

# q-VAE for Disentangled Representation Learning and Latent Dynamical Systems

Taisuke Kobayashi<sup>1</sup>

**Abstract**—A variational autoencoder (VAE) derived from Tsallis statistics called q-VAE is proposed. In the proposed method, a standard VAE is employed to statistically extract latent space hidden in sampled data, and this latent space helps make robots controllable in feasible computational time and cost. To improve the usefulness of the latent space, this paper focuses on disentangled representation learning, e.g.,  $\beta$ -VAE, which is the baseline for it. Starting from a Tsallis statistics perspective, a new lower bound for the proposed q-VAE is derived to maximize the likelihood of the sampled data, which can be considered an adaptive  $\beta$ -VAE with deformed Kullback-Leibler divergence. To verify the benefits of the proposed q-VAE, a benchmark task to extract the latent space from the MNIST dataset was performed. The results demonstrate that the proposed q-VAE improved disentangled representation while maintaining the reconstruction accuracy of the data. In addition, it relaxes the independency condition between data, which is demonstrated by learning the latent dynamics of nonlinear dynamical systems. By combining disentangled representation, the proposed q-VAE achieves stable and accurate long-term state prediction from the initial state and the action sequence.

## I. INTRODUCTION

Recently, deep learning [1] has become the most powerful tool to resolve analytically unsolvable problems. In robotics and robot control, several studies have applied deep learning to perform complicated tasks that rely on raw images from cameras [2], [3]. Typically, deep learning requires a large number of samples for end-to-end learning, i.e., from extracting features hidden in the inputs (e.g., images) to optimal control based on the extracted features.

Modularizing various functions, e.g., feature extraction [4], [5] and optimal control [6]–[9], is effective at reducing the number of samples. In particular, control theory has long been studied relative to stability, convergence speed, etc. Therefore, exploiting conventional but powerful control technology is a solution to reducing the number of samples; however, features extracted without ingenuity would be unsuitable.

Thus, this paper focuses on methods to extract features hidden in inputs. To this end, the variational autoencoder (VAE) [4] is a promising method. The VAE encodes inputs to a latent space with stochastic latent variables (and decodes them to inputs) in an unsupervised manner. In fact, control methods for the latent space gained by VAE have been proposed previously [10]–[12].

To improve the usefulness of such a latent space, recent VAE research has investigated disentangled representation

learning [13], [14], which assigns independent attributes hidden in the inputs to the axes of the latent space without supervisory signals. However, a representative of this methodology, i.e.,  $\beta$ -VAE [15], would lose the value of the VAE as a data generation model because the reconstruction accuracy of the inputs is reduced. Although variants of  $\beta$ -VAE have been proposed to resolve this problem, they have some drawbacks from a practicality perspective, e.g., heuristic designs are difficult to optimize [16], the assumption of batch data [17], and optimization of an additional discriminator network [18].

To realize simple and low-cost disentangled representation learning that is practically sufficient, this paper proposes a derivation of VAE combined with Tsallis statistics [19]–[22], which refer to the extended version of general statistics based on real parameter  $q$ . This derivation, which we refer to as q-VAE, mathematically provides adaptive  $\beta$  according to the amount of information in latent variables. Due to the adaptive  $\beta$ , the proposed q-VAE achieves proper extraction of the essential information of inputs while maintaining the reconstruction accuracy of the inputs. In addition, deformed Kullback-Leibler (KL) divergence [22], [23] weakens the magnitude of regularization in the vicinity of the center location of prior, which may construct the meaningful latent space.

In addition, the proposed q-VAE relaxes the condition that inputs be independent and identically distributed (i.i.d.) [21], which, generally, cannot be satisfied in a functional manner when controlling robots. Therefore, q-VAE is also extended to a model to learn the latent dynamics of inputs when manipulated variables are given. As a result, the encoded latent variables are transited to the next ones that are decoded to the next inputs. With a disentangled representation, even if the latent dynamics model is constrained as a diagonal system (i.e., a model in which latent variables are independent), which could be easily exploited to control robots with low computational cost, the proposed q-VAE alleviates the deterioration of prediction performance.

We verified the proposed q-VAE using an MNIST benchmark. The results indicate that q-VAE with an appropriate parameter outperforms  $\beta$ -VAE (i.e., fixed  $\beta$ ) in terms of disentangled representation and the reconstruction accuracy of inputs. In addition, learning the latent dynamics in nonlinear systems was performed. Although conventional methods fail to predict future states in a stable manner, q-VAE realizes stable and accurate long-term state prediction from the initial state and the action sequence.

<sup>1</sup>Taisuke Kobayashi is with the Division of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan [kobayashi@is.naist.jp](mailto:kobayashi@is.naist.jp)

## II. PRELIMINARIES

### A. Variational autoencoder

In this section, we briefly introduce the VAE [4] (see the upper structure of Fig. 2). Here, a generative model of inputs  $\mathbf{x}$  from latent variables  $\mathbf{z}$  is considered. This decoder is approximated by (deep) neural networks with parameters  $\theta$ :  $p(\mathbf{x} | \mathbf{z}; \theta)$ . The VAE attempts to maximize the log likelihood of  $N$  inputs,  $\log p(X)$ , where  $X = \{\mathbf{x}_n\}_1^N$ . By assuming that the inputs are i.i.d.,  $\log p(X)$  can be simplified to the sum of the respective log likelihoods  $\sum_{n=1}^N \log \mathbf{x}_n$ . To maximize  $\log p(X)$  indirectly, an evidence lower bound (ELBO)  $\mathcal{L}(X)$ , which is derived using Jensen's inequality and an encoder with parameters  $\phi$ ,  $\rho(\mathbf{z} | \mathbf{x}; \phi)$ , is maximized. Then,  $\mathcal{L}(X)$  is derived as follows.

$$\begin{aligned} \log p(X) &= \sum_{n=1}^N \log \int p(\mathbf{x}_n | \mathbf{z}; \theta) p(\mathbf{z}) d\mathbf{z} \\ &= \sum_{n=1}^N \log \int \frac{\rho(\mathbf{z} | \mathbf{x}_n; \phi)}{\rho(\mathbf{z} | \mathbf{x}_n; \phi)} p(\mathbf{x}_n | \mathbf{z}; \theta) p(\mathbf{z}) d\mathbf{z} \\ &\geq \sum_{n=1}^N \mathbb{E}_{\rho(\mathbf{z} | \mathbf{x}_n; \phi)} \left[ \log \frac{p(\mathbf{x}_n | \mathbf{z}; \theta) p(\mathbf{z})}{\rho(\mathbf{z} | \mathbf{x}_n; \phi)} \right] \\ &=: \mathcal{L}(X) \end{aligned} \quad (1)$$

where  $p(\mathbf{z})$  is a prior of the latent variables.  $\mathcal{L}(X)$  can be expressed as follows.

$$\begin{aligned} \mathcal{L}(X) &= \sum_{n=1}^N \mathbb{E}_{\rho(\mathbf{z} | \mathbf{x}_n; \phi)} [\log p(\mathbf{x}_n | \mathbf{z}; \theta)] \\ &\quad - \text{KL}(\rho(\mathbf{z} | \mathbf{x}_n; \phi) || p(\mathbf{z})) \\ &\simeq \sum_{n=1}^N \log p(\mathbf{x}_n | \mathbf{z}_n; \theta) - \text{KL}(\rho(\mathbf{z} | \mathbf{x}_n; \phi) || p(\mathbf{z})) \end{aligned} \quad (2)$$

where the first term denotes the negative reconstruction error in the standard autoencoder, and the second term, i.e., the KL divergence between the posterior and prior, is the regularization term, which is used to attempt to  $\mathbf{z} \sim p(\mathbf{z})$ . Generally,  $p(\mathbf{z})$  is given as standard normal distribution  $\mathcal{N}(\mathbf{0}, I)$ .

In the case of  $\beta$ -VAE [15], the second term is multiplied by  $\beta \in (0, \infty)$ , although its derivation is given by solving a constrained optimization problem. By giving  $\beta > 1$ , regularization to  $p(\mathbf{z})$  is strengthened, and the role of each axis of the latent space, which has limited expressive ability, is clarified to reconstruct the inputs.

### B. Tsallis statistics

Tsallis statistics refers to the organization of mathematical functions and associated probability distributions proposed by Tsallis [19]. This concept is organized based on  $q$ -deformed exponential and logarithmic functions, which are extensions of general exponential and logarithmic functions by real number  $q \in \mathbb{R}$ . We introduce the following definitions From Tsallis statistics.

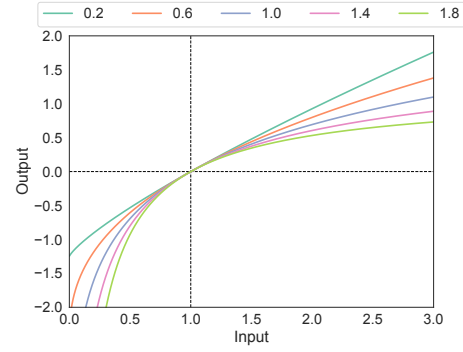


Fig. 1. Examples of  $q$ -logarithm with different  $q$  values. When  $q > 0$ , this function is considered as a concave function. As can be seen in the intuitive,  $\ln_{q+\epsilon}(x)$  with  $\epsilon$  a positive small number is larger than  $\ln_q(x)$ .

First, the  $q$ -logarithm,  $\ln_q(x)$  with  $x > 0$ , is given as follows.

$$\ln_q(x) = \begin{cases} \ln(x) & q = 1 \\ \frac{x^{1-q} - 1}{1-q} & q \neq 1 \end{cases} \quad (3)$$

where  $q$  gives its shape, as shown in Fig. 1. As shown in Fig. 1, the  $q$ -logarithm with  $q > 0$  is a concave function. Note that when  $q \rightarrow 1$ , the lower equation converges to the natural logarithm.

In the  $q$ -logarithm, multiplication of two variables,  $x, y > 0$ , is no longer simply divided. In other words, the following pseudo-additivity is derived.

$$\ln_q(xy) = \ln_q(x) + \ln_q(y) + (1-q) \ln_q(x) \ln_q(y) \quad (4)$$

Rather than general multiplication, a new multiplication operation  $\otimes_q$  is introduced as follows [20].

$$x \otimes_q y = \begin{cases} (x^{1-q} + y^{1-q} - 1)^{\frac{1}{1-q}} & x^{1-q} + y^{1-q} > 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

This definition means that the following additivity is satisfied when using  $\otimes_q$ .

$$\ln_q(x \otimes_q y) = \ln_q(x) + \ln_q(y) \quad (6)$$

By using  $\otimes_q$ , the  $q$ -likelihood of data  $X = \{\mathbf{x}_n\}_1^N$ , which is maximized when the probability is given as  $q$ -Gaussian, can be defined with  $q$ -i.i.d. (a relaxed version of the i.i.d. condition [20], [21]).

$$p(X) = p(\mathbf{x}_1) \otimes_q p(\mathbf{x}_2) \otimes_q \cdots \otimes_q p(\mathbf{x}_N) \quad (7)$$

Finally, the deformed version of KL divergence (referred to as Tsallis divergence [22]),  $\text{KL}_q$ , is expressed as follows.

$$\text{KL}_q(p_1 || p_2) = - \int p_1(x) \ln_q \frac{p_2(x)}{p_1(x)} dx \quad (8)$$

where  $p_1$  and  $p_2$  are arbitrary probability density functions. The  $q$ -logarithm for all  $x$  decreases as  $q$  increases; thus,  $\text{KL}_q$  increases as  $q$  increases. Tsallis divergence can be derived by transforming Renyi divergence, which has closed-form solutions for commonly-used distributions [23]. Therefore, the proposed  $q$ -VAE can be integrated with different priors, e.g., the Laplace distribution and mixture models.

### III. Q-VARIATIONAL AUTOENCODER

#### A. Derivation of ELBO

By combining Eqs. (6) and (7), a new  $q$ -log likelihood to be maximized is derived. In addition, if  $q > 0$ , the  $q$ -logarithm is a concave function (Fig. 1); therefore, Jensen's inequality can be used in a manner similar to Eq. (1).

$$\begin{aligned} \ln_q p(X) &= \sum_{n=1}^N \ln_q \int p(\mathbf{x}_n | \mathbf{z}; \boldsymbol{\theta}) p(\mathbf{z}) d\mathbf{z} \\ &\geq \sum_{n=1}^N \mathbb{E}_{\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi})} \left[ \ln_q \frac{p(\mathbf{x}_n | \mathbf{z}; \boldsymbol{\theta}) p(\mathbf{z})}{\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi})} \right] \\ &=: \mathcal{L}_q(X) \end{aligned} \quad (9)$$

Note again that  $X$  can be relaxed to be  $q$ -i.i.d. data, unlike the log likelihood used in Eq. (1).

According to Eq. (4),  $\mathcal{L}_q(X)$  can be divided into three terms and summarized to two terms similar to Eq. (2).

$$\begin{aligned} \mathcal{L}_q(X) &= \sum_{n=1}^N \mathbb{E}_{\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi})} \left[ \ln_q p(\mathbf{x}_n | \mathbf{z}; \boldsymbol{\theta}) \right. \\ &\quad \left. + \ln_q \frac{p(\mathbf{z})}{\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi})} \right. \\ &\quad \left. + (1-q) \ln_q p(\mathbf{x}_n | \mathbf{z}; \boldsymbol{\theta}) \ln_q \frac{p(\mathbf{z})}{\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi})} \right] \\ &\simeq \sum_{n=1}^N \ln_q p(\mathbf{x}_n | \mathbf{z}_n; \boldsymbol{\theta}) \left\{ 1 + (1-q) \ln_q \frac{p(\mathbf{z}_n)}{\rho(\mathbf{z}_n | \mathbf{x}_n; \boldsymbol{\phi})} \right\} \\ &\quad - \text{KL}_q(\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi}) || p(\mathbf{z})) \\ &= \sum_{n=1}^N \frac{\ln_q p(\mathbf{x}_n | \mathbf{z}_n; \boldsymbol{\theta})}{\beta_q(\mathbf{x}_n, \mathbf{z}_n)} - \text{KL}_q(\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi}) || p(\mathbf{z})) \end{aligned} \quad (10)$$

where the first term is denotes the negative reconstruction error with a new adaptive variable  $1/\beta_q(\mathbf{x}_n, \mathbf{z}_n)$ , and the second term, i.e., Tsallis divergence between the posterior and prior, is the regularization term used to attempt to  $\mathbf{z} \sim p(\mathbf{z})$ .

#### B. Practical implementation

Here, we present four practical statements. First, the probability output from decoder  $p(\mathbf{x} | \mathbf{z}; \boldsymbol{\theta})$  should be assumed as  $q$ -Gaussian distribution (or a Bernoulli distribution for binary inputs) to match the reconstruction error term with that of the standard VAE. Although a Gaussian distribution can be assumed, it would be numerically unstable because its exponential function is not canceled by logarithm and the power function is added by the  $q$ -logarithm. In addition,  $q$ -Gaussian distribution includes the student- $t$  distribution, which yields robust estimation [24]; therefore, further investigation of the decoder model may further unlock the potential of  $q$ -VAE.

Second, the computational graph of  $\beta_q(\mathbf{x}, \mathbf{z})$  is cut to simplify backpropagation and regard it as simply the input-dependent coefficient. Even with the computational graph, the learning direction of parameters (see the next section)

would be the same as the case without it. However, in the case with the computational graph, the gradient scale varies, thereby making learning unstable if the same hyperparameters as the standard VAE are employed. In fact, that case happened learning failure in debugging.

Third, we employ the latent distribution as Gaussian, which has a closed-form solution of Tsallis divergence [22], [23]. Given  $p_1$  and  $p_2$  as  $d$ -dimensional normal distributions with parameters  $\boldsymbol{\mu}_1$ ,  $\Sigma_1$ ,  $\boldsymbol{\mu}_2$ , and  $\Sigma_2$ , Tsallis divergence is solved as follows.

$$\text{KL}_q(p_1 || p_2) = \begin{cases} \frac{1}{2} \left\{ \text{tr}(\Sigma_2^{-1} \Sigma_1) + \ln \frac{|\Sigma_2|}{|\Sigma_1|} - d \right. \\ \quad \left. + (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^\top \Sigma_2^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) \right\} & q = 1 \\ \frac{\exp(\frac{1}{2} I_q(p_1 || p_2)) - 1}{q - 1} & q \neq 1 \end{cases} \quad (11)$$

where,

$$\begin{aligned} I_q(p_1 || p_2) &= \ln \frac{|\Sigma_2|^q |\Sigma_1|^{1-q}}{|\Sigma|} \\ &\quad + q(1-q) (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^\top \Sigma^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) \end{aligned} \quad (12)$$

$$\Sigma = q \Sigma_2 + (1-q) \Sigma_1 \quad (13)$$

Finally, according to the derivation of  $\beta$ -VAE [15],  $q$ -VAE can be integrated with  $\beta$ -VAE. In other words, the following constrained optimization problem is solved.

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \sum_{n=1}^N \frac{\ln_q p(\mathbf{x}_n | \mathbf{z}_n; \boldsymbol{\theta})}{\beta_q(\mathbf{x}_n, \mathbf{z}_n)} \\ \text{s.t. } \text{KL}_q(\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi}) || p(\mathbf{z})) < \epsilon \end{aligned}$$

where  $\epsilon$  denotes a threshold of Tsallis divergence. This problem can be rewritten as a Lagrangian under the KKT conditions as follows.

$$\begin{aligned} \mathcal{L}_{(\beta, q)}(X) &= \sum_{n=1}^N \frac{\ln_q p(\mathbf{x}_n | \mathbf{z}_n; \boldsymbol{\theta})}{\beta_q(\mathbf{x}_n, \mathbf{z}_n)} \\ &\quad - \beta \text{KL}_q(\rho(\mathbf{z} | \mathbf{x}_n; \boldsymbol{\phi}) || p(\mathbf{z})) \end{aligned} \quad (14)$$

where  $\beta$  denotes the hyperparameter to tune the tradeoff between reconstruction and regularization.

#### C. Analysis

In the proposed  $q$ -VAE, the above ELBO is maximized by optimizing parameter  $\boldsymbol{\theta}$  of the decoder and parameter  $\boldsymbol{\phi}$  of the encoder. As shown in Eqs. (2) and (10),  $q$ -VAE extends the standard VAE by adding parameter  $q$  because, if  $q \rightarrow 1$ ,  $\ln_q$ ,  $\beta_q$ , and  $\text{KL}_q$  converge on  $\ln$ , 1, and  $\text{KL}$ , respectively.

In addition, the proposed  $q$ -VAE can be considered a type of  $\beta$ -VAE with adaptive  $\beta_q(\mathbf{x}_n, \mathbf{z}_n)$ . According to range of  $q$ , the following three cases are expected.

- 1)  $q = 1$ : As mentioned previously,  $\beta_q$  is always equal to 1 in this case. Therefore, this case is considered the standard VAE or  $\beta$ -VAE with hyperparameter  $\beta$ .
- 2)  $q < 1$ : Generally, the posterior  $\rho(\mathbf{z}_n | \mathbf{x}_n; \boldsymbol{\phi})$  contains more information than the prior  $p(\mathbf{z}_n)$ . Thus,  $\beta_q$  would likely be greater than 1. In other words, the posterior is strongly constrained to discard information about

inputs only when it has a significant amount of information. Consequently, only essential information of the inputs is expected to be left without duplication in the latent space (and its axes). However, smaller  $q$  leads to smaller  $\text{KL}_q$ ; thus, if  $q$  is too small, no matter how large  $\beta_q$  is, the constraint to the prior will not work as described above because it was originally small.

- 3)  $q > 1$ : In contrast to the  $q < 1$  case,  $\beta_q$  would likely be less than 1. In other words, reconstruction would be prioritized without the constraint to the prior. In addition, as shown in Eq. (13), the covariance matrix for the Tsallis divergence calculation is likely to violate its positive-semidefinite condition.

From the above, we expect that a  $q$  value that is less than 1 is better relative to achieving disentangled representation learning with stable computation. To the best of our knowledge, no methods have been proposed to date to automatically tune  $\beta$  in  $\beta$ -VAE. Therefore, the proposed  $q$ -VAE method is the first method that can tune  $\beta$  via a mathematically natural derivation.

To avoid adverse effects caused by excessively small  $q$  values, the simplified version of Eq. (14) is investigated in the following.

$$\mathcal{L}_{(\beta,q)}^{\text{sm}}(X) = \sum_{n=1}^N \frac{\ln_q p(\mathbf{x}_n | \mathbf{z}_n; \theta)}{\beta_q(\mathbf{x}_n, \mathbf{z}_n)} - \beta \text{KL}(\rho(\mathbf{z} | \mathbf{x}_n; \phi) || p(\mathbf{z})) \quad (15)$$

This means that KL (rather than  $\text{KL}_q$ ) is employed for the strong constraint. When  $q < 1$ ,  $\mathcal{L}_{(\beta,q)}^{\text{sm}}$  is less than the original  $\mathcal{L}_{(\beta,q)}$  because KL is stronger than  $\text{KL}_q$ . Even though this case would avoid an overly weak constraint by  $\text{KL}_q$  with  $q \ll 1$ , it may make the constraint stronger when  $q \simeq 1$ .

#### D. Extension to latent dynamical systems

Another advantage of Tsallis statistics is given by the  $q$ -likelihood defined in Eq. (7). Its  $q$ -logarithm can be converted to the sum of the  $q$ -log likelihoods of respective samples without independency (i.e., with  $q$ -independency) between each sample. In other words, even if the inputs are sampled from dynamical systems with a transition probability, the proposed  $q$ -VAE can be applied as is.

A simple dynamical system is designed in reference to the literature [10], [25] (Fig. 2). Here, a trajectory of inputs  $X = \{\mathbf{x}_t\}_1^T$ , where  $T$  is the maximum time step, is generated from the following dynamics with parameters  $\eta$ .

$$\mathbf{z}_t \sim \rho(\mathbf{z}_t | \mathbf{x}_t; \phi) \quad (16)$$

$$\mathbf{z}_{t+1} \sim p(\mathbf{z}_{t+1} | \mathbf{z}_t, \mathbf{u}_t; \eta) \quad (17)$$

$$\mathbf{x}_{t+1} \sim p(\mathbf{x}_{t+1} | \mathbf{z}_{t+1}; \theta) \quad (18)$$

where  $\mathbf{u}$  denotes the manipulated variables of the dynamical system.

In that time,  $\mathcal{L}_{(\beta,q)}(X)$  in Eq. (14) is redefined as follows.

$$\mathcal{L}_{(\beta,q)}^{\text{dyn}}(X) = \sum_{t=1}^T \frac{\ln_q p(\mathbf{x}_{t+1} | \mathbf{z}_{t+1}^\eta; \theta)}{\beta_q(\mathbf{x}_t, \mathbf{z}_t)} - \beta \text{KL}_q(\rho(\mathbf{z} | \mathbf{x}_t; \phi) || p(\mathbf{z})) \quad (19)$$

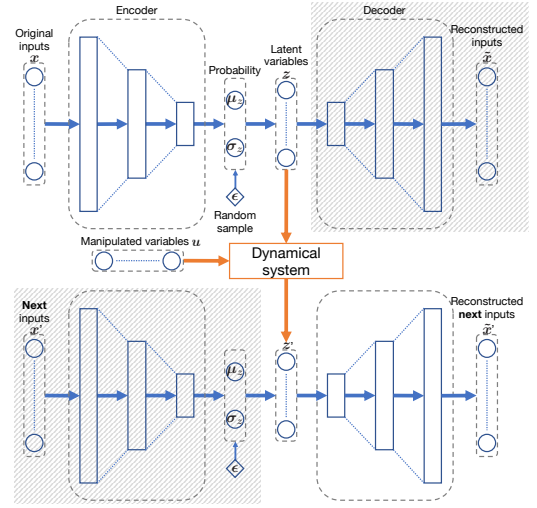


Fig. 2. Dynamics model in latent space: upper and lower components are the same network structure as the standard VAE; the next inputs  $\mathbf{x}'$  is predicted according to the current inputs  $\mathbf{x}$  and the given manipulated variables  $\mathbf{u}$ ; in this paper, a dynamical system is given to be time-varying linear and diagonal.

where  $\mathbf{z}_{t+1}^\eta$  denotes  $\mathbf{z}_{t+1}$  as predicted by the latent dynamics. When maximizing this function,  $\eta$  must also be optimized implicitly. However, the information about the dynamics would penetrate the decoder (and encoder), and  $\eta$  would fail to represent the latent dynamics. Therefore, in addition to  $\text{KL}_q$ , a further constraint is applied to the constrained optimization problem (Section III-B).

$$-\sum_{t=1}^T \ln \rho(\mathbf{z}_{t+1}^\eta | \mathbf{x}_{t+1}; \phi) < \epsilon$$

Similarly, this constraint can be rewritten as an additional maximization target with weight  $\gamma$  as follows.

$$\mathcal{L}^{\text{latent}}(X) = \gamma \sum_{t=1}^T \ln \rho(\mathbf{z}_{t+1}^\eta | \mathbf{x}_{t+1}; \phi) \quad (20)$$

Such a constraint is also introduced in reference to the literature [10]; however, this can be applied even if only the sampled latent variables are transited to the next ones using Eq. (17).

Note that the general form of the latent dynamics is simplified as much as possible to reduce computational cost. The general (i.e., nonlinear) dynamics are regarded to be time-varying linear, as is used in general nonlinear control via first-order Taylor expansion. As mentioned previously, the proposed  $q$ -VAE is suitable for disentangled representation; therefore, latent variables are ultimately independent of each other. The latent dynamics are simplified as follows.

$$\mathbf{z}_{t+1} = \text{diag}(\mathbf{a}(\mathbf{z}_t; \eta))\mathbf{z}_t + B(\mathbf{z}_t; \eta)\mathbf{u}_t \quad (21)$$

where  $\mathbf{a}$  and  $B$  are the outputs of the network with inputs  $\mathbf{z}$  and parameters  $\eta$ . In fact, this latent dynamics design is expected to cause modeling an error unless the latent variables are encoded to match this dynamics using disentangled representation learning.

#### IV. MNIST BENCHMARK

Here, we verify the performance of disentangled representation using the proposed q-VAE. The MNIST dataset, which contains  $28 \times 28 = 784$ -dimensional images with handwritten numbers  $0 \sim 9$ , was used in this evaluation.

##### A. Criteria of disentangled representation

To evaluate how the latent space obtains a disentangled representation, we consider the fact that a disentangled representation attempts to gain independent axes with essential information. This concept is closely related to independent component analysis (ICA) [26]. Therefore, kurtosis in the latent space is employed according to ICA as a criterion of disentangled representation.

Specifically, kurtosis from the sampled data  $Z = \{z_n\}_1^N$  where  $z_n \sim \rho(z \mid x_n; \phi)$ ,  $\kappa$ , is defined as Mardia's kurtosis [27].

$$\kappa = \frac{1}{N} \sum_{n=1}^N \{ (z_n - \mu_z)^\top \Sigma_z^{-1} (z_n - \mu_z) \} - d(d+2) \quad (22)$$

where  $\mu_z$  and  $\Sigma_z$  denote the mean and covariance of  $Z$ , respectively. A greater  $\kappa$  values indicates better disentangled representation.

In addition, reconstruction error criterion is important to demonstrate how much the informative latent space is achieved. Therefore, binary cross entropy (BCE) is employed. Here, a smaller BCE value indicates better reconstruction.

Note that these two criteria are computed only from the test data.

##### B. Network structure

With the proposed q-VAE, the difference from the baselines is only related to ELBO (i.e., the loss function); therefore, the same network structure was employed for all compared methods. In addition, the networks were constructed using PyTorch [28].

The images were converted to 784-dimensional vectors. The encoder included five fully-connected layers: 500 neurons, 275 neurons, 50 neurons, and 20 neurons corresponding to  $\mu$  and  $\sigma$  of the posterior (i.e.,  $d = 10$ ). The decoder also had five fully-connected layers: 255 neurons, 500 neurons, and 784 neurons corresponding to the reconstructed inputs.

In addition, the outputs from the intermediate layers of both the encoder and decoder were processed via layer normalization [29] and Swish functions [30], [31]. We found that appropriate normalization techniques, e.g., layer normalization, are important relative to reduce the variance of kurtosis, and the activation function in the intermediate layers only improved the BCE value.

##### C. Results

The effects of  $\beta$  in  $\beta$ -VAE [15] and  $q$  value were investigated. Here, (a)  $\beta = [1, 10]$  with an increment of 1 and (b)  $q = [0.1, 1]$  an increment of 0.1 were tested over 50 trials with 100 epochs. The results are plotted in Figs. 3(a) and

(b), respectively. Note that random seeds in the respective trials are given as the number of trials.

As expected,  $\beta$ -VAE increased the kurtosis (i.e., the disentangled representation ability) as  $\beta$  increased (Fig. 3(a)). However, the BCE valued deteriorated linearly along with  $\beta$ . Here,  $\beta$ -VAE suffered a tradeoff problem between disentangled representation and reconstruction abilities, which may make the design of  $\beta$  difficult.

In contrast, as shown in Fig. 3(b), the kurtosis has a peak at  $q = 0.8$ , and BCE was decreased monotonically as the value of  $q$  increased. In addition, BCEs in the  $q > 0.7$  cases appeared to be sufficiently small, which suggests the value of  $q$  should be approximately 0.8 to achieve disentangled representation and effective reconstruction abilities. In other words, the above different behaviors from  $\beta$ -VAE are achieved by the proposed q-VAE with a simple implementation, a slight increase in computational cost, and minimal effort to optimize the hyperparameters. Here, one concern is that a small  $q$  value makes Tsallis divergence  $KL_q$  small and reduces reconstruction accuracy by increasing  $\beta_q$ .

The results obtained with Eq. (15), which was proposed to avoid such behavior, are shown in Fig. 3(c). Due to the stronger and stable constraint to the prior, kurtosis increased extremely as  $q$  decreased. However, as expected, the stronger constraint caused the BCE to increase, even in  $q \simeq 1$ . In other words, the same tradeoff as in  $\beta$ -VAE occurred; thus, the original derivation in Eq. (10) is more desirable compared to the simplified version in Eq. (15).

Next, q-VAE was combined with  $\beta$ -VAE, as introduced in Eq. (14). The learning curves of several combinations are shown in Fig. 4, where tuples denote  $(\beta, q)$ . By focusing on pairs of i) the (1.0, 0.8)- and (3.0, 1.0)-VAEs, and ii) the (3.0, 0.8)- and (5.0, 1.0)-VAEs, we found that the proposed q-VAE increased the kurtosis as much as  $\beta$ -VAE while making BCE smaller than that of  $\beta$ -VAE. In particular, the (1.0, 0.8)-VAE achieved superior disentangled representation ability than the (3.0, 1.0)-VAE, while also achieving similar reconstruction performance as the (1.0, 1.0)-VAE.

#### V. SIMULATION TO LEARN LATENT DYNAMICS

##### A. Dataset

As a proof of concept, a dynamical simulation provided by Open AI Gym [32], i.e., LunarLanderContinuous-v2 (Fig. 5), was learned by the proposed q-VAE using Eqs. (19) to (21). This simulation observes eight-dimensional state space data and is controlled via a three-dimensional action space. Here, an expert controls the lander to land it safely in the permitted area via a keyboard interface. In addition, the state and action pairs were collected as a non-i.i.d. expert's trajectory data. In total, the dataset comprises 150 trajectories for training (and validation) and 50 trajectories for testing. All tuples  $(x_t, u_t, x_{t+1})$  in the 150 trajectories were divided into 80 % training data and 20 % validation data. Note that, during validation (after each training epoch), only the state prediction error was evaluated to investigate underfitting and overfitting. After training, the compared methods attempted to predict the state (and latent variables) trajectories.

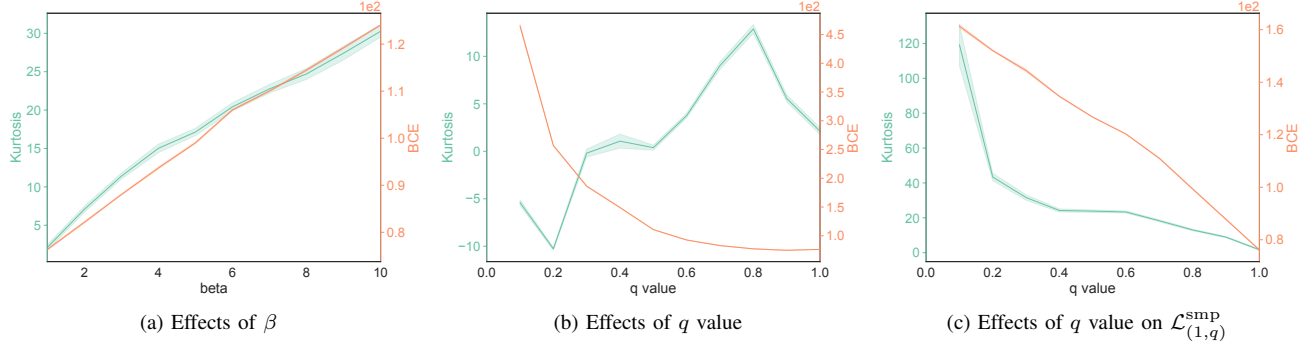


Fig. 3. Criteria relative to hyperparameters. The shapes of the respective plots imply the qualitative effects of hyperparameters. (a) Kurtosis and BCE increase approximately linearly as  $\beta$  increases. In other words, a tradeoff between the disentangled representation and reconstruction abilities is inevitable in  $\beta$ -VAE; (b) BCE is inversely proportional to  $q$  value, while kurtosis became a unimodal-like shape. (c) The simplified q-VAE defined in Eq. (15) shows that kurtosis and BCE are inversely proportional to the value of  $q$ .

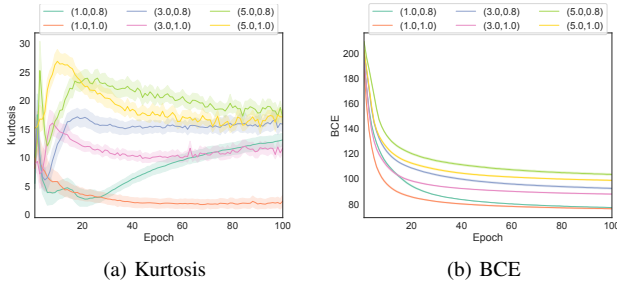


Fig. 4. Learning curves of the  $(\beta, q)$ -VAEs. The respective performances converge at 100 epoch.  $\beta$ -VAE effectively acquires large kurtosis at the expense of BCE. q-VAE increased kurtosis and suppressed the deterioration of BCE.

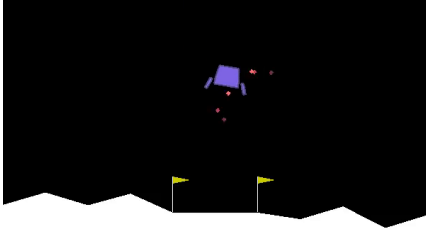


Fig. 5. LunarLanderContinuous-v2 [32]. A lander attempts to land safely inside two flags by controlling a main engine and two side engines. This problem has an eight-dimensional state and three-dimensional action spaces.

In this experiment, the prediction errors were the primary concern to demonstrate that the proposed q-VAE can extract latent dynamics, which is useful to reduce the computational cost of nonlinear model predictive control (MPC) [8], [9]. Therefore, the mean squared error (MSE) between the predicted and true states (or the predicted and encoded latent variables) is evaluated. In addition, when applying the learned model to MPC, long-term prediction is more important; therefore, all states (and latent variables) were predicted from the initial state by repeatedly going through the latent dynamics. Totally, four types of MSEs are given, i.e., 1-step state, 1-step latent,  $T$ -step state, and  $T$ -step latent, and these were computed using only the test data with 50

TABLE I  
NETWORK DESIGNS FOR LEARNING LATENT DYNAMICS

Version	Encoder	Latent dynamics
V1	[500, 400, 300, 200, 100]	[100, 100, 100, 100, 100]
V2	[250, 200, 150, 100]	[50, 50, 50]
V3	[100, 100, 100]	[100, 100, 100]

trajectories. Note that each trajectory has approximately 500 steps; thus, the proposed q-VAE had to predict states from one to approximately 500 steps future only using the initial state and actions at the respective times.

### B. Network structure

Due to real number inputs, all network layers were fully connected. In this study, three network structures were prepared to demonstrate the network-invariant performance. The numbers of layer neurons are listed in Table I. As can be seen, V1 is the largest structure, V2 is a moderate encoder with the smallest latent dynamics, and V3 is the simplest design. Note that a decoder inverted the hidden layers in the encoder with different parameters. In addition to the MNIST benchmark, layer normalization [29] and Swish functions [30], [31] were employed in this evaluation.

### C. Results

Three conditions with different  $(\beta, q)$  were compared:  $(0.0, 1.0)$  to demonstrate that the standard autoencoder does not have continuity in the latent space,  $(0.01, 1.0)$  as a baseline, and  $(0.01, 0.8)$  as the proposed q-VAE. Here,  $\beta$  was less than the MNIST case because the ratio between the input and latent dimensions was approximately 100 times different, and the  $\beta = 1$  case demonstrated relatively strong regularization. As common conditions, a three dimensional latent space was given, and  $\gamma$  in Eq. (20) was set to 0.1. For each condition, 50 trials with different random seeds were conducted, and the mean of the prediction error during each trajectory was computed for the test data.

The results are summarized in Fig. 6. Regardless of the conditions and network structures, ability to predict the 1-step state and latent variables was demonstrated. In



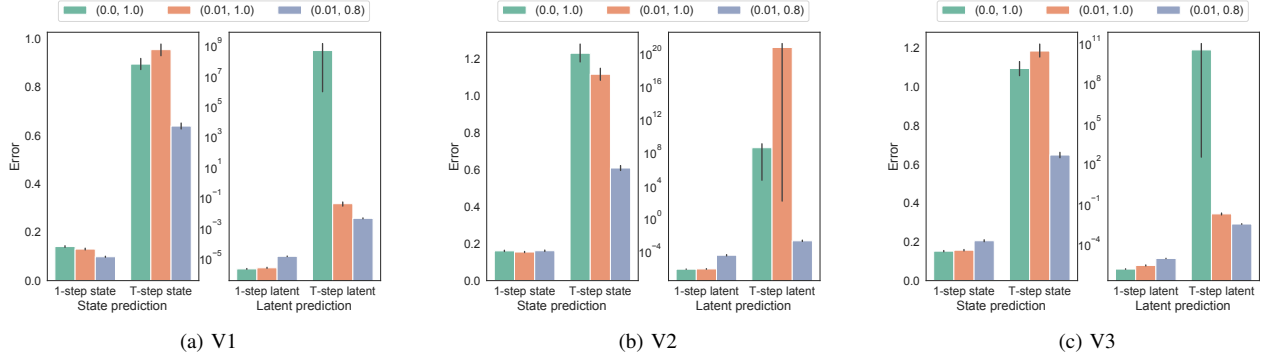


Fig. 6. Prediction errors of respective models. All methods succeeded in 1-step prediction with the same accuracy. With all models, the proposed q-VAE gained better  $T$ -step prediction than the compared methods. In particular, the compared methods made the  $T$ -step latent diverge occasionally; however, this trend did not occur with the proposed q-VAE.

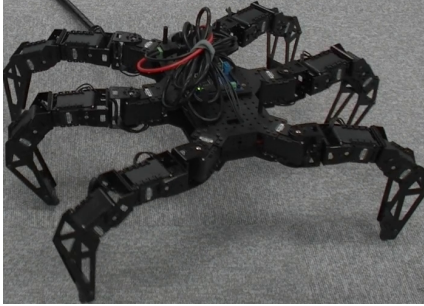


Fig. 7. PhantomX hexapod developed by Trossen Robotics. Each leg is controlled by inverse kinematics toward its reference given as an oscillator's phase.

contrast,  $T$ -step prediction demonstrated the superiority of q-VAE. For  $T$ -step latent variable prediction, the proposed q-VAE method yielded stable results, although the compared methods made it diverge occasionally. This implies that the latent space extracted by the proposed q-VAE provides a natural representation of dynamics. As a result,  $T$ -step state prediction was improved by q-VAE. Interestingly, although the other methods with the largest network structure (i.e., V1) gained better results compared to the other structures, the proposed q-VAE achieved the same excellent prediction performance regardless of network structure.

## VI. REAL ROBOT EXPERIMENT

### A. Conditions

In this experiment, the walking motion of a hexapod robot (Trossen Robotics; Fig. 7) was evaluated to learn its dynamics. The hexapod robot uses central pattern generators (CPGs) [33], [34] to generate periodic walking. The dynamics of six oscillators corresponding to the robot's legs are given as follows.

$$\dot{\xi} = e^{-ke}\omega + \alpha(\xi) \quad (23)$$

where  $\xi$  denotes the phases of the oscillators with natural frequency  $\omega$ , which is slowed according to the position errors of the legs  $e$  and gain  $k$ . Note that this error term imitates the tegotae-based [34].  $\alpha$  are computed according to the oscillator network, which attracts walking gait to tripod.

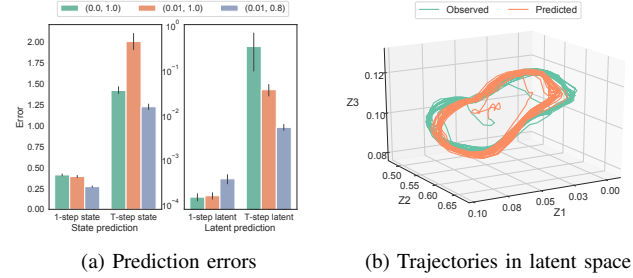


Fig. 8. Learning results of  $(\beta, q)$ -VAEs. (a) The proposed q-VAE improved prediction performance. (b) The trajectory predicted by the proposed q-VAE was attracted to almost the same periodic attractor as the observed attractor.

Here, the observed and reference joint angles are given as inputs. In total, this experiment considered 36-dimensional state and six-dimensional action spaces. Similar to the previous experiment, the dataset comprised 150 trajectories for training (and validation) and 50 trajectories for testing, and each trajectory involved 500 steps. In addition, the criteria and hyperparameters used in the previous experiment were considered in this experiment. Among the three network structures detailed in Table I, V3 was employed in this experiment. In addition, due to noisy observations, a robust optimizer [35] was also employed.

### B. Results

The results are shown in Fig. 8. As can be seen, the proposed q-VAE improved prediction performance compared to the baselines. As an example, prediction from the initial state and actions are visualized in the attached video. The latent variables are shown in the right of Fig. 8. As can be seen, the trajectory of the predicted latent variables were attracted to nearly the same periodic attractor as the observed ones. In other words, the proposed q-VAE revealed the natural latent dynamics corresponding to the oscillator dynamics from real observation data.

## VII. CONCLUSION

In this paper, we have proposed the q-VAE method, which is derived according to Tsallis statistics. Due to Tsallis statistics, the proposed q-VAE has three primary beneficial features, i.e., input-dependent  $\beta_q$  according to the amount of

information in the latent space, it demonstrates Tsallis divergence regularization rather than the standard KL divergence, and it relaxes the assumption of i.i.d. input data. The first two features are suitable for disentangled representation learning, and the proposed q-VAE outperformed the baseline  $\beta$ -VAE approach on the MNIST benchmark. In addition, the second feature, i.e., Tsallis divergence, was verified by testing a simplified version of q-VAE. The third feature allows the proposed q-VAE to be used to learn latent dynamics from non i.i.d. data. As a proof of concept, the proposed q-VAE was demonstrated to stably and accurately predicting future states (approximately 500 steps) from only the initial state and the action sequence.

Although the proposed q-VAE outperformed the compared baselines, it has several practical approximations. In future, we expect that further improvement can be obtained by removing these approximations. In addition, the proposed q-VAE was derived as a new base of VAE variants; thus, it can be integrated with the latest disentangled representation learning methods [16]–[18] to improve their performance. Finally, the proposed q-VAE can be applied to real complex systems with high-dimensional input, e.g., vision systems, to control them relative to the prediction of future states in real time using the latest control theory.

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI, Grant-in-Aid for Scientific Research (B), Grant Number 20H04265.

## REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [2] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, *et al.*, "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Conference on Robot Learning*, 2018, pp. 651–673.
- [3] Y. Tsurumine, Y. Cui, E. Uchibe, and T. Matsubara, "Deep reinforcement learning with smooth policy update: Application to robotic cloth manipulation," *Robotics and Autonomous Systems*, vol. 112, pp. 72–83, 2019.
- [4] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," in *International Conference on Learning Representations*, 2014.
- [5] T. Kobayashi, T. Aoyama, K. Sekiyama, and T. Fukuda, "Selection algorithm for locomotion based on the evaluation of falling risk," *IEEE Transactions on Robotics*, vol. 31, no. 3, pp. 750–765, 2015.
- [6] M. Neunert, C. De Crousaz, F. Furrer, M. Kamel, F. Farshidian, R. Siegwart, and J. Buchli, "Fast nonlinear model predictive control for unified trajectory optimization and tracking," in *IEEE international conference on robotics and automation*. IEEE, 2016, pp. 1398–1404.
- [7] T. Kobayashi, K. Sekiyama, Y. Hasegawa, T. Aoyama, and T. Fukuda, "Unified bipedal gait for autonomous transition between walking and running in pursuit of energy minimization," *Robotics and Autonomous Systems*, vol. 103, pp. 27–41, 2018.
- [8] J. Koenemann, A. Del Prete, Y. Tassa, E. Todorov, O. Stasse, M. Bennewitz, and N. Mansard, "Whole-body model-predictive control applied to the hrp-2 humanoid," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2015, pp. 3346–3351.
- [9] B. Amos, I. Jimenez, J. Sacks, B. Boots, and J. Z. Kolter, "Differentiable mpc for end-to-end planning and control," in *Advances in Neural Information Processing Systems*, 2018, pp. 8289–8300.
- [10] M. Watter, J. Springenberg, J. Boedecker, and M. Riedmiller, "Embed to control: A locally linear latent dynamics model for control from raw images," in *Advances in neural information processing systems*, 2015, pp. 2746–2754.
- [11] M. Fraccaro, S. Kamronn, U. Paquet, and O. Winther, "A disentangled recognition and nonlinear dynamics model for unsupervised learning," in *Advances in Neural Information Processing Systems*, 2017, pp. 3601–3610.
- [12] D. Hafner, T. Lillicrap, I. Fischer, R. Villegas, D. Ha, H. Lee, and J. Davidson, "Learning latent dynamics for planning from pixels," in *International Conference on Machine Learning*, 2019, pp. 2555–2565.
- [13] I. Higgins, D. Amos, D. Pfau, S. Racaniere, L. Matthey, D. Rezende, and A. Lerchner, "Towards a definition of disentangled representations," *arXiv preprint arXiv:1812.02230*, 2018.
- [14] S. van Steenkiste, F. Locatello, J. Schmidhuber, and O. Bachem, "Are disentangled representations helpful for abstract visual reasoning?" in *Advances in Neural Information Processing Systems*, 2019, pp. 14 222–14 235.
- [15] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, "beta-vae: Learning basic visual concepts with a constrained variational framework," in *International Conference on Learning Representations*, 2017.
- [16] C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner, "Understanding disentangling in  $\beta$ -vae," *arXiv preprint arXiv:1804.03599*, 2018.
- [17] T. Q. Chen, X. Li, R. B. Grosse, and D. K. Duvenaud, "Isolating sources of disentanglement in variational autoencoders," in *Advances in Neural Information Processing Systems*, 2018, pp. 2610–2620.
- [18] H. Kim and A. Mnih, "Disentangling by factorising," in *International Conference on Machine Learning*, 2018, pp. 2654–2663.
- [19] C. Tsallis, "Possible generalization of boltzmann-gibbs statistics," *Journal of statistical physics*, vol. 52, no. 1-2, pp. 479–487, 1988.
- [20] H. Suyari and M. Tsukada, "Law of error in tsallis statistics," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 753–757, 2005.
- [21] S. Umarov, C. Tsallis, and S. Steinberg, "On a q-central limit theorem consistent with nonextensive statistical mechanics," *Milan journal of mathematics*, vol. 76, no. 1, pp. 307–328, 2008.
- [22] F. Nielsen and R. Nock, "A closed-form expression for the sharma-mittal entropy of exponential families," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 3, p. 032003, 2011.
- [23] M. Gil, F. Alajaji, and T. Linder, "Rényi divergence measures for commonly used univariate continuous distributions," *Information Sciences*, vol. 249, pp. 124–131, 2013.
- [24] H. Takahashi, T. Iwata, Y. Yamanaka, M. Yamada, and S. Yagi, "Student-t variational autoencoder for robust density estimation," in *International Joint Conference on Artificial Intelligence*, 2018, pp. 2696–2702.
- [25] M. Karl, M. Soelch, J. Bayer, and P. van der Smagt, "Deep variational bayes filters: Unsupervised learning of state space models from raw data," in *International Conference on Learning Representations*, 2017.
- [26] M. Girolami and C. Fyfe, "Negentropy and kurtosis as projection pursuit indices provide generalised ica algorithms," in *Advances in Neural Information Processing Systems Workshop*, vol. 9, 1996.
- [27] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, vol. 57, no. 3, pp. 519–530, 1970.
- [28] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *Advances in Neural Information Processing Systems Workshop*, 2017.
- [29] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.
- [30] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv preprint arXiv:1710.05941*, 2017.
- [31] S. Elfving, E. Uchibe, and K. Doya, "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning," *Neural Networks*, vol. 107, pp. 3–11, 2018.
- [32] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *arXiv preprint arXiv:1606.01540*, 2016.
- [33] M. A. Lewis, A. H. Fagg, and G. A. Bekey, "Genetic algorithms for gait synthesis in a hexapod robot," in *Recent trends in mobile robots*. World Scientific, 1993, pp. 317–331.
- [34] D. Owaki, M. Goda, S. Miyazawa, and A. Ishiguro, "A minimal model describing hexapedal interlimb coordination: the tegotae-based approach," *Frontiers in neurorobotics*, vol. 11, p. 29, 2017.
- [35] W. E. L. Ilboudo, T. Kobayashi, and K. Sugimoto, "Tadam: A robust stochastic gradient optimizer," *arXiv preprint arXiv:2003.00179*, 2020.