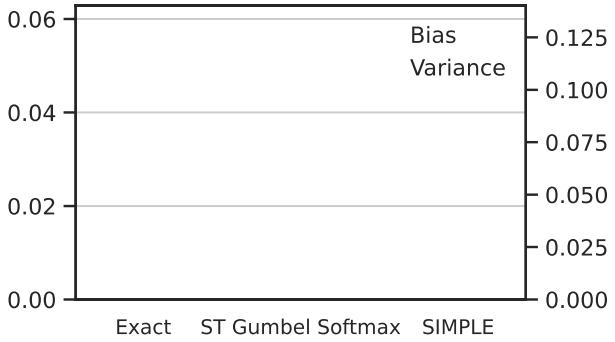


SIMPLE: Gradient Estimator for *k*-Subset Sampling

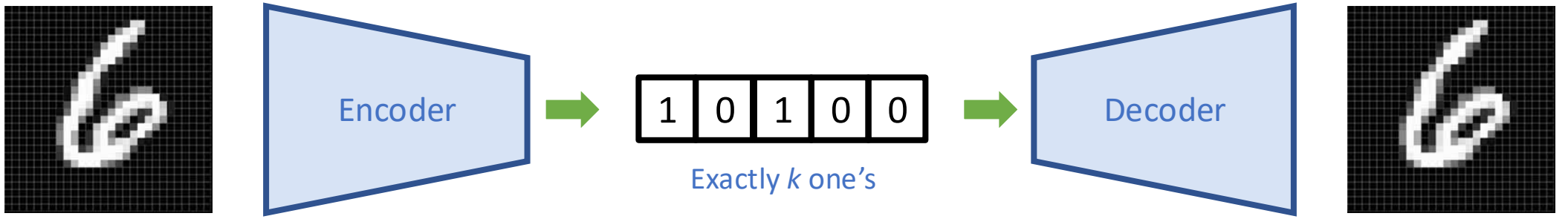


We achieve ***lower bias and variance*** by exact, discrete samples and exact derivative of conditional marginals.

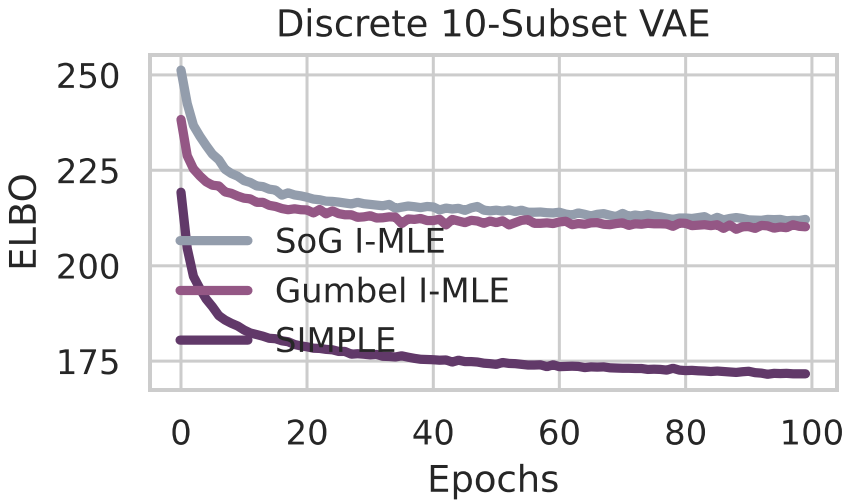


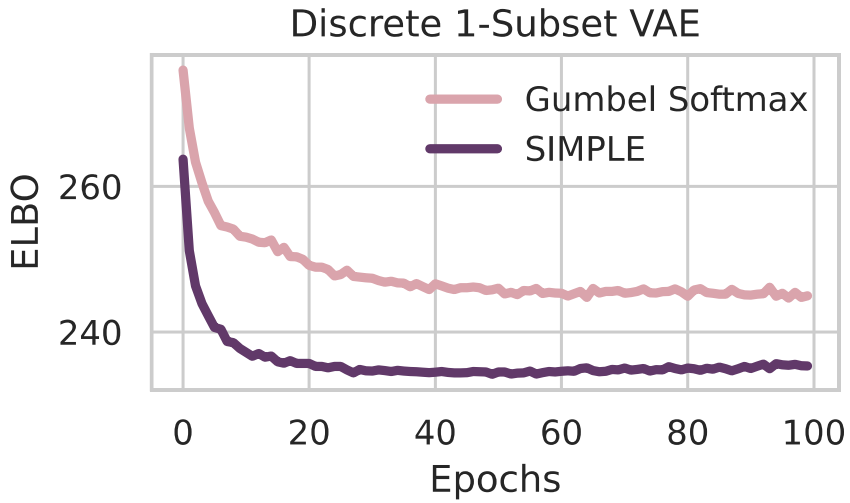


Experiment: Discrete VAE



Metric: *exact* ELBO





Experiment: Learn to Explain (L2X)

Key Words (*k* = 10)	**Taste Score**
a lite bodied beer with a pleasant taste. was like a reddish color. a little like wood and caramel with a hop finish. has a sort of fruity flavor like grapes or cherry that is sort of buried in there. mouth feel was lite, sort of bubbly. not hard to down, though a bit harder then one would expect given the taste.	0.7

Results for three aspects with $k = 10$

Method	Appearance		Palate		Taste	
	Test MSE	Precision	Test MSE	Precision	Test MSE	Precision
SIMPLE (Ours)	**2.35 ± 0.28**	**66.81 ± 7.56**	**2.68 ± 0.06**	**44.78 ± 2.75**	**2.11 ± 0.02**	**42.31 ± 0.61**
L2X (t = 0.1)	10.70 ± 4.82	30.02 ± 15.82	6.70 ± 0.63	**50.39 ± 13.58**	6.92 ± 1.61	32.23 ± 4.92
SoftSub (t = 0.5)	**2.48 ± 0.10**	52.86 ± 7.08	2.94 ± 0.08	39.17 ± 3.17	2.18 ± 0.10	**41.98 ± 1.42**
I-MLE (τ = 30)	**2.51 ± 0.05**	**65.47 ± 4.95**	2.96 ± 0.04	40.73 ± 3.15	2.38 ± 0.04	**41.38 ± 1.55**

Results for aspect Aroma, for k in {5, 10, 15}

Method	$k = 5$		$k = 10$		$k = 15$	
	Test MSE	Precision	Test MSE	Precision	Test MSE	Precision
SIMPLE (Ours)	**2.27 ± 0.05**	**57.30 ± 3.04**	**2.23 ± 0.03**	**47.17 ± 2.11**	3.20 ± 0.04	**53.18 ± 1.09**
L2X (t = 0.1)	5.75 ± 0.30	33.63 ± 6.91	6.68 ± 1.08	26.65 ± 9.39	7.71 ± 0.64	23.49 ± 10.93
SoftSub (t = 0.5)	2.57 ± 0.12	**54.06 ± 6.29**	2.67 ± 0.14	44.44 ± 2.27	**2.52 ± 0.07**	37.78 ± 1.71
I-MLE (τ = 30)	2.62 ± 0.05	**54.76 ± 2.50**	2.71 ± 0.10	**47.98 ± 2.26**	2.91 ± 0.18	39.56 ± 2.07

Thanks!

- code: github.com/UCLA-StarAI/SIMPLE