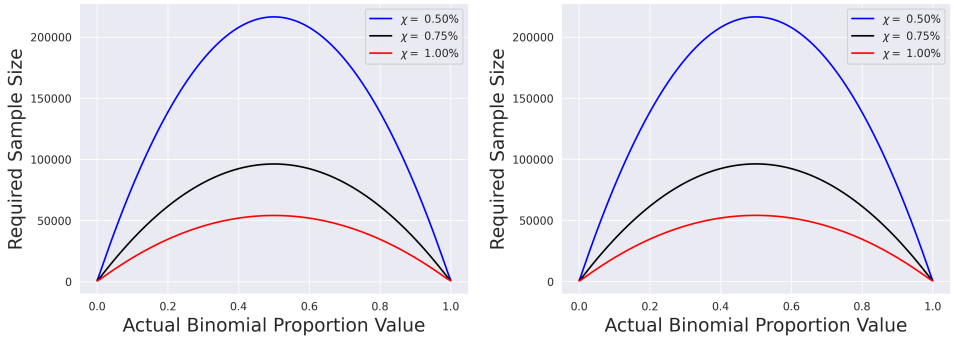
A Appendix

A.1 Observation for Binomial Confidence Interval Methods

In this section, we show the plots for the number of samples required to estimate an unknown binomial proportion parameter through two popular estimation techniques - the Wilson (Wilson, 1927) and Agresti-Coull method (Agresti and Coull, 1998). For this experiment, we use three different values of the target error χ = 0.5 %, 0.75 %, and 1.0 % and a fixed confidence value $(1-\alpha) = 0.99$ for both estimation methods. As shown in Fig 4, for a fixed target error χ, confidence $(1-\alpha)$, and estimation technique, the number of samples required for estimation peaks, when the actual parameter value is around 0.5 and is the smallest around the boundaries. This is consistent with the observation described in Section 3.1.



(a) Agresti-Coull method

(b) Wilson method

Figure 4: The number of samples for the Agresti-Coull and Wilson method to achieve a target error χ with confidence $(1-\alpha)$ where $\alpha = 0.01$. The plots show that the number of required samples for different methods peaks at 0.5 and decreases towards the boundaries.

A.2 Theorems

Theorem 2. *If a classifier f^p is such that for all $x \in \mathbb{R}^m, \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon)) \leq \zeta_x$, and classifier f satisfies $\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \geq \underline{p_A} \geq \overline{p_B} \geq \max_{c \neq c_A} \mathbb{P}_\epsilon(f(x+\epsilon) = c)$ and $\underline{p_A} - \zeta_x \geq \overline{p_B} + \zeta_x$ then g^p satisfies $g^p(x+\delta) = c_A$ for all δ satisying $\|\delta\|_2 \leq \frac{\sigma}{2}(\Phi^{-1}(\underline{p_A} - \zeta_x) - \Phi^{-1}(\overline{p_B} + \zeta_x))$*

Proof. If $f(x+\epsilon) = c_A$ and $f^p(x+\epsilon) = f(x+\epsilon)$ then $f^p(x+\epsilon) = c_A$.
Thus, if $f^p(x+\epsilon) \neq c_A$ then $f(x+\epsilon) \neq c_A$ or $f^p(x+\epsilon) \neq f(x+\epsilon)$.
Using union bound,

$$\mathbb{P}_\epsilon(f^p(x+\epsilon) \neq c_A) \leq \mathbb{P}_\epsilon(f(x+\epsilon) \neq c_A) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$(1 - \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A)) \leq (1 - \mathbb{P}_\epsilon(f(x+\epsilon) = c_A)) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$\underline{p_A} - \zeta_x \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A)$$

Similarly, if $f(x+\epsilon) \neq c$ then $f^p(x+\epsilon) \neq c$ or $f^p(x+\epsilon) \neq f(x+\epsilon)$.
Hence, using union bound,

$$\mathbb{P}_\epsilon(f(x+\epsilon) \neq c) \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) \neq c) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$(1 - \mathbb{P}_\epsilon(f(x+\epsilon) = c)) \leq (1 - \mathbb{P}_\epsilon(f^p(x+\epsilon) = c)) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$\mathbb{P}_\epsilon(f^p(x+\epsilon) = c) \leq \mathbb{P}_\epsilon(f(x+\epsilon) = c) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$\max_{c \neq c_A} \mathbb{P}_\epsilon(f^p(x+\epsilon) = c) \leq \max_{c \neq c_A} \mathbb{P}_\epsilon(f(x+\epsilon) = c) + \zeta_x$$
$$\max_{c \neq c_A} \mathbb{P}_\epsilon(f^p(x+\epsilon) = c) \leq \overline{p_B} + \zeta_x$$

Hence, using Theorem 1, g^p satisfies $g^p(x+\delta) = c_A$ for all δ satisying $\|\delta\|_2 \leq \frac{\sigma}{2}(\Phi^{-1}(\underline{p_A} - \zeta_x) - \Phi^{-1}(\overline{p_B} + \zeta_x))$

□

Theorem 3. *If* $\underline{p_A} - \zeta_x \geq \frac{1}{2}$, *then* $\sigma\Phi^{-1}(\underline{p_A} - \zeta_x) \leq \frac{\sigma}{2}(\Phi^{-1}(\underline{p_A} - \zeta_x) - \Phi^{-1}(\overline{p_B} + \zeta_x))$

Proof. Since $\underline{p_A} - \zeta_x \geq \frac{1}{2}$, $0 \leq \underline{p_A} \leq 1$ and $\zeta_x \geq 0$, we get $0 \leq \underline{p_A} - \zeta_x \leq 1$

And since $1 - \underline{p_A} \geq \overline{p_B}$, we get $\overline{p_B} + \zeta_x \leq \frac{1}{2}$, and thus, $0 \leq \overline{p_B} + \zeta_x \leq 1$

Since $\Phi^{-1}(1-t) = -\Phi^{-1}(t)$

$$\Phi^{-1}(\overline{p_B} + \zeta_x) = -\Phi^{-1}(1 - (\overline{p_B} + \zeta_x))$$
$$= -\Phi^{-1}((1 - \overline{p_B}) - \zeta_x)$$

Since $1 - \underline{p_A} \geq \overline{p_B}$

$$\leq -\Phi^{-1}(\underline{p_A} - \zeta_x)$$

Hence,

$$\Phi^{-1}(\underline{p_A} - \zeta_x) \leq -\Phi^{-1}(\overline{p_B} + \zeta_x)$$
$$\frac{\sigma}{2}\Phi^{-1}(\underline{p_A} - \zeta_x) \leq -\frac{\sigma}{2}\Phi^{-1}(\overline{p_B} + \zeta_x)$$

Adding $\frac{\sigma}{2}\Phi^{-1}(\underline{p_A} - \zeta_x)$ on both sides,

$$\sigma\Phi^{-1}(\underline{p_A} - \zeta_x) \leq \frac{\sigma}{2}(\Phi^{-1}(\underline{p_A} - \zeta_x) - \Phi^{-1}(\overline{p_B} + \zeta_x))$$

□

Theorem 4. *If* $\mathbb{P}_\epsilon(f(x+\epsilon) = f^p(x+\epsilon)) > 1 - \zeta_x$ *with confidence at least* $1 - \alpha_\zeta$. *If classifier* f *satisfies* $\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \geq \underline{p_A}$ *with confidence at least* $1 - \alpha$. *Then for classifier* f^p, $\mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A) \geq \underline{p_A} - \zeta_x$ *with confidence at least* $1 - (\alpha + \alpha_\zeta)$

Proof. Suppose f and f^p are classifiers such that for a fixed $x \in \mathbb{R}^m$, $\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \geq \underline{p_A}$ and $\mathbb{P}_\epsilon(f(x+\epsilon) = f^p(x+\epsilon)) > 1 - \zeta_x$. Note that this is true by the definition of $\underline{p_A}$, and is a separate $\underline{p_A}$ for each x. The statement is not true for all x with single $\underline{p_A}$
Let E_1 denote the event that $\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \geq \underline{p_A}$.
Let E_2 denote the event that $\mathbb{P}_\epsilon(f(x+\epsilon) = f^p(x+\epsilon)) > 1 - \zeta_x$.
By Theorem 2,

$$\mathbb{P}_\epsilon(f(x+\epsilon) = c_A) \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A) + \mathbb{P}_\epsilon(f(x+\epsilon) \neq f^p(x+\epsilon))$$
$$\underline{p_A} - \zeta_x \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A)$$

Let E_3 denote the event that $\underline{p_A} - \zeta_x \leq \mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A)$
Since, E_1 and E_2 imply E_3 i.e. $E_1 \cap E_2 \subseteq E_3$,

$$\mathbb{P}(E_3) \geq \mathbb{P}(E_1 \cap E_2)$$

By the additive rule of probability,

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cup E_2)$$
$$\mathbb{P}(E_3) \geq (1 - \alpha) + (1 - \alpha_\zeta) - 1$$
$$\mathbb{P}(E_3) \geq 1 - (\alpha + \alpha_\zeta)$$

Hence, for classifier f^p, $\mathbb{P}_\epsilon(f^p(x+\epsilon) = c_A) \geq \underline{p_A} - \zeta_x$ has confidence at least $1 - (\alpha + \alpha_\zeta)$ □

A.3 EVALUATION NETWORKS

Table 5 and Table 6 respectively present the standard top-1 accuracy of the original and approximated base classifiers and smoothed classifiers respectively.

Table 5: Standard top-1 accuracy for (non-smoothed) networks for combinations of approximations and σ's.

Dataset	Architecture	σ	original	Quantization			Prune		
				fp16	bf16	int8	5%	10%	20%
CIFAR10	ResNet-20	0.25	67.2	67.2	66.8	67.2	67.4	66.6	66.6
		0.5	56.8	56.8	57.2	56.8	57	57.4	58
		1.0	47.2	47.2	47.0	47.2	47	46.2	45.2
CIFAR10	ResNet-110	0.25	69.0	69.0	69.4	69.0	69.2	68.8	68.2
		0.5	59.4	59.4	59.4	59.4	59.6	59	58.8
		1.0	47.0	47.0	46.8	46.8	46.8	47.2	47
ImageNet	ResNet-50	0.5	24.2	24.2	24.4	24.2	24.2	24.4	24.2
		1.0	9.6	9.6	9.6	9.6	9.6	9.6	9.6
		2.0	6.4	6.4	6.4	6.4	6.4	6.4	6.4

Table 6: standard top-1 accuracy for smoothed networks for combinations of approximations and σ's.

Dataset	Architecture	σ	original	Quantization			Prune		
				fp16	bf16	int8	5%	10%	20%
CIFAR10	ResNet-20	0.25	77.2	77	77.2	77.2	77.6	77.2	77.6
		0.5	67.8	67.4	67.8	67.8	67.8	67.4	67.8
		1.0	55.6	55.6	55.6	55.8	54.8	55.2	55.6
CIFAR10	ResNet-110	0.25	76.6	76.4	76.2	76.4	76.2	76.2	76.4
		0.5	66.2	67	68	66.4	67	66.8	66.6
		1.0	55.6	55.4	56.2	56.2	55	55	54.8
ImageNet	ResNet-50	0.5	63.8	63.4	63.2	63.4	63.6	64	63
		1.0	48.8	48.6	48.8	48.6	48.8	48.6	47.8
		2.0	34.4	34.2	33.8	34.2	34.2	34.4	33.4

Table 8: ζ_x for approximate networks trained on different Gaussian augmentation σ's.

Dataset	Architecture	σ	Quantization			Prune		
			fp16	bf16	int8	5%	10%	20%
CIFAR10	ResNet-20	0.25	0.01	0.01	0.006	0.01	0.02	0.04
		0.5	0.006	0.008	0.01	0.01	0.02	0.03
		1.0	0.006	0.007	0.006	0.007	0.02	0.02
CIFAR10	ResNet-110	0.25	0.006	0.01	0.006	0.009	0.02	0.04
		0.5	0.006	0.006	0.006	0.008	0.02	0.03
		1.0	0.006	0.008	0.009	0.007	0.01	0.02
ImageNet	ResNet-50	0.5	0.006	0.009	0.006	0.01	0.02	0.09
		1.0	0.007	0.01	0.006	0.01	0.02	0.08
		2.0	0.006	0.01	0.006	0.007	0.02	0.07

A.4 ζ_x EVALUATION

We compute ζ_x value as the binomial confidence upper limit using (Clopper and Pearson, 1934) method with $n = 1000$ samples. For an experiment that adds Gaussian corruptions with σ to the input, we use the network that is trained with Gaussian data augmentation with variance σ^2.

A.5 SENSITIVITY TO CHANGING n

In Section 5, to save time due to a large number of approximations and DNNs tested, we used $n = 10^4$ samples for g's certification. Here, we present the effect of certifying with a larger n by comparing the ACR vs certification time on the IRS and baseline approaches for ResNet-20 on CIFAR10. On average, for larger n, we demonstrate greater speedup for larger σ. For instance, for int8 quantization with $\sigma = 1.0$, the speedup for certifying with $n = 10^6$ samples was 5.85x as compared to certification with $n = 10^4$ which yielded at 2.65x speedup. However, for smaller σ, certification with a larger n results in less speedup. For $\sigma = 0.25$, we observe speedups from 1.29x to 1.37x for $n = 10^4$ whereas from 0.93x to 1.15x for $n = 10^6$.

Table 7: Average IRS speedup for combinations of n, σ's, and quantizations for ResNet-20 on CIFAR10.

n	σ	Quantization		
		fp16	bf16	int8
10^4	0.25	1.37x	1.29x	1.3x
	0.5	1.79x	1.7x	1.77x
	1.0	2.85x	2.41x	2.65x
10^5	0.25	1.22x	1.11x	1.27x
	0.5	1.73x	1.4x	1.86x
	1.0	3.88x	2.40x	4.31x
10^6	0.25	1.12x	0.93x	1.15x
	0.5	1.97x	1.04x	2.25x
	1.0	4.58x	1.25x	5.85x

A.6 EVALUATION WITH LARGER n_p

The objective of IRS is to certify the approximated DNN with few samples. Thus, we consider n_p ranging from 1% to 10%. Nevertheless, we check IRS effectiveness for larger n_p values in this ablation study.

Since, IRS certifies radius $\sigma\Phi^{-1}(\underline{p_A} - \zeta_x)$ that is always smaller than original certified radius. When $n_p = n$, the baseline running from scratch should perform better than IRS, as it will reach a certification radius close to $\sigma\Phi^{-1}(\underline{p_A})$.

In this experiment, on CIFAR10 ResNet-20 with $\sigma = 1$, we let $n_p \in \{5\%, 10\% \dots 80\%\}$ of n. Figure 5 shows the ACR vs mean time plot for the baseline and IRS. We see that IRS gives speedup for $n_p = 70\%$. For $n_p = 75\%$ and $n_p = 80\%$, we see that baseline ACR is higher and IRS cannot achieve that ACR.

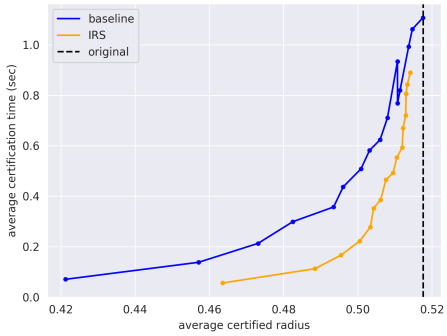
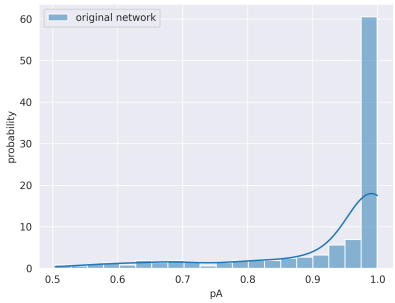
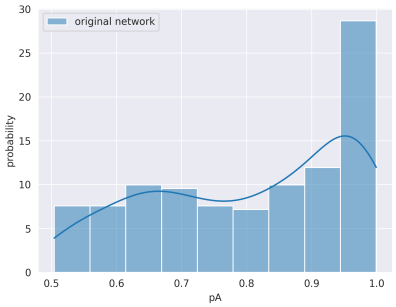


Figure 5: CIFAR10 ResNet-20 with $\sigma = 1$, for $n_p \in \{5\%, 10\% \dots 80\%\}$ of n

A.7 EFFECT OF STANDARD DEVIATION σ ON IRS SPEEDUP.

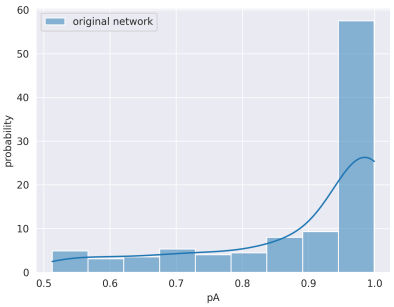


(a) ResNet-110 on CIFAR-10 ($\sigma = 0.25$)

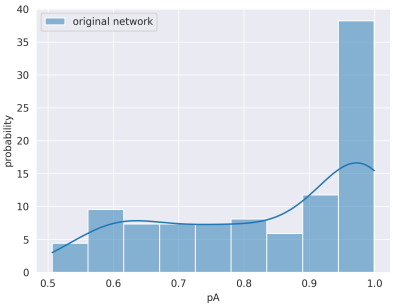


(b) ResNet-110 on CIFAR-10 ($\sigma = 1.0$)

Figure 6: Distribution of $\underline{p_A}$ values greater than 0.5 with different σ for ResNet-110 on CIFAR-10.



(a) ResNet-50 on ImageNet ($\sigma = 1.0$)



(b) ResNet-50 on ImageNet ($\sigma = 2.0$)

Figure 7: Distribution of $\underline{p_A}$ values greater than 0.5 with different σ for ResNet-50 on ImageNet.

Figure 6 and Figure 7, present the $\underline{p_A}$ distribution between 0.5 to 1, for ResNet-110 on CIFAR-10 and ResNet-50 on ImageNet respectively. The x-axis represents the range of $\underline{p_A}$ values and the y-axis represents their respective proportion. The results show that while certifying larger σ, on average the $\underline{p_A}$ values are smaller. As shown in Figure 7a, for $\sigma = 0.25$, less than 35% of $\underline{p_A}$ values are smaller than 0.95. On the other hand, in Figure 7b, when $\sigma = 1.0$, the distribution is less left-skewed as nearly 75% of $\underline{p_A}$ values are less than 0.95. When the σ is larger, the values of $\underline{p_A}$ tend to be farther away from 1. Therefore, the estimation of $\underline{p_A}$ is less precise in such cases, as observed in insight 2. As a result, non-incremental RS performs poorly compared to IRS in these situations, leading to a greater speedup with IRS.

A.8 THRESHOLD PARAMETER γ

Table 9 presents the proportion of cases for which $\underline{p_A} > \gamma$ for the γ chosen through hyperparameter search in Section 5.4 for different σ and networks.

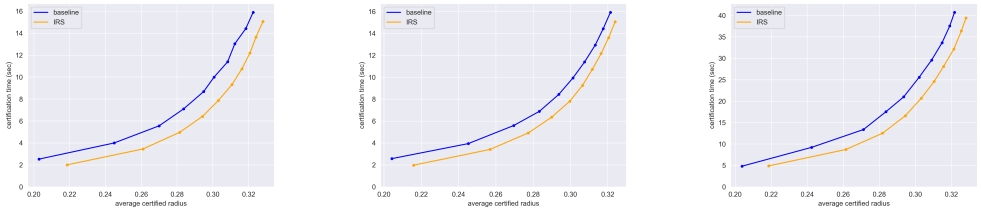
Table 9: Proportion of $\underline{p_A} > \gamma$ for different σ and networks.

Dataset	Architecture	γ	σ	$\underline{p_A} > \gamma$
CIFAR10	ResNet-20	0.99	0.25	0.346
			0.5	0.162
			1.0	0.034
CIFAR10	ResNet-110	0.99	0.25	0.362
			0.5	0.146
			1.0	0.034
ImageNet	ResNet-50	0.995	0.5	0.292
			1.0	0.14
			2.0	0.04

For CIFAR10 ResNet-20, we observe that $\underline{p_A} > \gamma = 0.346$ when $\sigma = 0.25$ and $\underline{p_A} > \gamma = 0.034$ when $\sigma = 1.0$. Additionally, for ImageNet ResNet-50, the results show $\underline{p_A} > \gamma = 0.292$ when $\sigma = 0.50$ and $\underline{p_A} > \gamma = 0.04$ when $\sigma = 2.0$. As shown in Section 5, certifying larger σ yields on average smaller $\underline{p_A}$. Expectedly, we see a smaller proportion of $\underline{p_A} > \gamma$ for larger σ and vice versa.

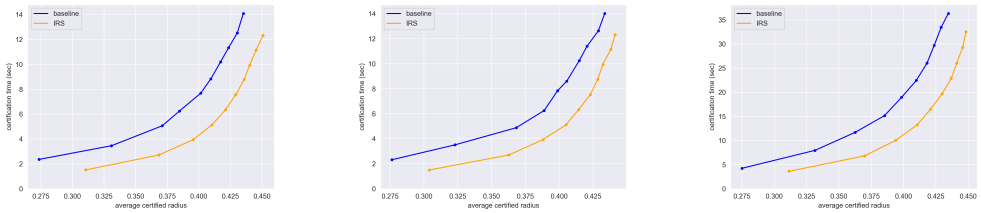
A.9 Quantization Plots

In this section, we present the ACR vs. time plots for all the quantization experiments. We use $n = 10^4$ for samples for certification of g. For certifying g^p, we consider n_p values from $\{1\%, \ldots 10\%\}$ of n. Note that these smaller values of n, n_p compared to Section 5.1 allow us to perform a large number of experiments.



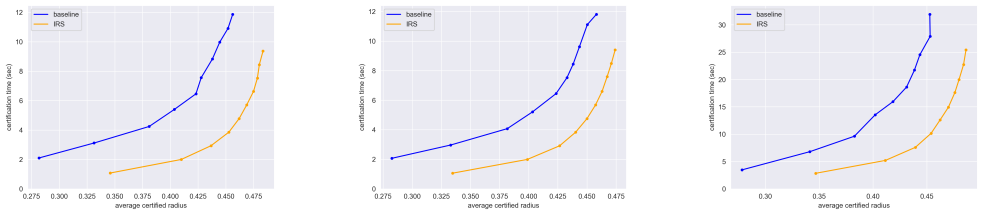
(a) fp16 (b) bf16 (c) int8

Figure 8: ResNet-20 on CIFAR10 with $\sigma = 0.25$.



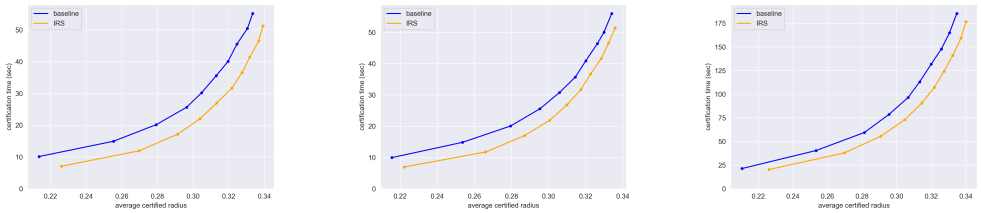
(a) fp16 (b) bf16 (c) int8

Figure 9: ResNet-20 on CIFAR10 with $\sigma = 0.5$.



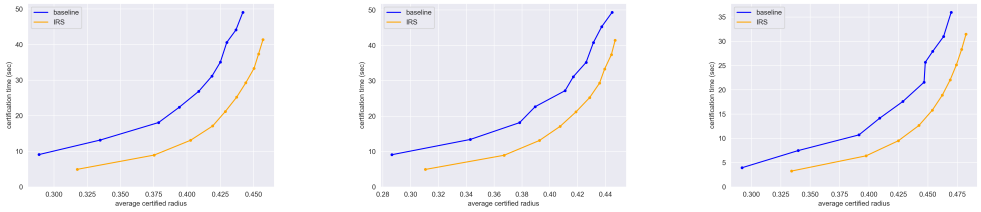
(a) fp16 (b) bf16 (c) int8

Figure 10: ResNet-20 on CIFAR10 with $\sigma = 1.0$.



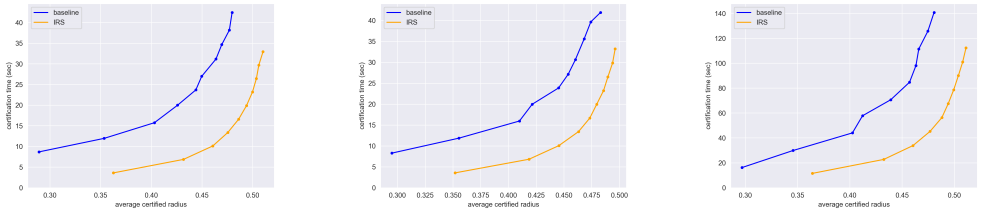
(a) fp16 (b) bf16 (c) int8

Figure 11: ResNet-110 on CIFAR10 with $\sigma = 0.25$.



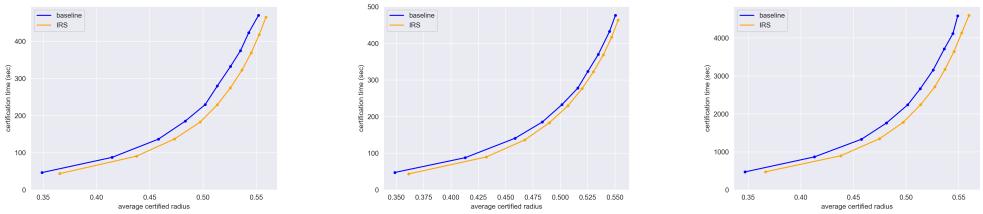
(a) fp16 (b) bf16 (c) int8

Figure 12: ResNet-110 on CIFAR10 with $\sigma = 0.5$.



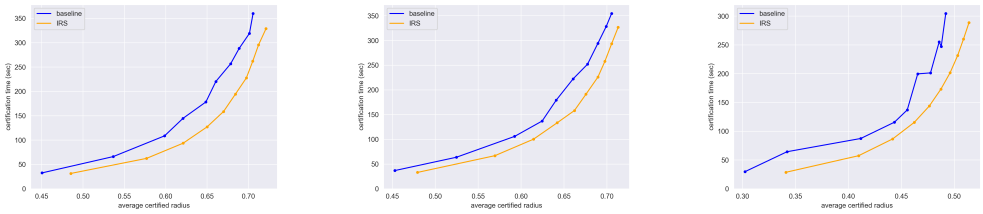
(a) fp16 (b) bf16 (c) int8

Figure 13: ResNet-110 on CIFAR10 with $\sigma = 1.0$.



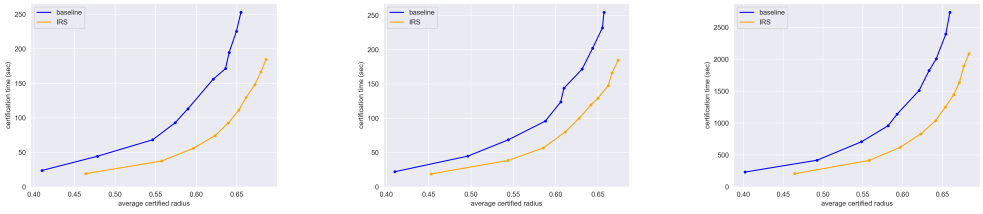
(a) fp16 (b) bf16 (c) int8

Figure 14: ResNet-50 on ImageNet with $\sigma = 0.5$.



(a) fp16 (b) bf16 (c) int8

Figure 15: ResNet-50 on ImageNet with $\sigma = 1.0$.



(a) fp16 (b) bf16 (c) int8

Figure 16: ResNet-50 on ImageNet with $\sigma = 2.0$.