

## A. Proof of Corollary 1

**Corollary** (Discrete case). *Suppose  $\mathbf{x} \in \mathcal{X}$  where  $\mathcal{X}$  is the set of all one-hot vectors of length  $n$ , and let  $f_\psi : \mathcal{Z} \rightarrow \Delta^{n-1}$  be a deterministic function that goes from the latent space  $\mathcal{Z}$  to the  $n - 1$  dimensional simplex  $\Delta^{n-1}$ . Further let  $G_\psi : \mathcal{Z} \rightarrow \mathcal{X}$  be a deterministic function such that  $G_\psi(\mathbf{z}) = \arg \max_{\mathbf{w} \in \mathcal{X}} \mathbf{w}^\top f_\psi(\mathbf{z})$ , and as above let  $\mathbb{P}_\psi(\mathbf{x} | \mathbf{z})$  be the dirac distribution derived from  $G_\psi$  such that  $p_\psi(\mathbf{x} | \mathbf{z}) = \mathbb{1}\{\mathbf{x} = G_\psi(\mathbf{z})\}$ . Then the following is an upper bound on  $\|\mathbb{P}_\psi - \mathbb{P}_\star\|_{TV}$ , the total variation distance between  $\mathbb{P}_\star$  and  $\mathbb{P}_\psi$ .*

$$\inf_{Q: \mathbb{P}_Q = \mathbb{P}_\mathbf{z}} \mathbb{E}_{\mathbb{P}_\star} \mathbb{E}_{Q(\mathbf{z} | \mathbf{x})} \left[ -\frac{2}{\log 2} \log \mathbf{x}^\top f_\psi(\mathbf{z}) \right]$$

*Proof.* Let our cost function be  $c(\mathbf{x}, \mathbf{y}) = \mathbb{1}\{\mathbf{x} \neq \mathbf{y}\}$ . We first note that for all  $\mathbf{x}, \mathbf{z}$

$$\log 2 \mathbb{1}\{\mathbf{x} \neq \arg \max_{\mathbf{w} \in \mathcal{X}} \mathbf{w}^\top f_\psi(\mathbf{z})\} < -\log \mathbf{x}^\top f_\psi(\mathbf{z})$$

This holds since if  $\mathbb{1}\{\mathbf{x} \neq \arg \max_{\mathbf{w} \in \mathcal{X}} \mathbf{w}^\top f_\psi(\mathbf{z})\} = 1$ , we have  $\mathbf{x}^\top f_\psi(\mathbf{z}) < 0.5$ , and  $-\log \mathbf{x}^\top f_\psi(\mathbf{z}) > -\log 0.5 = \log 2$ . If on the other hand  $\mathbf{x} = \arg \max_{\mathbf{w} \in \mathcal{X}} \mathbf{w}^\top f_\psi(\mathbf{z})$ , then the LHS is 0 and RHS is always positive since  $f_\psi(\mathbf{z}) \in \Delta^{n-1}$ . Then,

$$\begin{aligned} & \inf_{Q: \mathbb{P}_Q = \mathbb{P}_\mathbf{z}} \mathbb{E}_{\mathbb{P}_\star} \mathbb{E}_{Q(\mathbf{z} | \mathbf{x})} \left[ -\frac{2}{\log 2} \log \mathbf{x}^\top f_\psi(\mathbf{z}) \right] \\ & > \inf_{Q: \mathbb{P}_Q = \mathbb{P}_\mathbf{z}} \mathbb{E}_{\mathbb{P}_\star} \mathbb{E}_{Q(\mathbf{z} | \mathbf{x})} [2 \mathbb{1}\{\mathbf{x} \neq \arg \max_{\mathbf{w} \in \mathcal{X}} \mathbf{w}^\top f_\psi(\mathbf{z})\}] \\ & = 2 \inf_{Q: \mathbb{P}_Q = \mathbb{P}_\mathbf{z}} \mathbb{E}_{\mathbb{P}_\star} \mathbb{E}_{Q(\mathbf{z} | \mathbf{x})} [\mathbb{1}\{\mathbf{x} \neq G_\psi(\mathbf{z})\}] \\ & = 2 \inf_{Q: \mathbb{P}_Q = \mathbb{P}_\mathbf{z}} \mathbb{E}_{\mathbb{P}_\star} \mathbb{E}_{Q(\mathbf{z} | \mathbf{x})} [c(\mathbf{x}, G_\psi(\mathbf{z}))] \\ & = 2W_c(\mathbb{P}_\star, \mathbb{P}_\psi) \\ & = \|\mathbb{P}_\star - \mathbb{P}_\psi\|_{TV} \end{aligned}$$

The fifth line follows from Theorem 1, and the last equality uses the well-known correspondence between total variation distance and optimal transport with the indicator cost function (Gozlan & Léonard, 2010).  $\square$

## B. Optimality Property

One can interpret the ARAE framework as a dual pathway network mapping two distinct distributions into a similar one;  $\text{enc}_\phi$  and  $g_\theta$  both output code vectors that are kept similar in terms of Wasserstein distance as measured by the critic. We provide the following proposition showing that under our parameterization of the encoder and the generator, as the Wasserstein distance converges, the encoder distribution ( $\mathbb{P}_Q$ ) converges to the generator distribution ( $\mathbb{P}_\mathbf{z}$ ), and further, their moments converge.

This is ideal since under our setting the generated distribution is simpler than the encoded distribution, because the

input to the generator is from a simple distribution (e.g. spherical Gaussian) and the generator possesses less capacity than the encoder. However, it is not so simple that it is overly restrictive (e.g. as in VAEs). Empirically we observe that the first and second moments do indeed converge as training progresses (Section 7).

**Proposition 1.** *Let  $\mathbb{P}$  be a distribution on a compact set  $\mathcal{X}$ , and  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  be a sequence of distributions on  $\mathcal{X}$ . Further suppose that  $W(\mathbb{P}_n, \mathbb{P}) \rightarrow 0$ . Then the following statements hold:*

(i)  $\mathbb{P}_n \rightsquigarrow \mathbb{P}$  (i.e. convergence in distribution).

(ii) All moments converge, i.e. for all  $k > 1, k \in \mathbb{N}$ ,

$$\mathbb{E}_{X \sim \mathbb{P}_n} \left[ \prod_{i=1}^d X_i^{p_i} \right] \rightarrow \mathbb{E}_{X \sim \mathbb{P}} \left[ \prod_{i=1}^d X_i^{p_i} \right]$$

for all  $p_1, \dots, p_d$  such that  $\sum_{i=1}^d p_i = k$

*Proof.* (i) has been proved in (Villani, 2008) Theorem 6.9.

For (ii), using *The Portmanteau Theorem*, (i) is equivalent to the following statement:

$\mathbb{E}_{X \sim \mathbb{P}_n} [f(X)] \rightarrow \mathbb{E}_{X \sim \mathbb{P}} [f(X)]$  for all bounded and continuous function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ , where  $d$  is the dimension of the random variable.

The  $k$ -th moment of a distribution is given by

$$\mathbb{E} \left[ \prod_{i=1}^d X_i^{p_i} \right] \text{ such that } \sum_{i=1}^d p_i = k$$

Our encoded code is bounded as we normalize the encoder output to lie on the unit sphere, and our generated code is also bounded to lie in  $(-1, 1)^n$  by the tanh function. Hence  $f(X) = \prod_{i=1}^d X_i^{q_i}$  is a bounded continuous function for all  $q_i \geq 0$ . Therefore,

$$\mathbb{E}_{X \sim \mathbb{P}_n} \left[ \prod_{i=1}^d X_i^{p_i} \right] \rightarrow \mathbb{E}_{X \sim \mathbb{P}} \left[ \prod_{i=1}^d X_i^{p_i} \right]$$

where  $\sum_{i=1}^d p_i = k$   $\square$

## C. Sample Generations

In Figure 6 we show some generated samples from the ARAE, AE, and a LM.

## D. Sentence Interpolations

In Figure 7 we show generations from interpolated latent vectors. Specifically, we sample two points  $\mathbf{z}_0$  and  $\mathbf{z}_1$  from

ARAE Samples	AE Samples	LM Samples
A woman preparing three fish . A woman is seeing a man in the river . There passes a woman near birds in the air . Some ten people is sitting through their office . The man got stolen with young dinner bag . Monks are running in court .  The Two boys in glasses are all girl . The man is small sitting in two men that tell a children .  The two children are eating the balloon animal . A woman is trying on a microscope . The dogs are sleeping in bed .	Two Three woman in a cart tearing over of a tree . A man is hugging and art . The fancy skier is starting under the drag cup in . A dog are <unk> a A man is not standing . The Boys in their swimming . A surfer and a couple waiting for a show . A couple is a kids at a barbecue . The motorcycles is in the ocean loading I 's bike is on empty The actor was walking in a small dog area . no dog is young their mother	a man walking outside on a dirt road , sitting on the dock . A large group of people is taking a photo for Christmas and at night . Someone is avoiding a soccer game . The man and woman are dressed for a movie . Person in an empty stadium pointing at a mountain . Two children and a little boy are <unk> a man in a blue shirt . A boy rides a bicycle . A girl is running another in the forest . the man is an indian women .

Figure 6: Text samples generated from ARAE, a simple AE, and from a baseline LM trained on the same data. To generate from an AE we fit a multivariate Gaussian to the learned code space and generate code vectors from this Gaussian.

A man is on the corner in a sport area . A man is on corner in a road all . A lady is on outside a race-track . A lady is outside on a race-track . A lot of people is outdoors in an urban setting . A lot of people is outdoors in an urban setting . A lot of people is outdoors in an urban setting .	A man is on a ship path with the woman . A man is on a ship path with the woman . A man is passing on a bridge with the girl . A man is passing on a bridge with the girl . A man is passing on a bridge with the girl . A man is passing on a bridge with the dogs . A man is passing on a bridge with the dogs .	A man in a cave is used an escalator . A man in a cave is used an escalator A man in a cave is used chairs . A man in a number is used many equipment A man in a number is posing so on a big rock . People are posing in a rural area . People are posing in a rural area.
---	--	---

Figure 7: Sample interpolations from the ARAE. Constructed by linearly interpolating in the latent space and decoding to the output space. Word changes are highlighted in black.

$p(\mathbf{z})$  and construct intermediary points  $\mathbf{z}_\lambda = \lambda \mathbf{z}_1 + (1 - \lambda) \mathbf{z}_0$ . For each we generate the argmax output  $\tilde{\mathbf{x}}_\lambda$ .

### E. Vector Arithmetic

We generate 1 million sentences from the ARAE and parse the sentences to obtain the main verb, subject, and modifier. Then for a given sentence, to change the main verb we subtract the mean latent vector ( $\mathbf{t}$ ) for all other sentences with the same main verb (in the first example in Figure 5 this would correspond to all sentences that had “sleeping” as the main verb) and add the mean latent vector for all sentences that have the desired transformation (with the running example this would be all sentences whose main verb was “walking”). We do the same to transform the subject and the modifier. We decode back into sentence space with the transformed latent vector via sampling from  $p_\psi(g(\mathbf{z} + \mathbf{t}))$ . Some examples of successful transformations are shown in Figure 5 (right). Quantitative evaluation of the success of the vector transformations is given in Figure 5 (left). For each original vector  $\mathbf{z}$  we sample 100 sentences from  $p_\psi(g(\mathbf{z} + \mathbf{t}))$  over the transformed new latent vector and

consider it a match if *any* of the sentences demonstrate the desired transformation. Match % is proportion of original vectors that yield a match post transformation. As we ideally want the generated samples to only differ in the specified transformation, we also calculate the average word precision against the original sentence (Prec) for any match.

## F. Experimental Details

### MNIST experiments

- The encoder is a three-layer MLP, 784–800–400–100.
- Additive Gaussian noise is injected into  $\mathbf{c}$  then gradually decayed to 0.
- The decoder is a four-layer MLP, 100–400–800–1000–784
- The autoencoder is optimized by Adam, with learning rate  $5e-04$ .
- An MLP generator 32–64–100–150–100.
- An MLP critic 100–100–60–20–1 with weight clipping  $\epsilon = 0.05$ . The critic is trained 10 iterations in every loop.
- GAN is optimized by Adam, with learning rate  $5e-04$  on the generator and  $5e-05$  on the critic.
- Weighing factor  $\lambda^{(1)} = 0.2$ .

### Text experiments

- The encoder is an one-layer LSTM with 300 hidden units.
- Additive Gaussian noise is injected into  $\mathbf{c}$  then gradually decayed to 0.
- The decoder is an one-layer LSTM with 300 hidden units.
- The LSTM state vector is augmented by the hidden code  $\mathbf{c}$  at every decoding time step, before forwarding into the output softmax layer.
- The word embedding is of size 300.
- The autoencoder is optimized by SGD with learning rate 1. A grad clipping on the autoencoder, with max `grad_norm` set to 1.
- An MLP generator 100–300–300.
- An MLP critic 300–300–1 with weight clipping  $\epsilon = 0.01$ . The critic is trained 5 iterations in every loop.
- GAN is optimized by Adam, with learning rate  $5e-05$  on the generator, and  $1e-05$  on the critic.

### Semi-supervised experiments

The following changes are made based on the SNLI experiments:

- Larger network to GAN components: an MLP generator 100-150-300-500 and an MLP critic 500-500-150-80-20-1 with weight clipping factor  $\epsilon = 0.02$ .

### Yelp/Yahoo transfer

- An MLP style adversarial classifier 300-200-100, trained by SGD learning rate 0.1.
- Weighing factor from both adversarial forces  $\lambda_a^{(1)} = 1$ ,  $\lambda_b^{(1)} = 10$ .

### G. Style Transfer Samples

In the following pages we show randomly sampled style transfers from the Yelp/Yahoo corpus.

Yelp Sentiment Transfer

Positive to Negative		Negative to Positive	
Original	great indoor mall .	Original	hell no !
ARAE	no smoking mall .	ARAE	hell great !
Cross-AE	terrible outdoor urine .	Cross-AE	incredible pork !
Original	great blooming onion .	Original	highly disappointed !
ARAE	no receipt onion .	ARAE	highly recommended !
Cross-AE	terrible of pie .	Cross-AE	highly clean !
Original	i really enjoyed getting my nails done by peter .	Original	bad products .
ARAE	i really needed getting my nails done by now .	ARAE	good products .
Cross-AE	i really really told my nails done with these things .	Cross-AE	good prices .
Original	definitely a great choice for sushi in las vegas !	Original	i was so very disappointed today at lunch .
ARAE	definitely a _num_ star rating for _num_ sushi in las vegas .	ARAE	i highly recommend this place today .
Cross-AE	not a great choice for breakfast in las vegas vegas !	Cross-AE	i was so very pleased to this .
Original	the best piece of meat i have ever had !	Original	i have n't received any response to anything .
ARAE	the worst piece of meat i have ever been to !	ARAE	i have n't received any problems to please .
Cross-AE	the worst part of that i have ever had had !	Cross-AE	i have always the desert vet .
Original	really good food , super casual and really friendly .	Original	all the fixes were minor and the bill ?
ARAE	really bad food , really generally really low and decent food .	ARAE	all the barbers were entertaining and the bill did n't disappoint .
Cross-AE	really good food , super horrible and not the price .	Cross-AE	all the flavors were especially and one !
Original	it has a great atmosphere , with wonderful service .	Original	small , smokey , dark and rude management .
ARAE	it has no taste , with a complete jerk .	ARAE	small , intimate , and cozy friendly staff .
Cross-AE	it has a great horrible food and run out service .	Cross-AE	great , , , chips and wine .
Original	their menu is extensive , even have italian food .	Original	the restaurant did n't meet our standard though .
ARAE	their menu is limited , even if i have an option ,	ARAE	the restaurant did n't disappoint our expectations though .
Cross-AE	their menu is decent , i have gotten italian food .	Cross-AE	the restaurant is always happy and knowledge .
Original	everyone who works there is incredibly friendly as well .	Original	you could not see the stage at all !
ARAE	everyone who works there is incredibly rude as well .	ARAE	you could see the difference at the counter !
Cross-AE	everyone who works there is extremely clean and as well .	Cross-AE	you could definitely get the fuss !
Original	there are a couple decent places to drink and eat in here as well .	Original	room is void of all personality , no pictures or any sort of decorations .
ARAE	there are a couple slices of options and _num_ wings in the place .	ARAE	room is eclectic , lots of flavor and all of the best .
Cross-AE	there are a few night places to eat the car here are a crowd .	Cross-AE	it 's a nice that amazing , that one 's some of flavor .
Original	if you 're in the mood to be adventurous , this is your place !	Original	waited in line to see how long a wait would be for three people .
ARAE	if you 're in the mood to be disappointed , this is not the place .	ARAE	waited in line for a long wait and totally worth it .
Cross-AE	if you 're in the drive to the work , this is my place !	Cross-AE	another great job to see and a lot going to be from dinner .
Original	we came on the recommendation of a bell boy and the food was amazing .	Original	the people who ordered off the menu did n't seem to do much better .
Cross-AE	we came on the recommendation and the food was a joke .	ARAE	the people who work there are super friendly and the menu is good .
Cross-AE	we went on the car of the time and the chicken was awful .	Cross-AE	the place , one of the office is always worth you do a business .
Original	service is good but not quick , just enjoy the wine and your company .	Original	they told us in the beginning to make sure they do n't eat anything .
ARAE	service is good but not quick , but the service is horrible .	ARAE	they told us in the mood to make sure they do great food .
Cross-AE	service is good , and horrible , is the same and worst time ever .	Cross-AE	they 're us in the next for us as you do n't eat .
Original	the steak was really juicy with my side of salsa to balance the flavor .	Original	the person who was teaching me how to control my horse was pretty rude .
ARAE	the steak was really bland with the sauce and mashed potatoes .	ARAE	the person who was able to give me a pretty good price .
Cross-AE	the fish was so much , the most of sauce had got the flavor .	Cross-AE	the owner 's was gorgeous when i had a table and was friendly .
Original	other than that one hell hole of a star bucks they 're all great !	Original	he was cleaning the table next to us with gloves on and a rag .
ARAE	other than that one star rating the toilet they 're not allowed .	ARAE	he was prompt and patient with us and the staff is awesome .
Cross-AE	a wonder our one came in a _num_ months , you 're so better !	Cross-AE	he was like the only thing to get some with with my hair .

Figure 8: Full sheet of sentiment transfer result on the Yelp corpus.



**Adversarially Regularized Autoencoders**

**Yahoo Topic Transfer on Answers**

	from Science		from Music		from Politics
Original	take 1ml of hcl ( concentrated ) and dilute it to 50ml .	Original	all three are fabulous artists , with just incredible talent ! !	Original	4 years of an idiot in office + electing the idiot again = ?
Music	take em to you and shout it to me	Science	all three are genetically bonded with water , but just as many substances , are capable of producing a special case .	Science	4 years of an idiot in the office of science ?
Politics	take bribes to islam and it will be punished .	Politics	all three are competing with the government , just as far as i can .	Music	4 ) <unk> in an idiot , the idiot is the best of the two points ever !
Original	oils do not do this , they do not &quot; set &quot; ;	Original	she , too , wondered about the underwear outside the clothes .	Original	send me \$ 100 and i &apos;ll send you a copy - honest .
Music	cucumbers do not do this , they do not &quot; do &quot; ;	Science	she , too , i know , the clothes outside the clothes .	Science	send me an email and i &apos;ll send you a copy .
Politics	corporations do not do this , but they do not .	Politics	she , too , i think that the cops are the only thing about the outside of the u.s .	Music	send me \$ 100 and i &apos;ll send you a copy .
Original	the average high temps in jan and feb are about 48 deg .	Original	i like rammstein and i don &apos;t speak or understand german .	Original	willis can be <unk> , or typed and signed without needing an attorney .
Music	the average high school in seattle and is about 15 minutes .	Science	i like googling and i don &apos;t understand or speak .	Science	euler can be <unk> , and without any type of operations , or <unk> .
Politics	the average high infantry division is in afghanistan and alaska .	Politics	i like mccain and i don &apos;t care about it .	Music	madonna can be <unk> , and signed without opening or <unk> .
Original	the light from you lamps would move away from you at light speed	Original	mark is great , but the guest hosts were cool too !	Original	hungary : 20 january 1945 , ( formerly a member of the axis )
Music	the light from you tube would move away from you	Science	mark is great , but the water will be too busy for the same reason .	Science	nh3 : 20 january , 78 ( a )
Politics	the light from you could go away from your state	Politics	mark twain , but the great lakes , the united states of america is too busy .	Music	1966 - 20 january 1961 ( a ) 1983 song
Original	van <unk> , on the other hand , had some serious issues ...	Original	they all offer terrific information about the cast and characters , ...	Original	bulgaria : 8 september 1944 , ( formerly a member of the axis )
Music	van <unk> on the other hand , had some serious issues .	Science	they all offer insight about the characteristics of the earth , and are composed of many stars .	Science	moreover , $8\sqrt{3} + (x+7)(x^2) = (a^2)$
Politics	van <unk> , on the other hand , had some serious issues .	Politics	they all offer legitimate information about the invasion of iraq and the u.s . , and all aspects of history .	Music	harrison : 8 september 1961 ( a ) ( 1995 )
Original	just multiply the numerator of one fraction by that of the other .	Original	but there are so many more i can &apos;t think of !	Original	anyone who doesnt have a billion dollars for all the publicity cant win .
Music	just multiply the fraction of the other one that &apos;s just like it .	Science	but there are so many more of the number of questions .	Science	anyone who doesnt have a decent chance is the same for all the other .
Politics	just multiply the same fraction of other countries .	Politics	but there are so many more of the can i think of today .	Music	anyone who doesnt have a lot of the show for the publicity .
Original	civil engineering is still an umbrella field comprised of many related specialties .	Original	i love zach he is sooo sweet in his own way !	Original	the theory is that cats don &apos;t take to being tied up but thats <unk> .
Music	civil rights is still an art union .	Science	the answer is he &apos;s definitely in his own way !	Science	the theory is that cats don &apos;t grow up to <unk> .
Politics	civil law is still an issue .	Politics	i love letting he is sooo smart in his own way !	Music	the theory is that dumb but don &apos;t play <unk> to <unk> .
Original	h2o2 ( hydrogen peroxide ) naturally decomposes to form o2 and water .	Original	remember the industry is very shady so keep your eyes open !	Original	the fear they are trying to instill in the common man is based on what ?
Music	jackie and brad pitt both great albums and they are my fav .	Science	remember the amount of water is so very important .	Science	the fear they are trying to find the common ancestor in the world .
Politics	kennedy and blair hate america to invade them .	Politics	remember the amount of time the politicians are open your mind .	Music	the fear they are trying to find out what is wrong in the song .
Original	the quieter it gets , the more white noise you can here .	Original	but can you fake it , for just one more show ?	Original	think about how much planning and people would have to be involved in what happened .
Music	the fray it gets , the more you can hear .	Science	but can you fake it , just for more than one ?	Science	think about how much time would you have to do .
Politics	the gop gets it , the more you can here .	Politics	but can you fake it for more than one ?	Music	think about how much money and what would be <unk> about in the world ?
Original	h2co3 ( carbonic acid ) naturally decomposes to form water and co2 .	Original	i am going to introduce you to the internet movie database .	Original	this restricts the availability of cash to them and other countries too start banning them .
Music	phoebe and jack , he &apos;s gorgeous and she loves to get him !	Science	i am going to investigate the internet to google .	Science	this reduces the intake of the other molecules to produce them and thus are too large .
Politics	nixon ( captured ) he lied and voted for bush to cause his country .	Politics	i am going to skip the internet to get you checked .	Music	this is the cheapest package of them too .

Figure 10: Full sheet of Yahoo topic transfer on answers.