LEARNING TO DISCOVER

AI and physics conference, Wednesday 27th of April 2022

LEARNING TO DISCOVER

AI and physics conference, Wednesday 27th of April 2022

"Automatic selection of latent variables in Variational Auto-Encoders"

Who are we ?

EMMA JOUFFROY

2nd year PhD student in machine learning, in collaboration with CEA CESTA and IMS

YANNICK BERTHOUMIEU

Since 2007 Yannick Berthoumieu is a full Professor of the Bordeaux Institute of Technology, where he currently a member of the Signal and Image Processing Group (GSI) belonging to the CNRS IMS Laboratory.

AUDREY GIREMUS

HdR and university professor, Audrey Giremus is currently teaching signal processing at Bordeaux University and is at the head of the signal & image group of IMS (CNRS - UMR 5218)

SUMMARY

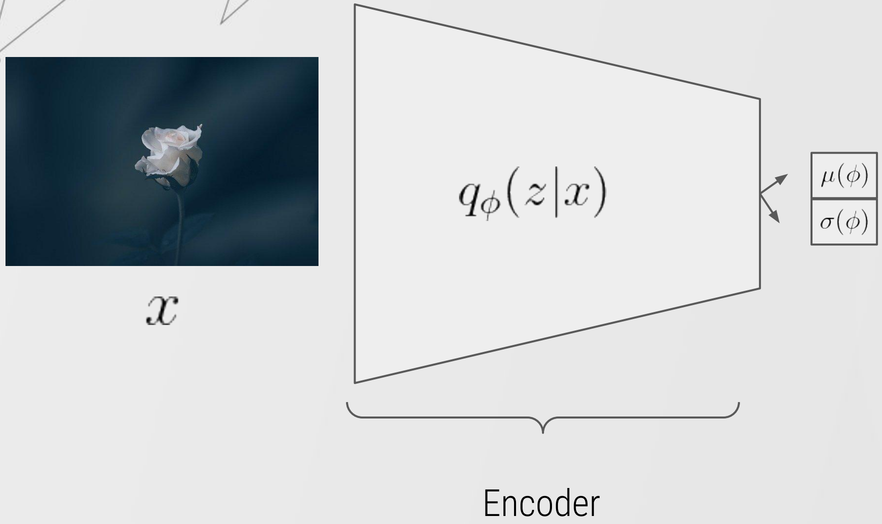
01

INTRODUCTION

Variational autoencoders

VARIATIONAL AUTOENCODER

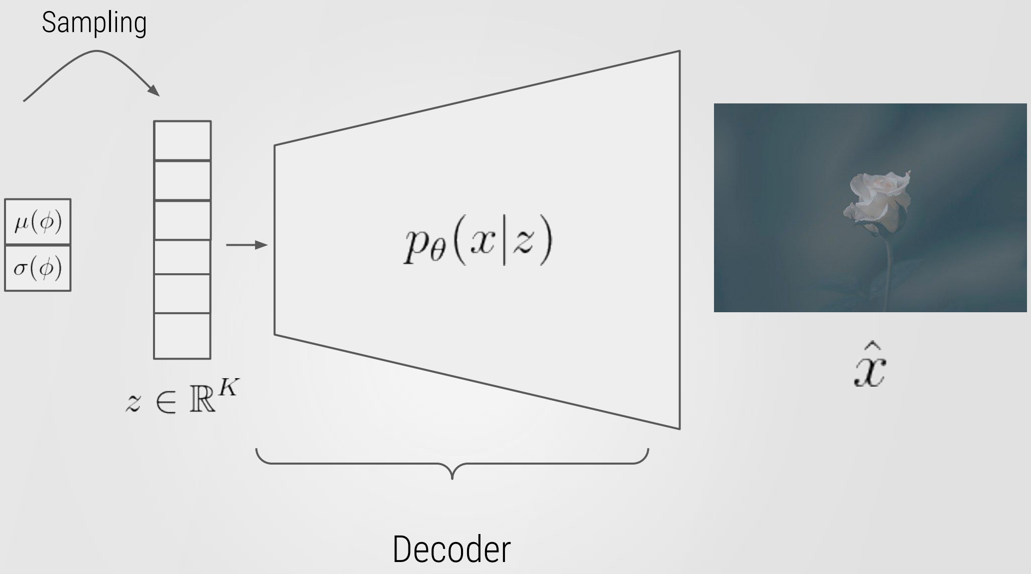
A generative model based on a latent space



Kingma & Welling 2013

VARIATIONAL AUTOENCODER

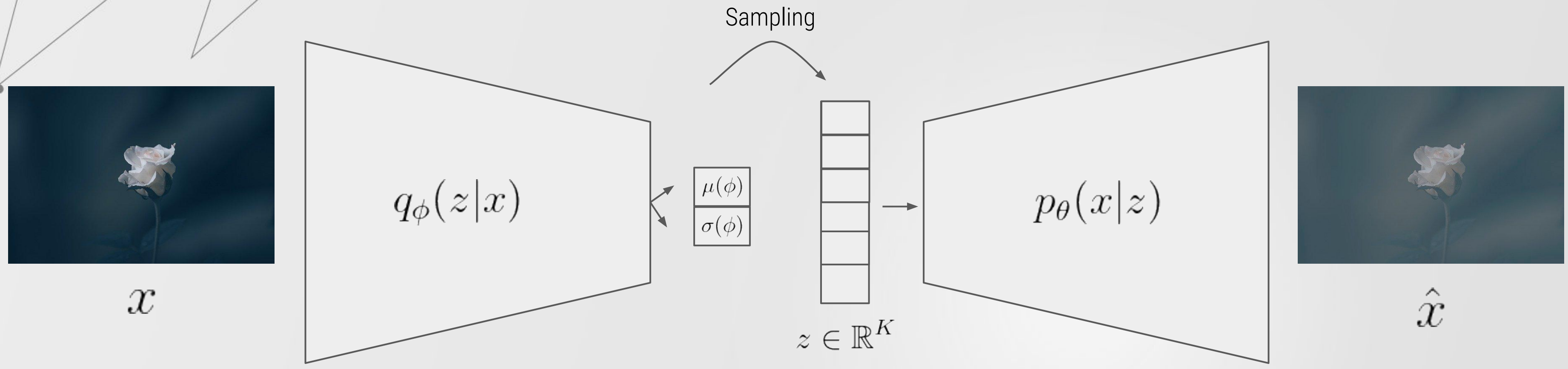
A generative model based on a latent space



Kingma & Welling 2013

VARIATIONAL AUTOENCODER

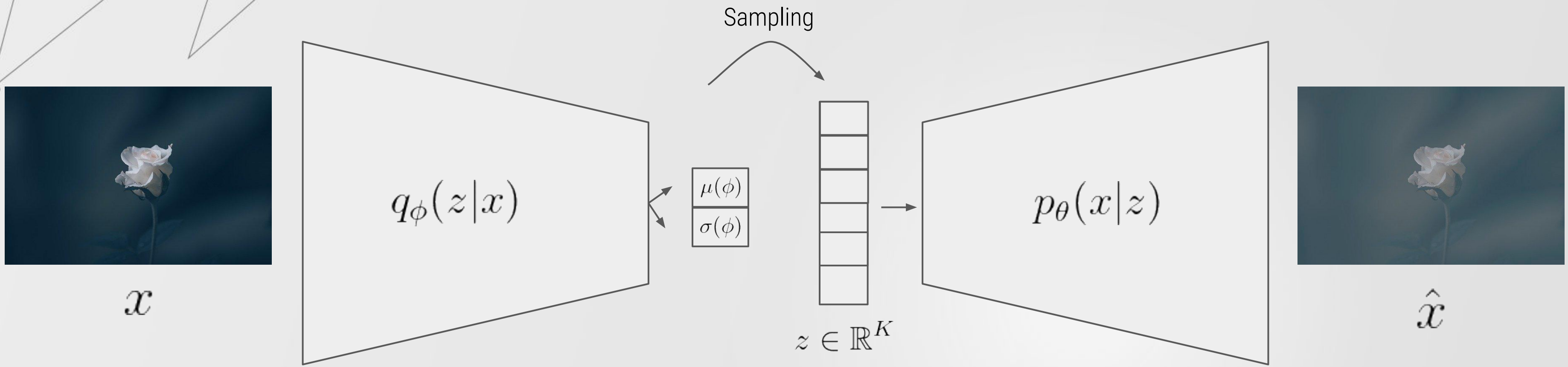
A generative model based on a latent space



Kingma & Welling 2013

VARIATIONAL AUTOENCODER

A generative model based on a latent space

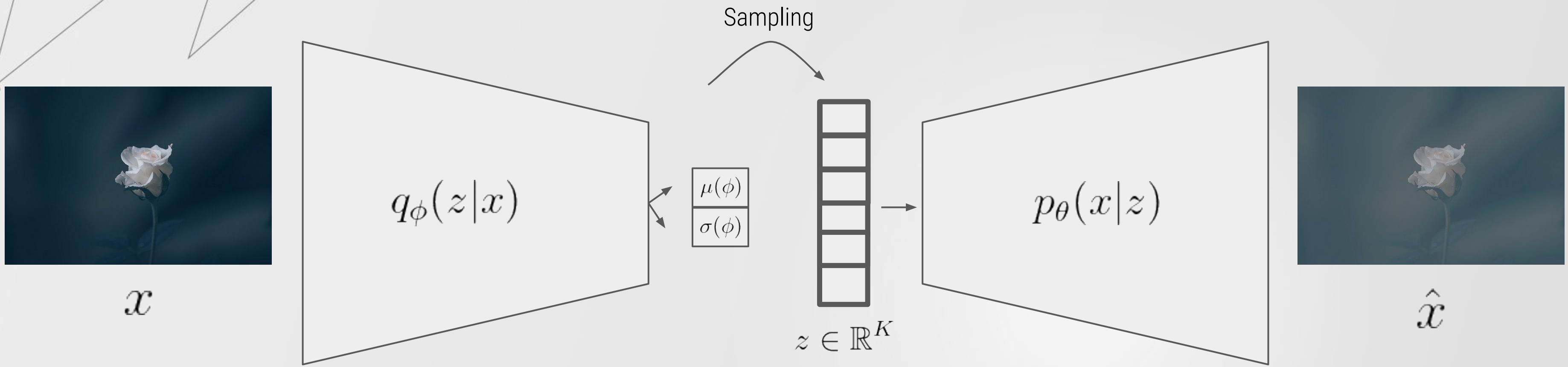


$$-ELBO = \mathcal{L}(\theta, \phi) = \mathbb{E}_{x \sim p(x)} \left[\mathbb{E}_{z \sim q_\phi(z|x)} [-\log p(x|z)] + \sum_{i=1}^{K} \mathrm{KL} \left[q_\phi(z_k|x) || p(z_k) \right] \right]$$

Kingma & Welling 2013

VARIATIONAL AUTOENCODER

A generative model based on a latent space



Kingma & Welling 2013

VARIATIONAL AUTOENCODER

And informative components

z_1 → Horizontal position

$z \in \mathbb{R}^K$

VARIATIONAL AUTOENCODER

And informative components

z_1 → Horizontal position

z_2 → Vertical position

$z \in \mathbb{R}^K$

VARIATIONAL AUTOENCODER

And informative components

$z_1 \longrightarrow$ Horizontal position

$z_2 \longrightarrow$ Vertical position

$z_3 \longrightarrow$ Color

$$z \in \mathbb{R}^K$$

VARIATIONAL AUTOENCODER

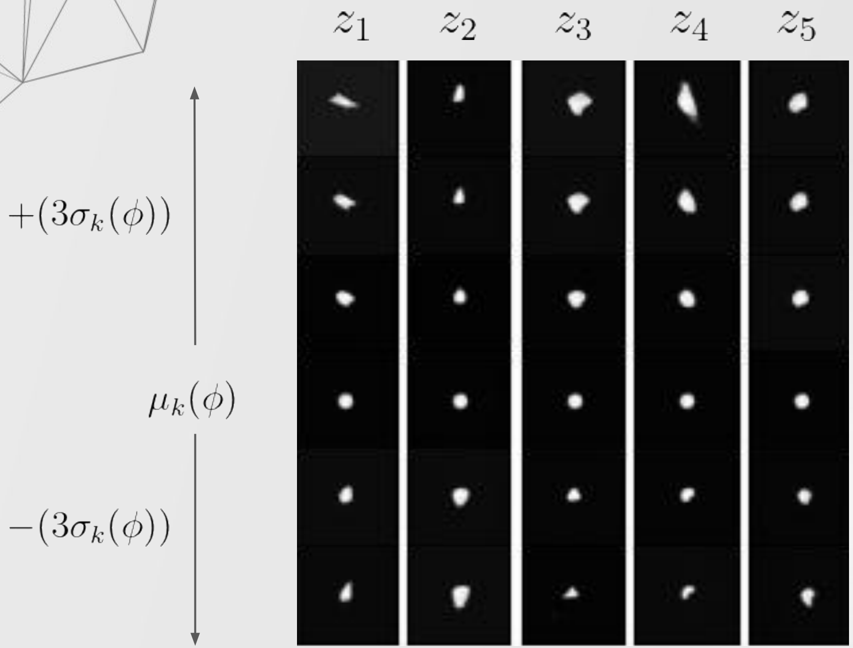
And informative components

z_1	→ Horizontal position
z_2	→ Vertical position
z_3	→ Color
...	
z_{K-1}	
z_K	

$$z \in \mathbb{R}^K$$

VARIATIONAL AUTOENCODER

And disentanglement



Each latent variable is informative, and entangled with the others

LIMITATIONS

Of the Variational Autoencoder

Information

Each latent variable is informative

Entanglement

Each latent variable is correlated with each other

Simple priors

The commonly used distribution for the prior law is a gaussian with identity covariance

Lower bound of the evidence

The maximized criterion is a lower bound of the true distribution

LIMITATIONS

Of the Variational Autoencoder

Information

Each latent variable is informative

Entanglement

Each latent variable is correlated with each other

Simple priors

The commonly used distribution for the prior law is a gaussian with identity covariance

Lower bound of the evidence

The maximized criterion is a lower bound of the true distribution

VARIATIONAL AUTOENCODER

And one alternative

$$\mathcal{L}(\theta, \phi) = \mathbb{E}_{x \sim p(x)} \left[\mathbb{E}_{z \sim q_{\phi}(z|x)} [-\log p(x|z)] + \beta \sum_{i=1}^{K} \mathrm{KL} \left[q_{\phi}(z_k|x) || p(z_k) \right] \right]$$

With

$$\beta > 1$$

Higgins & al. 2013

VARIATIONAL AUTOENCODER

And one alternative

$$\mathcal{L}(\theta, \phi) = \mathbb{E}_{x\sim p(x)}\left[\mathbb{E}_{z\sim q_\phi(z|x)}[-\log p(x|z)] + \boxed{\beta}\sum_{i=1}^{K} \mathrm{KL}\left[q_\phi(z_k|x)||p(z_k)\right]\right]$$

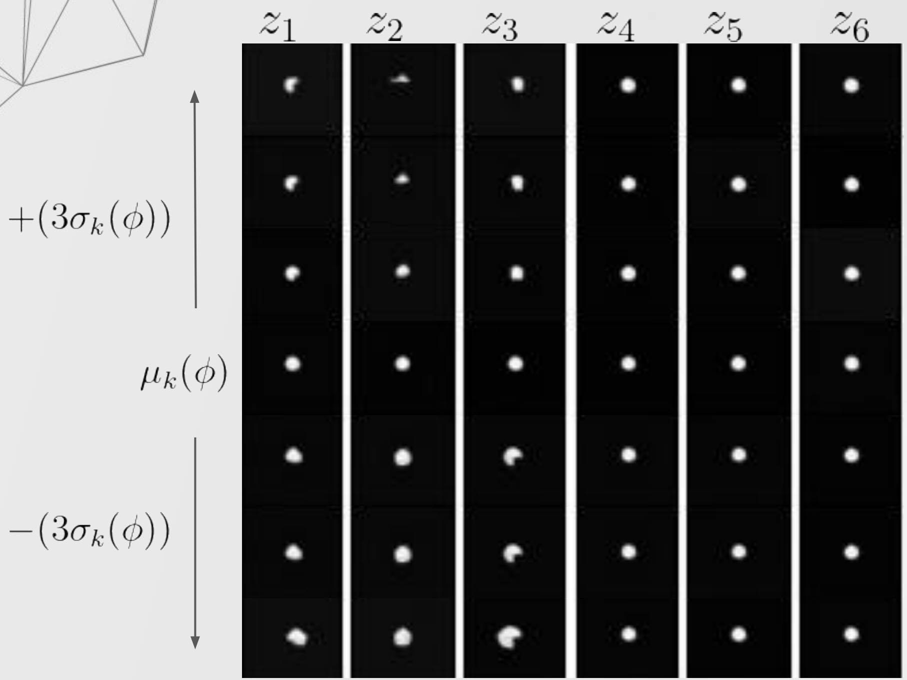
With

$$\beta > 1$$

Higgins & al. 2013

VARIATIONAL AUTOENCODER

And one alternative



Better separation between informative and uninformative components, better disentanglement, but at the expense of the reconstruction quality

LIMITATIONS

From the latter alternative

$$\mathbb{E}_{x\sim p(x)}\left[\beta \mathrm{KL}[q_\phi(z|x)||p(z)]\right]$$

$$= \beta \mathrm{I}_q(z;x) + \beta \mathrm{KL}[q(z)||p(z)]$$

Kim & al. 2018

LIMITATIONS

From the latter alternative

$$\mathbb{E}_{x\sim p(x)}\left[\beta \mathrm{KL}[q_\phi(z|x)||p(z)]\right]$$

$$= \boxed{\beta \mathrm{I}_q(z;x)} + \beta \mathrm{KL}[q(z)||p(z)]$$

Kim & al. 2018

02
ASSUMPTIONS & OBJECTIVES

Inferred variances and informative components

OBJECTIVES

For the proposed modeling

Adjusting latent space

Auto adjustment of the latent space size depending on the data complexity

By not adding a weighting term that could impact the reconstruction quality

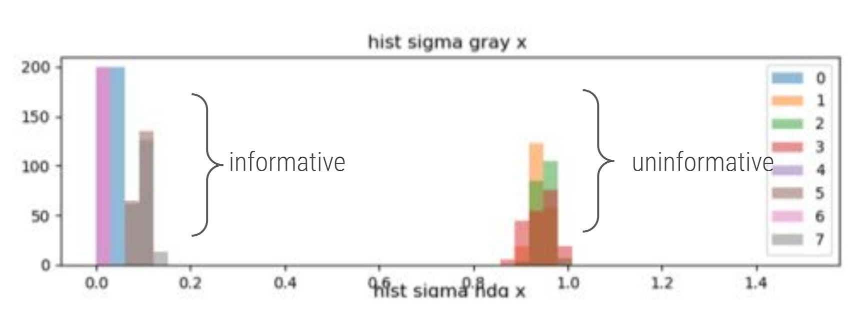
No compromise on reconstruction

Limit architectural complexity

By restraining the number of learnable parameters

ASSUMPTIONS
Underlying the current work

Informative components tend to have lower learned variances values



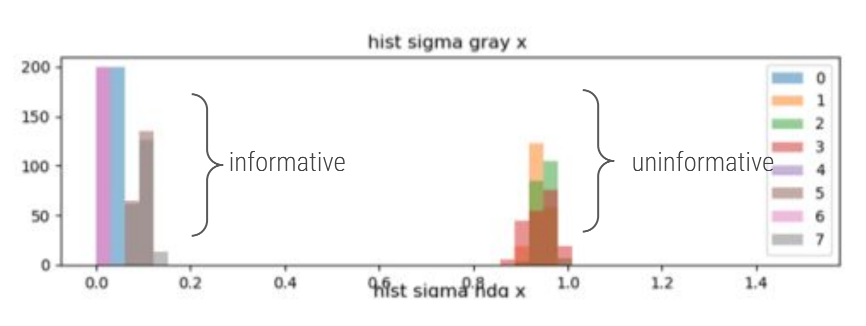
Histogram of inferred variance over 200 samples of a learned beta-VAE with beta = 150

ASSUMPTIONS

Underlying the current work

Informative components tend to have lower learned variances values

Through a hierarchical bayesian modeling, the learned variances could be forced to be either high or small depending on information contained in the latent variable



Histogram of inferred variance over 200 samples of a learned beta-VAE with beta = 150

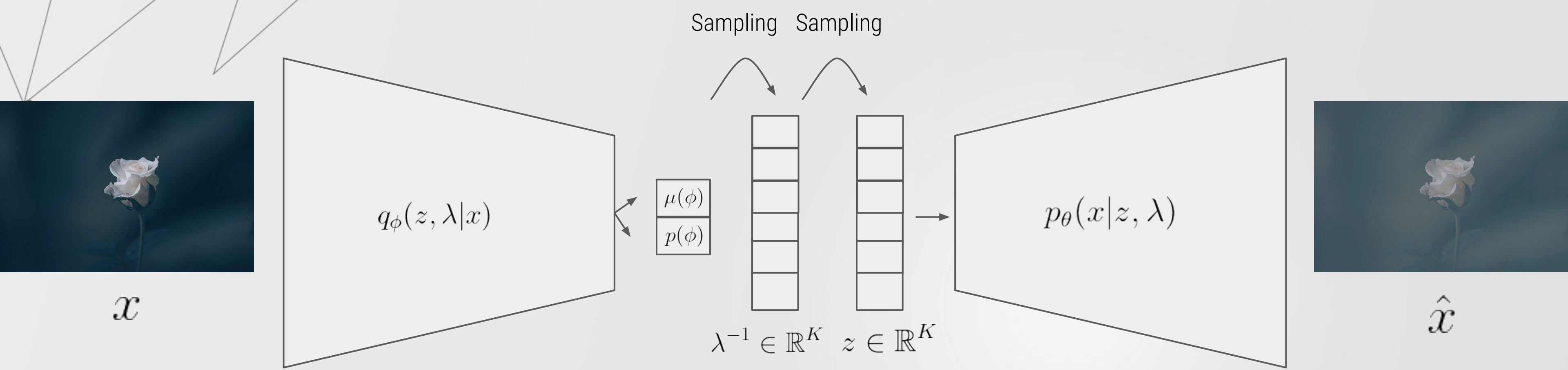
03

PROPOSED MODELING

A bayesian hierarchy based on variational autoencoder

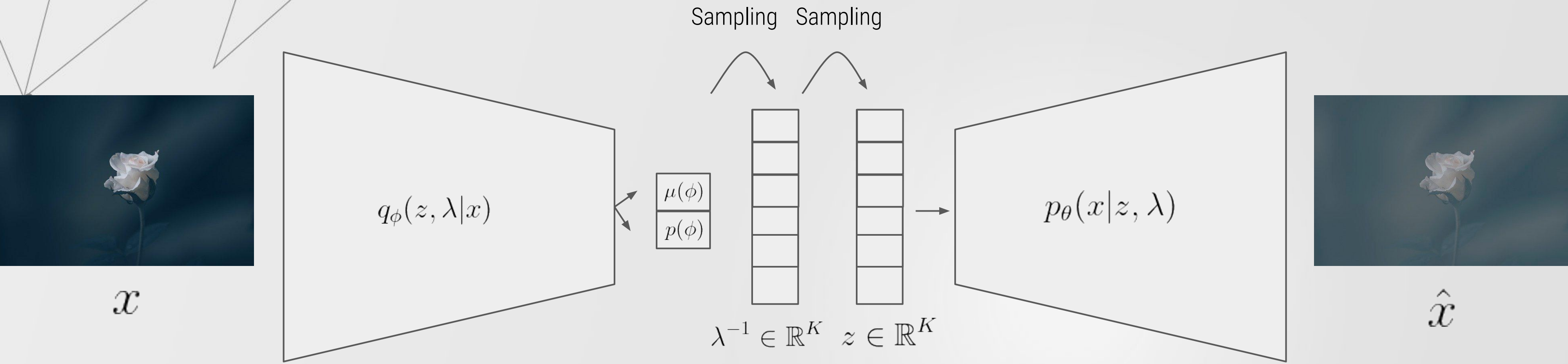
PROPOSED ARCHITECTURE

The Normal-Gamma Variational-Autoencoder
(NGVAE)



PROPOSED ARCHITECTURE

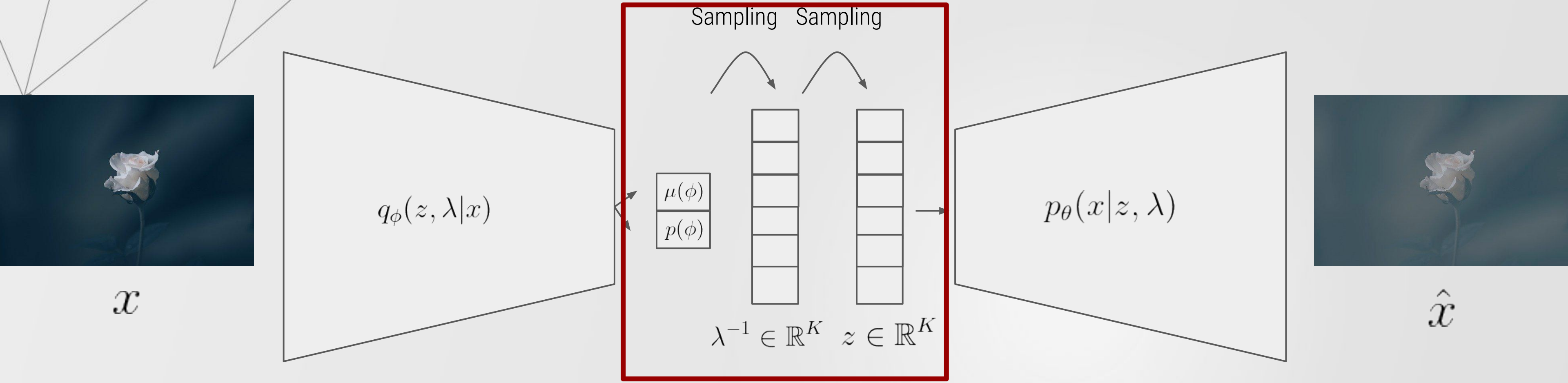
The Normal-Gamma Variational-Autoencoder (NGVAE)



$$q_\phi(z_k, \lambda_k^{-1}) = p_k \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_1, \beta_1) + (1 - p_k) \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

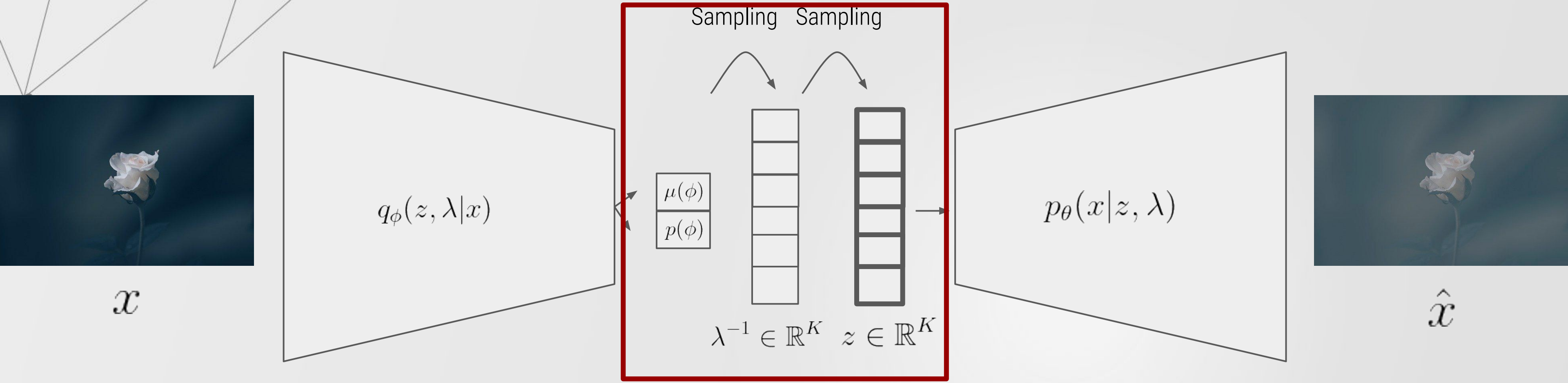
The Normal-Gamma Variational-Autoencoder
(NGVAE)



$$q_\phi(z_k, \lambda_k^{-1}) = p_k \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_1, \beta_1) + (1 - p_k)\mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

The Normal-Gamma Variational-Autoencoder
(NGVAE)



$$q_\phi(z_k, \lambda_k^{-1}) = p_k \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_1, \beta_1) + (1 - p_k)\mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

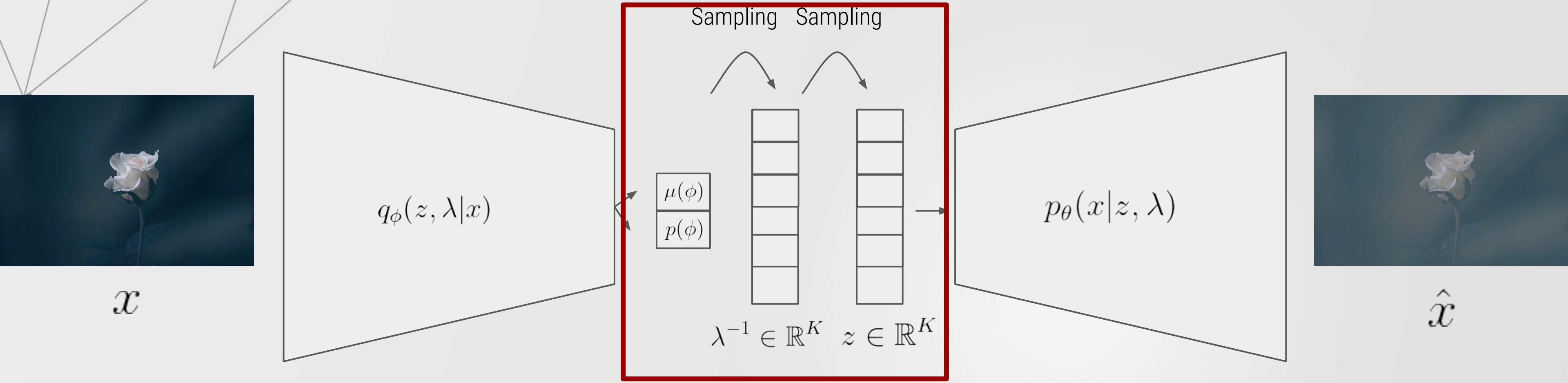
The Normal-Gamma Variational-Autoencoder
(NGVAE)

$$z \in \mathbb{R}^K$$

$$z_k \sim \mathcal{N}(x; \mu_k, \lambda_k^{-1})$$

PROPOSED ARCHITECTURE

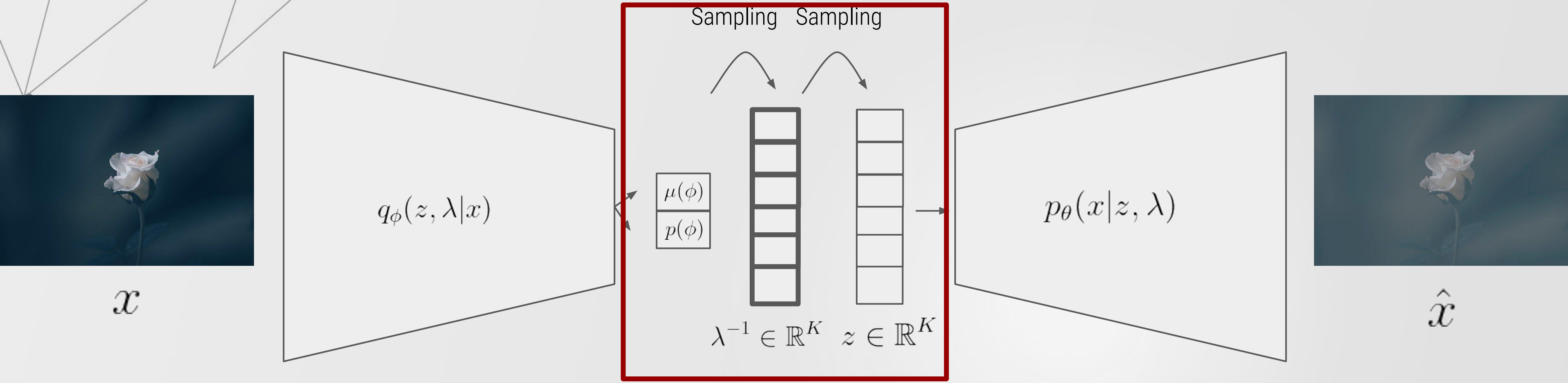
The Normal-Gamma Variational-Autoencoder
(NGVAE)



$$q_\phi(z_k, \lambda_k^{-1}) = p_k \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_1, \beta_1) + (1 - p_k) \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

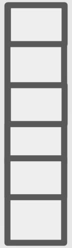
The Normal-Gamma Variational-Autoencoder (NGVAE)



$$q_\phi(z_k, \lambda_k^{-1}) = p_k \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_1, \beta_1) + (1 - p_k) \mathcal{NIG}(z_k, \lambda_k^{-1}; \mu_k, \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

The Normal-Gamma Variational-Autoencoder
(NGVAE)



$\lambda^{-1} \in \mathbb{R}^K$

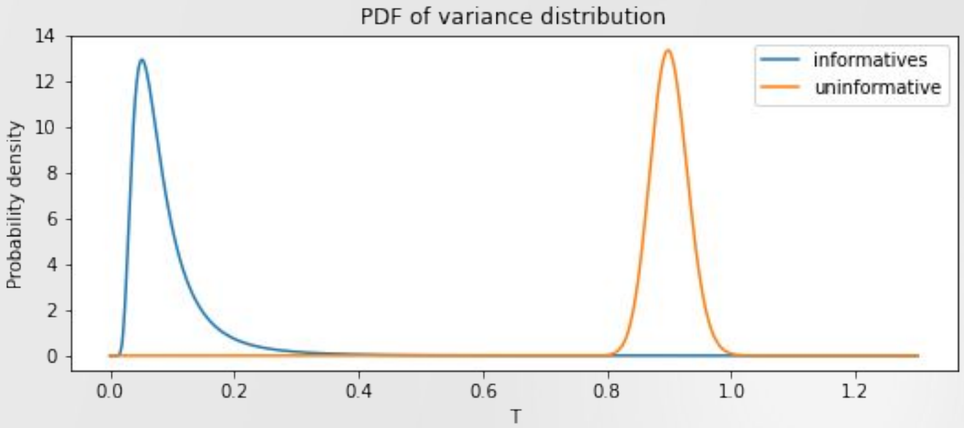
$$\lambda_k^{-1} \sim q_\phi(\lambda_k^{-1}|x)$$

$$q_\phi(\lambda_k^{-1}|x) = p_k \mathcal{IG}(\lambda_k^{-1}; \alpha_1, \beta_1) + (1 - p_k)\mathcal{IG}(\lambda_k^{-1}; \alpha_2, \beta_2)$$

PROPOSED ARCHITECTURE

The Normal-Gamma Variational-Autoencoder

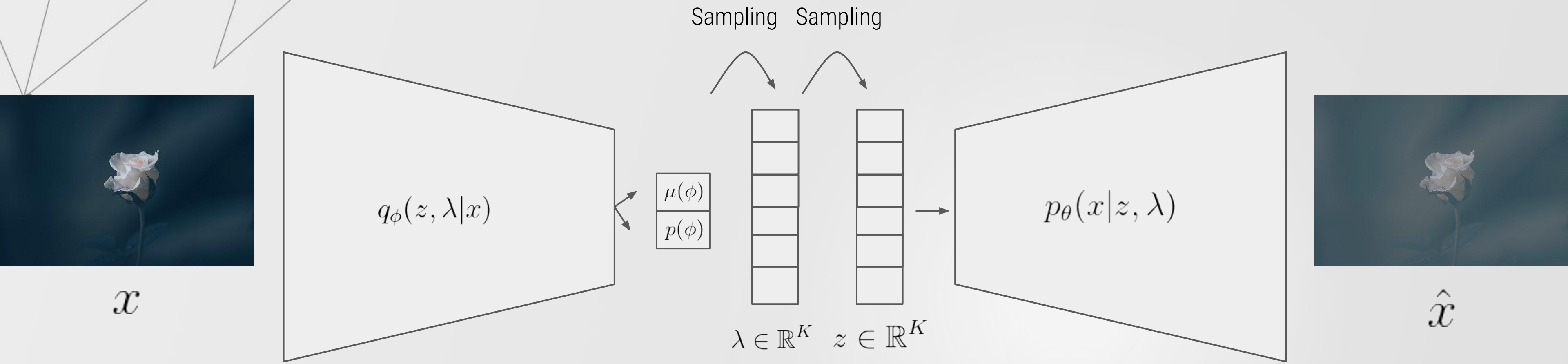
(NGVAE)



$$q_\phi(\lambda_k^{-1}|x) = \overbrace{p_k \mathcal{IG}(\lambda_k^{-1}; \alpha_1, \beta_1)}^{\text{uninformative}} + \overbrace{(1 - p_k)\mathcal{IG}(\lambda_k^{-1}; \alpha_2, \beta_2)}^{\text{informative}}$$

PROPOSED ARCHITECTURE

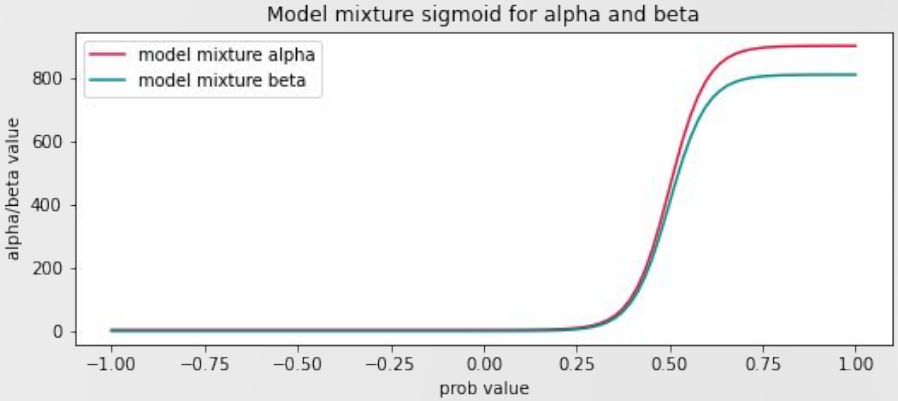
The Normal-Gamma Variational-Autoencoder
(NGVAE)



$$-ELBO = \mathcal{L}(\phi, \theta) = \mathbb{E}_{x \sim p(x)} \left[\mathbb{E}_{\sim q_\phi(z, \lambda|x)} \left[-\log p_\theta(x|z) \right] + \sum_{k=1}^{K} KL \left[q_\phi(z_k, \lambda_k|x) || p(z_k, \lambda_k) \right] \right]$$

RELAXING ASSUMPTIONS

To enable the gradient backpropagation



$$\alpha(p_k) = f_r(p_k) = \frac{\alpha_2 - \alpha_1}{1 + e^{-r(p_k - 0.5)}} + \alpha_1$$

$$\beta(p_k) = f_r(p_k) = \frac{\beta_2 - \beta_1}{1 + e^{-r(p_k - 0.5)}} + \beta_1$$

REPARAMETERIZATION TRICK

To enable the gradient backpropagation

$$\nabla_{\phi,\theta}\mathcal{L}(\phi,\theta) = -\nabla_{\phi,\theta}\mathbb{E}_{\sim q_\phi(z,\lambda|x)}[\log p_\theta(x|z)] + \nabla_\phi KL[q_\phi(z,\lambda|x)||p(z,\lambda)]$$

REPARAMETERIZATION TRICK

To enable the gradient backpropagation

$$\nabla_{\phi,\theta}\mathcal{L}(\phi,\theta) = -\nabla_{\phi,\theta}\mathbb{E}_{\sim q_\phi(z,\lambda|x)}[\log p_\theta(x|z)] + \underbrace{\nabla_\phi KL[q_\phi(z,\lambda|x)||p(z,\lambda)]}_{\text{Can be computed analytically}}$$

REPARAMETERIZATION TRICK

To enable the gradient backpropagation

$$\nabla_{\phi,\theta}\mathcal{L}(\phi,\theta) = -\underbrace{\nabla_{\phi,\theta}\mathbb{E}_{\sim q_\phi(z,\lambda|x)}[\log p_\theta(x|z)]}_{\text{Raises difficulties for the calculation of the gradient}} + \underbrace{\nabla_\phi KL[q_\phi(z,\lambda|x)||p(z,\lambda)]}_{\text{Can be computed analytically}}$$

REPARAMETERIZATION TRICK

To enable the gradient backpropagation

$$\nabla_{\phi,\theta}\mathcal{L}(\phi,\theta) = -\underbrace{\nabla_{\phi,\theta}\mathbb{E}_{\sim q_\phi(z,\lambda|x)}[\log p_\theta(x|z)]}_{} + \nabla_\phi KL[q_\phi(z,\lambda|x)||p(z,\lambda)]$$

- Changing of variables : weak correlations with variational parameters

REPARAMETERIZATION TRICK

To enable the gradient backpropagation

$$\nabla_{\phi,\theta}\mathcal{L}(\phi,\theta) = -\underbrace{\nabla_{\phi,\theta}\mathbb{E}_{\sim q_\phi(z,\lambda|x)}[\log p_\theta(x|z)]}_{} + \nabla_\phi KL[q_\phi(z,\lambda|x)||p(z,\lambda)]$$

- Changing of variables : weak correlations with variational parameters
- Applying the score trick

Ruiz & al. 2016

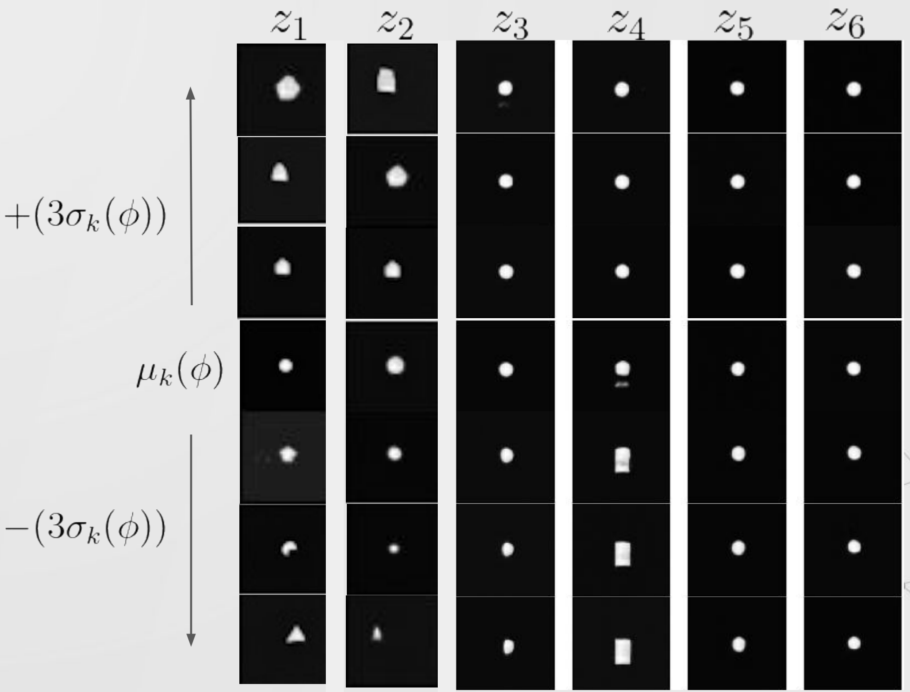
04

RESULTS

Separation in the latent space

RESULTS

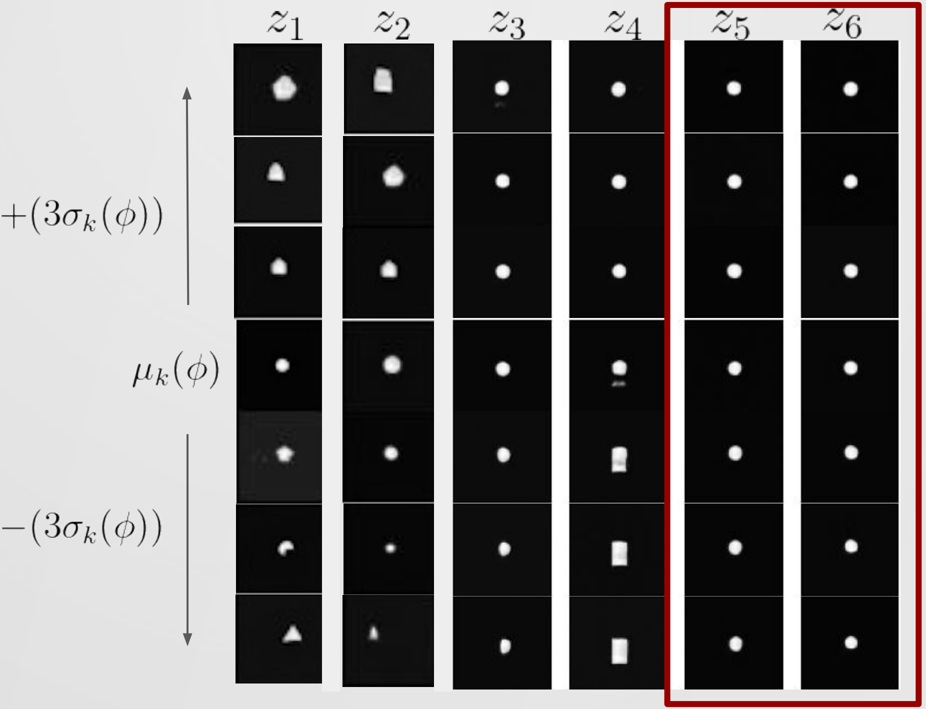
Separation in the latent space

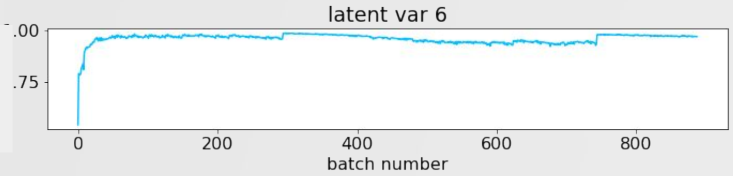


The latent space is clearly separated between informative and uninformative components, reconstruction is not impacted

RESULTS

Separation in the latent space

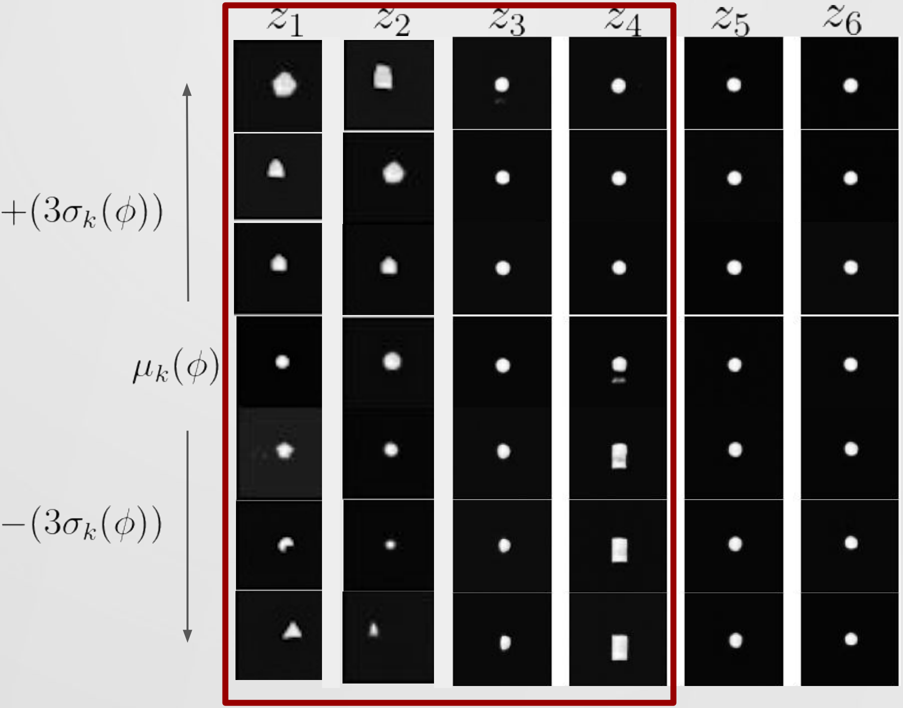


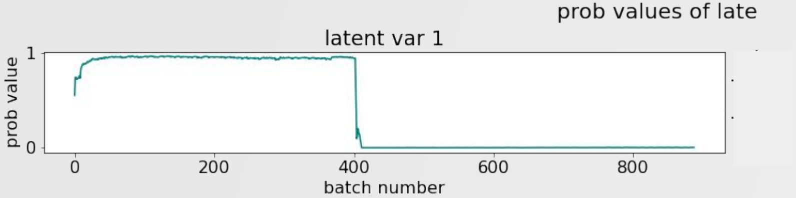


The corresponding p_k is high

RESULTS

Separation in the latent space

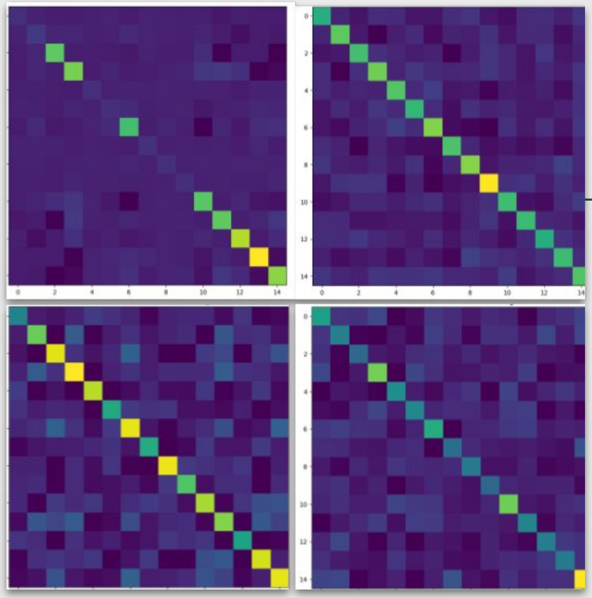




The corresponding p_k is low

RESULTS

Separation in the latent space



Scaled empirical covariance matrices of the latent variables over the whole dataset for, from top-left to bottom-right: the NGVAE, the beta-VAE with eta=150, the beta$VAE with beta=27 and the vanilla-VAE.

Clear separated latent space

Comparing to other state of the art architectures, the proposed model is **able to discriminate between informative** and uninformative components

RESULTS

Separation in the latent space

	Vanilla	Beta-VAE (27)	Beta-VAE (150)	NGVAE
Entropy (-)	2.6	2.6	2.6	**2.3**
Decorrelation (+)	12.7	7.8	11.9	**14.3**
Disentanglem ent (+)	0.67	0.64	0.62	**0.68**
Information	15	15	**7**	**7**

05

CONCLUSION

Improving disentanglement & measuring uncertainties

CONCLUSION

- Automatically defines the **number of informative components** of the latent space
- Number of informative components **matching** the **number of generative factors**
- **No needs** of adjusting dataset dependent **hyperparameters**
- However, the **reparametrization** involving jacobian **is computationally expensive**

FUTURE WORK

- Application on **state of the art dataset**
- Improving **disentanglement** within the latent space
- Extending to **flat hierarchical** representation for **uncertainty** measurement

THANKS

Does anyone have any **question**?
emma.jouffroy@u-bordeaux.fr

RELATED WORKS

- Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
- Bengio, Y., Courville, A., and Vincent, P. Representation learning: A review and new perspectives. IEEE transactions on Pattern Analysis and Machine Intelligence, 35 (8):1798–1828, 2013.
- Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., ... & Lerchner, A. (2016). beta-vae: Learning basic visual concepts with a constrained variational framework.
- Ruiz, F. R., AUEB, T. R., & Blei, D. (2016). The generalized reparameterization gradient. *Advances in neural information processing systems*, *29*.
- *R. T. Chen, X. Li, R. B. Grosse and D. K. Duvenaud, "Isolating sources of disentanglement in variational autoencoders" NeurIPS 2018, Vol. 31.*
- *H.Kim and A. Mnih, "Disentangling by factorising", ICML 2018, pp. 2649-2658*