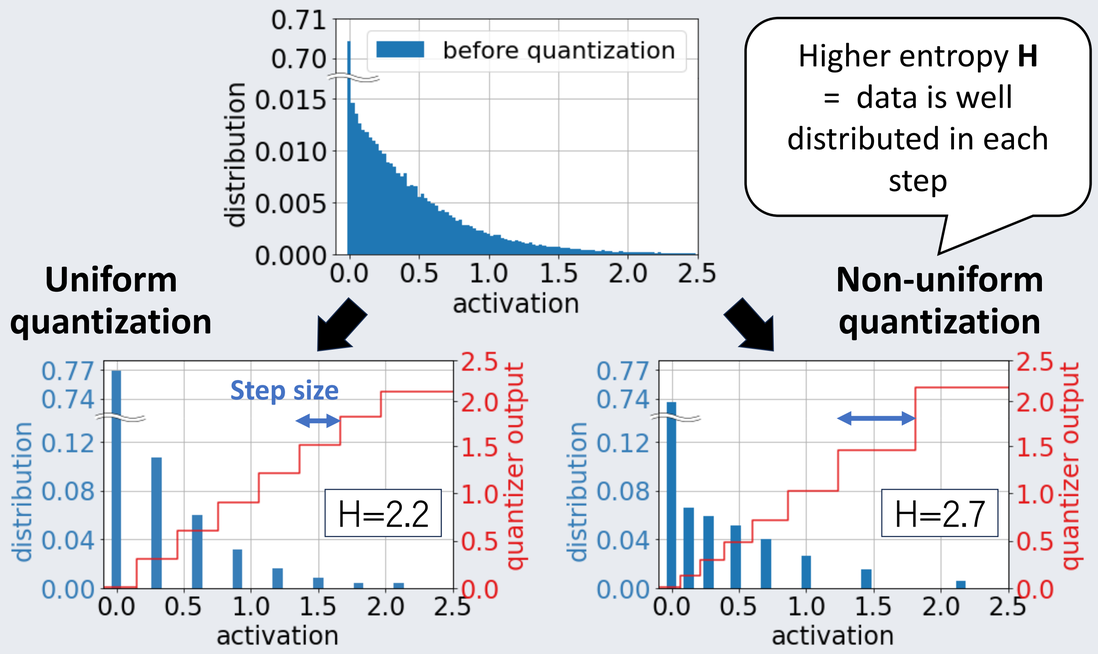
Learning Non-Uniform Step-Sizes for Neural Network Quantization

S. Gongyo[1], J.R. Liang[2], M. Ambai[1], R. Kawakami[2], I. Sato[1,2]

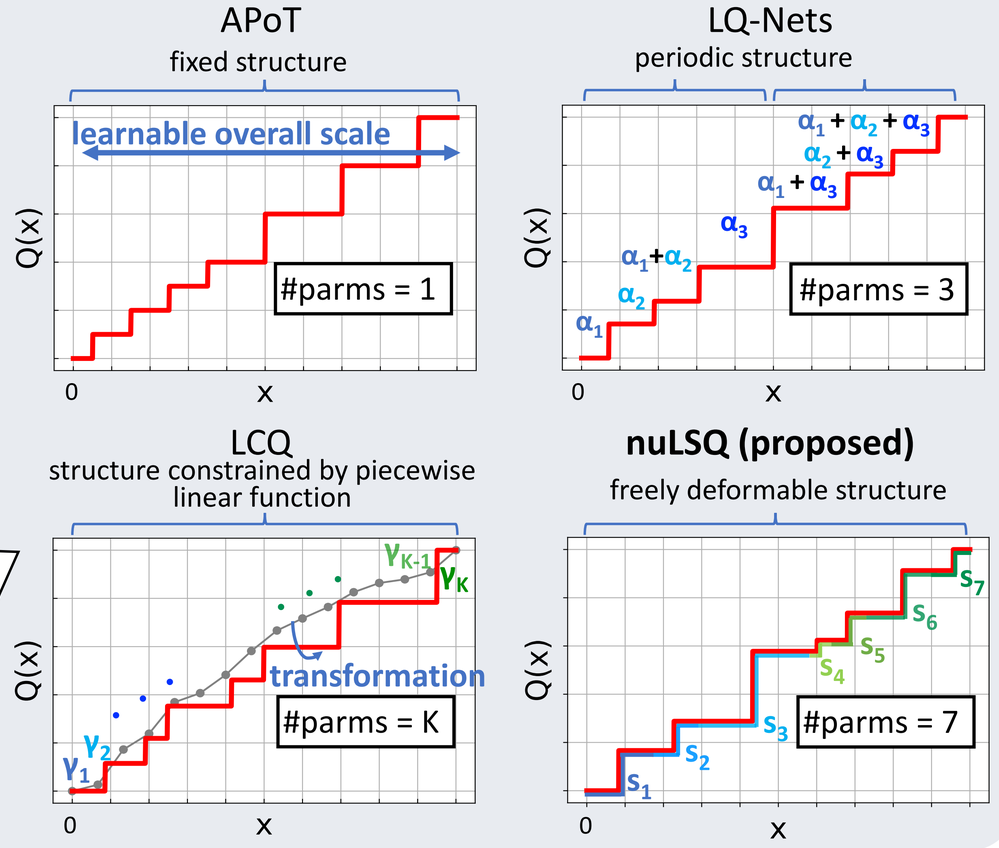
[1]Denso IT Laboratory, Inc., [2]Institute of Science Tokyo

Introduction

- Quantization typically assumes a uniform step size
- Use of **non-uniform** step size can better capture the distribution



nuLSQ has high flexibility among non-uniform methods



Learned Step-Size Quantization (LSQ)

[Esser, Steven K.+. ICLR 2020]

Forward Pass

$$Q_{LSQ}(x, s) = \sum_{n=1}^{N} s\sigma\left(x - ns + \frac{s}{2}\right)$$

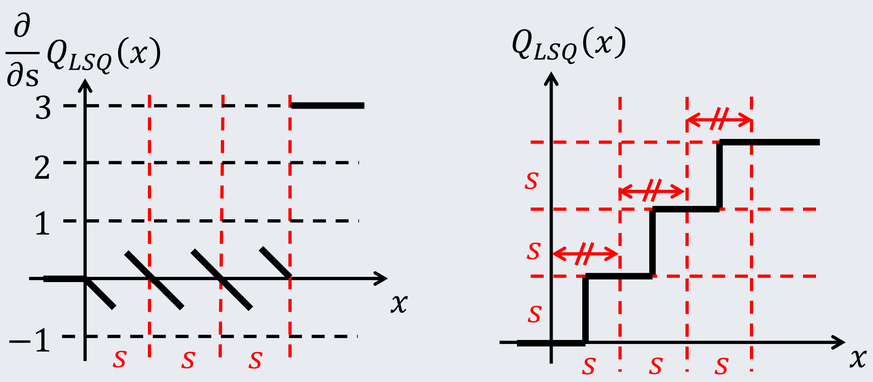
Uniform step size s

N ← The number of step-size N

$\sigma(\cdot)$ Unit step function

Backward Pass

- Uniform step-size gradient approximated with straight through estimator (STE).



Proposed Method

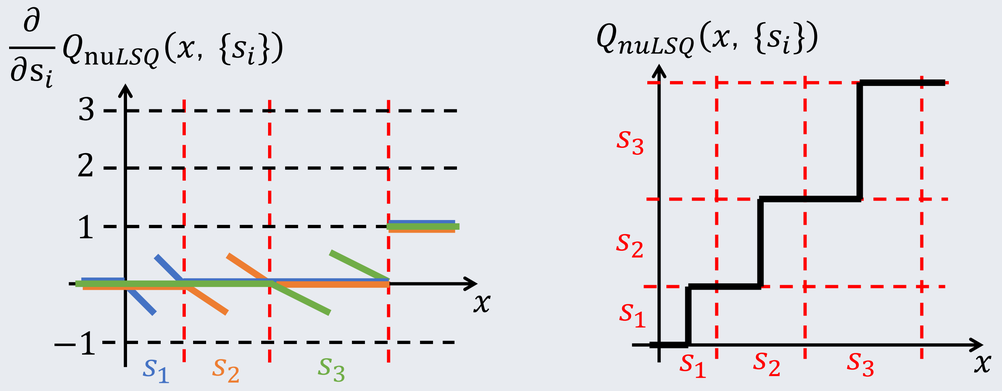
Forward Pass

$$Q_{nuLSQ}(x, \{s_i\}) = \sum_{n=1}^{N} s_n \sigma\left(x - \left(\sum_{m=1}^{n-1} s_m + \frac{s_n}{2}\right)\right)$$

Non-uniform step sizes $\{s_i\}$

Backward Pass

- Non-uniform step-sizes gradient approximated with straight through estimator (STE).



Experiments

ImageNet top-1 accuracy (%) of MobileNetV2 (FP:72.9)

Bits (WA)	LSQ	LSQ +BR	LSQ +Dp/Fz	UniQ	LCQ	LLSQ	EWGS	nuLSQA
2/2	46.7	50.6	-	50.5	-	-	-	**58.4**
3/3	65.3	67.4	67.8	65.0	-	-	-	**67.9**
4/4	69.5	70.4	70.6	68.2	70.8	67.4	70.3	**71.1**

ImageNet top-1 accuracy (%) of ResNet-18 (FP:69.76)

PACT*	DoReFa*	LSQ*	APoT†	LQ-Nets*	LCQ*	nuLSQA
62.48	63.28	64.51	64.41	63.71	64.67	**64.89**

CIFAR100 top-1 accuracy (%) comparison to LSQ

Network	Methods	Bits (W/A)		
		2/2	3/3	4/4
ResNet-20 (FP:69.8)	LSQ*	65.82	68.60	**69.56**
	nuLSQ-A	**66.02**	68.58	69.42
	nuLSQ-W	66.00	**68.70**	69.40
ResNet-56 (FP:74.9)	LSQ*	70.50	72.62	**73.48**
	nuLSQ-A	70.66	**72.98**	**73.48**
	nuLSQ-W	**70.82**	72.82	73.42

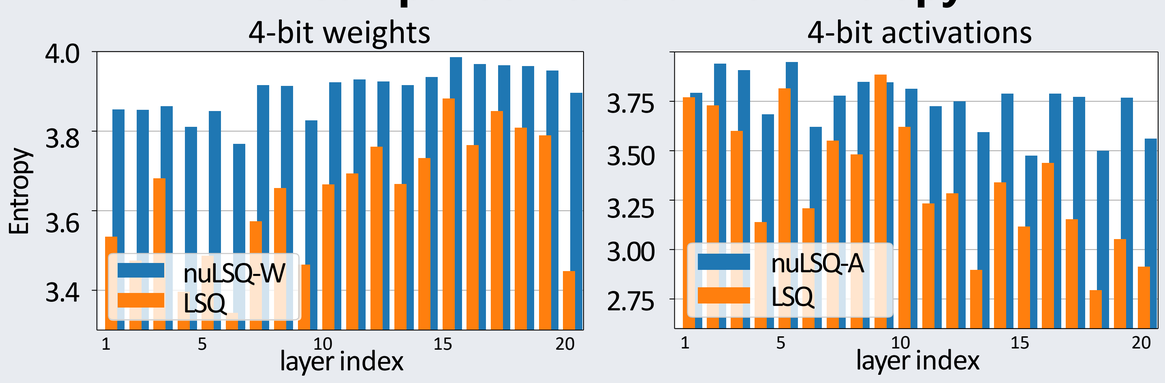
ImageNet top-1 accuracy (%) comparison

Methods	Swin-T (FP:81.2)			ConvNext (FP:81.87)
	2/2	3/3	4/4	3/3
LSQ*	74.58	77.48	78.33	72.9
nuLSQ-WA	**74.91**	**77.71**	**78.37**	**73.39**

Comparison to QAT methods with distillation

Network	Methods	Top-1 Acc (%)	Top-5 Acc (%)
MobileNet V2	QKD	67.4	87.0
	PROFIT	71.56	90.40
	nuLSQ-A	**71.89**	**90.44**

Comparison of Shannon Entropy



This work was supported by DENSO IT LAB Recognition and Learning Algorithm Collaborative Research Chair (Science Tokyo).