Effects of Exponential Gaussian Distribution on (Double Sampling) Randomized Smoothing

Youwei Shu [1] **Xi Xiao** [1 †] **Derui Wang** [2] **Yuxin Cao** [1] **Siji Chen** [1] **Minhui Xue** [2] **Linyi Li** [3 4] **Bo Li** [3 5]

Abstract

Randomized Smoothing (RS) is currently a scalable certified defense method providing robustness certification against adversarial examples. Although significant progress has been achieved in providing defenses against ℓ_p adversaries, the interaction between the smoothing distribution and the robustness certification still remains vague. In this work, we comprehensively study the effect of two families of distributions, named Exponential Standard Gaussian (ESG) and Exponential General Gaussian (EGG) distributions, on Randomized Smoothing and Double Sampling Randomized Smoothing (DSRS). We derive an analytic formula for ESG's certified radius, which converges to the origin formula of RS as the dimension d increases. Additionally, we prove that EGG can provide tighter constant factors than DSRS in providing $\Omega(\sqrt{d})$ lower bounds of ℓ_2 certified radius, and thus further addresses the curse of dimensionality in RS. Our experiments on real-world datasets confirm our theoretical analysis of the ESG distributions, that they provide almost the same certification under different exponents η for both RS and DSRS. In addition, EGG brings a significant improvement to the DSRS certification, but the mechanism can be different when the classifier properties are different. Compared to the primitive DSRS, the increase in certified accuracy provided by EGG is prominent, up to 6.4% on ImageNet. Our code is available at https://github.com/tdano1/eg-on-smoothing.

[†]Corresponding Author [1]Shenzhen International Graduate School, Tsinghua University [2]CSIRO's Data61 [3]University of Illinois Urbana-Champaign [4]Simon Fraser University [5]University of Chicago. Correspondence to: Youwei Shu <shuyw21@mails.tsinghua.edu.com>, Xi Xiao <xiaox@sz.tsinghua.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

1. Introduction

Deep neural networks (DNNs) have achieved great success in various applications. However, DNNs are susceptible to adversarial perturbations in their inputs. To tackle the problem of adversarial attacks, a series of empirical defenses, such as adversarial training (Goodfellow et al., 2015; Kurakin et al., 2017; Madry et al., 2018), have been proposed. Nevertheless, this strategy quickly evolved into an arms race because no matter how robust the DNNs are, well-crafted adversarial examples are capable of bypassing the defenses (Carlini & Wagner, 2017; Athalye et al., 2018; Uesato et al., 2018). Recently, researchers proposed and developed certified defenses (Wong & Kolter, 2018; Wong et al., 2018; Raghunathan et al., 2018), a series of methodologies that can output the bounds of perturbed inputs, and provide provable robustness for classifiers. Aligned with these exact certified defense methods, randomized smoothing (RS) (Lécuyer et al., 2019; Li et al., 2019; Cohen et al., 2019) appears as a certifying tool based on probability, and gains in popularity since it can provide scalable robustness certifications for black-box functions. Cohen et al. (2019) first introduced the Neyman-Pearson (NP) lemma into the certification, which provided tight ℓ_2 certified radii for linear classifiers. Later, a series of attempts further extended the certification process of RS using functional optimization frameworks (Zhang et al., 2020; Dvijotham et al., 2020).

However, current research on the distributions for randomized smoothing is far from sufficient. After Zhang et al. (2020), Yang et al. (2020) and Li et al. (2022), this branch of study seems to be dormant. Intuitively, investigating the interrelationship of distribution and the RS method is not only beneficial for understanding and solving the limit of RS, but also liable to excavate the mathematical rules hidden deep in RS. In this work, we systematically study the ESG and EGG distributions in the RS framework, providing a detailed theoretical analysis for both the new distribution families. In addition, we conduct extensive experiments on real-world datasets, testify our theory on ESG, and complete the analysis for EGG. Overall, we conclude that the ESG distributions share almost the same certification for RS, and EGG's certification can be significantly improved using different exponents η, showing the better potential to

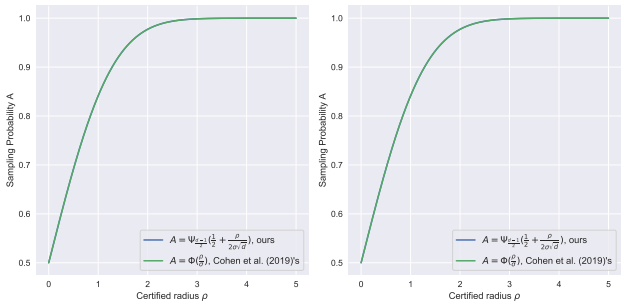


Figure 1. Our analytic formula for ESG highly approximates Cohen et al. (2019)'s at a sufficiently large dimension. (Both $\sigma = 1.0$. Left: $d = 3072$, Right: $d = 150224$.)

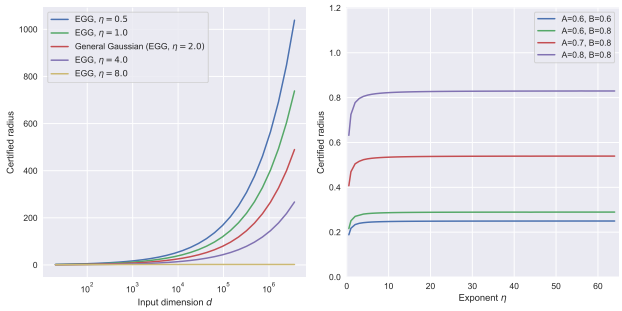


Figure 2. Numerical simulations for EGG in DSRS. Left: the concentration property holds ($B = 1$), smaller η provides tighter lower bounds. Right: the concentration property does not hold ($B < 1$), larger η provides better certified radius. For definitions of A and B, see Equation (5).

lessen the curse of dimensionality than the SOTA solution (Li et al., 2022).

The ESG distributions are extensions of Gaussian by generalizing the exponent η from 2 to $\mathbb{R}_+$. As shown by (Yang et al., 2020), the Gaussian distribution provides the SOTA certification for ℓ_2 certified radius in RS. In this work, we report that their views can be augmented because in addition to Gaussian, the family of ESG distributions can provide almost identical certification compared to Gaussian. Namely, we find that ESG can tie the SOTA distribution for RS in a high-dimensional setting. Concretely, we comprehensively analyze the computational method for the certified radius in ESG. By reversing the traditional calculation procedure and proposing two asymptotically mild assumptions, we figure out the simple analytic relation between sampling probability and certified radius, which is quite rare among all the distributions. Our theoretical analysis is perfectly enhanced by numerical simulations and experiments on real-world datasets. Moreover, we find the analytic formula for ESG converges to the results derived by Cohen et al. (2019), which reveals a mathematical connection between the beta distribution and the normal distribution.

Likewise, the EGG distributions are derivations of the General Gaussian distribution from the perspective of the exponent η. General Gaussian was introduced by Zhang et al. (2020), and exploited by DSRS (Li et al., 2022) to treat the curse of dimensionality in randomized smoothing. In this work, we further address the problem of the curse of dimensionality by tightening the lower bounds offered by DSRS via EGG. On the whole, the effect of EGG on RS certification largely depends on whether the base classifier satisfies a concentration property. In a more ideal case, where the concentration property is almost completely satisfied, EGG with a smaller η can provide tighter constant factors for the lower bounds provided by DSRS, which further alleviates the curse of dimensionality. But for more general cases, especially when real classifiers do not satisfy the concentration property well, EGG with a larger η gives better certified accuracy. To sum it up, despite different mechanisms, the introduction of EGG comprehensively improves robustness certifications provided by General Gaussian, both theoretically and practically.

Our main contributions include:

- For sufficiently large dimensions, we derive the analytic relation between certified radius and sampling probability for ESG distributions in the RS framework. The analytic formula obtained from ESG is in essence convergent to the formula derived by Cohen et al. (2019).
- Our theoretical analysis of EGG shows that the current solution to the curse of dimensionality in RS can be deepened: injecting some EGG into DSRS can tighten the constant factors of the lower bound of the certified radius.
- Extensive experiments on real datasets and classifiers verify our conclusion that the certification from ESG remains almost unchanged with η, and EGG provides comprehensively improved certifications compared to General Gaussian. On ImageNet, the increase in certified accuracy brought about by EGG can reach up to 6.4% compared to the baseline.

2. Related work

Randomized smoothing was first proposed as an extension for differential privacy, which provides a certified robustness bound for classifiers (Lécuyer et al., 2019). Subsequently, a series of improvements were made to obtain tighter ℓ_2 norm certificates through the Rényi divergence and the Neyman-Pearson Lemma (Li et al., 2019; Cohen et al., 2019). Furthermore, there are methods modeling the certification process

as functional optimization problems (Zhang et al., 2020; Dvijotham et al., 2020).

A line of work focuses on extending the robustness certification from ℓ_2-only to certifications against adversaries with other ℓ_p norms. ℓ_0 radius certification was made possible by constructing Neyman-Pearson sets for discrete random variables (Lee et al., 2019; Jia et al., 2020). Different from ℓ_2 certified regions, the ℓ_1 certified region is asymmetric in the space, which poses a new challenge to the certification algorithm. By perturbing input data under other noise distributions, such as Laplace and uniform distributions, the previous works have obtained ℓ_1 certificates (Teng et al.; Yang et al., 2020; Levine & Feizi, 2021). In addition to the asymmetry, some work discovered the phenomenon of the curse of dimensionality when trying to certify against ℓ_p adversaries whose $p > 2$ (Yang et al., 2020; Blum et al., 2020; Kumar et al., 2020). To deal with this issue, a recent work offered an $\Omega(1)$ bound w.r.t. the input dimension for ℓ_∞ certified radius by introducing a supplementary smoothing distribution (Li et al., 2022), which breaks the curse of dimensionality theoretically for the first time. Lately, a study found that the computation of certified radius can be improved by incorporating the geometric information from adjacent decision domains of the same class (Cullen et al., 2022).

There were also investigations using anisotropic or sample-specific smoothing noise to improve the certification (Eiras et al., 2022; Súkeník et al., 2022). In addition, a chain of work focused on improving the performance of base classifiers through adopting better training techniques (Salman et al., 2019; Zhai et al., 2020; Jeong & Shin, 2020; Jeong et al., 2021), or introducing denoising modules (Salman et al., 2020; Carlini et al., 2023; Wu et al., 2023). There were also attempts to adapt RS to broader application scenarios. RS was extended into defenses against adversarial patches (Levine & Feizi, 2020; Yatsura et al., 2023) and semantic perturbations (Li et al., 2021; Hao et al., 2022; Alfarra et al., 2022; Pautov et al., 2022). Moreover, RS has been shown to be capable of providing provable guarantees for tasks such as object detection, semantic segmentation, and watermarking (Chiang et al., 2020; Fischer et al., 2021; Bansal et al., 2022).

3. Preliminaries

Problem setup. We focus on the typical multi-class classification task in this work. Let $x_i \in \mathbb{R}^d$ be the i-th d-dimensional data point and $y_i \in \mathcal{Y} = \{1, 2, \cdots, N\}$ be its corresponding ground-truth label. We assume a dataset $\mathcal{J}$ contains data pairs $(x_i, y_i), i \in \mathbb{N}_{\leq n}$ that are i.i.d drawn from the sample space $\mathbb{R}^d \times \mathcal{Y}$. A N-way *base classifier* (neural networks in this work) $f : \mathbb{R}^d \to \mathcal{Y}$ can be trained to maximize the empirical classification accuracy $\frac{1}{|\mathcal{J}|}\sum_{(x,y)\in\mathcal{J}} \mathbb{1}_{f(x)=y}$ on dataset $\mathcal{J}$. Given an arbitrary data point x and its label y, it is known that in practice, most classifiers trained using standard training techniques are susceptible to adversarial perturbations within a small ϵ-ball.

Randomized smoothing. To mitigate adversarial perturbations, RS has been employed as a certified defense method that can provide a robustness guarantee on the correctness of the classification results from classifiers. It provides the robustness certification for the base classifier f by constructing its smoothed counterpart $\bar{f}$. Given a base classifier f, an input $x_0 \in \mathbb{R}^d$ and a smoothing distribution $\mathcal{P}$, the *smoothed classifier* is defined as follows:

$$\bar{f}_{\mathcal{P}}(x_0) = \operatorname*{argmax}_{a\in\mathcal{Y}} \mathbb{P}_{z\sim\mathcal{P}}\{f(x_0 + z) = a\}. \tag{1}$$

With the definition of $\bar{f}$, we can evaluate its ℓ_p certified robustness by ℓ_p certified radius defined below.

Definition 3.1. Given a base classifier $f : \mathbb{R}^d \to \mathcal{Y}$, its smoothed counterpart $\bar{f}_{\mathcal{P}} : \mathbb{R}^d \to \mathcal{Y}$ under a distribution $\mathcal{P}$ and a labeled example $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$. Then r is called ℓ_p **certified radius** of $\bar{f}_{\mathcal{P}}$ if

$$\forall x,\ \|x - x_0\|_p < r,\ \bar{f}_{\mathcal{P}}(x) = y_0. \tag{2}$$

To make the most conservative use of the information from smoothed classifiers (e.g., sampling probability A), we can consider a *true binary classifier*. Given a base classifier $f : \mathbb{R}^d \to \mathcal{Y}$, we call $\tilde{f}_{x_0} : \mathbb{R}^d \to \{0, 1\}$ a true binary classifier of f if for $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$ and random vector $z \in \mathbb{R}^d$:

$$\tilde{f}_{x_0}(z) = \mathbb{1}_{f(x_0+z)=y_0}. \tag{3}$$

In short, we need to construct the *worst* true binary classifier to compute the certified radius in RS. For example, Cohen et al. (2019) introduced the Neyman-Pearson lemma and successfully found such a worst true binary classifier. Given $\tilde{f}_{x_0}$ from the function space $\mathcal{F} = \{h(x) \mid h(x) \in [0, 1], \forall x \in \mathbb{R}^d\}$, their idea can be formulated as

$$\begin{aligned} \min_{\tilde{f}_{x_0}\in\mathcal{F}} \quad & \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(\delta + z)\right), \\ \text{s.t.} \quad & \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(z)\right) = A. \end{aligned} \tag{4}$$

Double sampling randomized smoothing. Essentially, the DSRS framework is based on a generalization of the Neyman-Pearson Lemma (Chernoff & Scheffe, 1952) that introduces one more subjection into the system. More specifically, DSRS provides a method to construct the worst true binary classifier based on the sampling probabilities of the base classifier under two different distributions. Like the Neyman-Pearson lemma, the problem of DSRS can be for-

Table 1. Properties and definitions of distributions.

Distribution	PDF	Notation	Formal Variance
Standard Gaussian	$\propto \exp\left(-\frac{r^2}{2\sigma^2}\right)$	$\mathcal{N}(\sigma)$	σ
Exponential Standard Gaussian	$\propto \exp\left(-\frac{r^\eta}{2\sigma_s^\eta}\right)$	$\mathcal{S}(\sigma,\eta)$	$\sigma_s = 2^{-\frac{1}{\eta}}\sqrt{\frac{d\Gamma(\frac{d}{\eta})}{\Gamma(\frac{d+2}{\eta})}}\sigma$
Truncated Exponential Standard Gaussian	$\propto \exp\left(-\frac{r^\eta}{2\sigma_s^\eta}\right)\mathbb{1}_{r\le T}$	$\mathcal{S}_t(\sigma,\eta,T)$	$\sigma_s = 2^{-\frac{1}{\eta}}\sqrt{\frac{d\Gamma(\frac{d}{\eta})}{\Gamma(\frac{d+2}{\eta})}}\sigma$
Exponential General Gaussian	$\propto r^{-2k}\exp\left(-\frac{r^\eta}{2\sigma_g^\eta}\right)$	$\mathcal{G}(\sigma,\eta,k)$	$\sigma_g = 2^{-\frac{1}{\eta}}\sqrt{\frac{d\Gamma(\frac{d-2k}{\eta})}{\Gamma(\frac{d-2k+2}{\eta})}}\sigma$
Truncated Exponential General Gaussian	$\propto r^{-2k}\exp\left(-\frac{r^\eta}{2\sigma_g^\eta}\right)\mathbb{1}_{r\le T}$	$\mathcal{G}_t(\sigma,\eta,k,T)$	$\sigma_g = 2^{-\frac{1}{\eta}}\sqrt{\frac{d\Gamma(\frac{d-2k}{\eta})}{\Gamma(\frac{d-2k+2}{\eta})}}\sigma$

mulated as a functional optimization problem as below:

$$\begin{aligned} \min_{\tilde{f}_{x_0}\in\mathcal{F}} \quad & \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(\delta+z)\right), \\ \text{s.t.} \quad & \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(z)\right) = A, \\ & \mathbb{E}_{z\sim\mathcal{Q}}\left(\tilde{f}_{x_0}(z)\right) = B. \end{aligned} \tag{5}$$

In the equations (5), $A, B \in [0,1]$ are probabilities that the base classifier f outputs the right label y_0 for example x_0 under noise distributions $\mathcal{P}$, $\mathcal{Q}$, respectively. Practically, they are usually estimated by the Monte Carlo sampling. For an example x_0 and a given combination of $\mathcal{P}$, $\mathcal{Q}$, A, B, we are able to derive a unique $\tilde{f}_{x_0}$. Finally, by finding the maximum $\|\delta\|_2$ that satisfies $\mathbb{P}_{z\sim\mathcal{P}}\{f(x_0+\delta+z)=y_0\} > 0.5$, we obtain the certified radius of x_0.

4. Exponential Gaussian distributions

In this section, we show the definition of ESG and EGG distributions used throughout the paper. In brief, ESG distributions are generalizations of the Gaussian distribution, which provides the SOTA performance in providing ℓ_2 certified radius for randomized smoothing (Cohen et al., 2019; Yang et al., 2020). Likewise, EGG distributions are generalizations of the General Gaussian distribution, which was introduced into the DSRS framework to provide a theoretical solution to the curse of dimensionality in randomized smoothing (Li et al., 2022). To the best of our knowledge, the EGG distributions belong to the Kotz-type distribution (Kotz, 1975), and we are the first to investigate their performance on randomized smoothing. In summary, our work systematically studies the interaction between multivariate distributions and (DS)RS through the lens of the exponent of distributions.

For conventional randomized smoothing tasks, only ESG or EGG distributions will be used. Additionally, under the DSRS framework, when using ESG as the smoothing distribution, the Truncated Exponential Standard Gaussian (TESG) distribution is employed as the supplementary distribution. Similarly, Truncated Exponential General Gaussian (TEGG) serves as the supplementary distribution when adopting EGG as the smoothing distribution. We let $\mathcal{S}(\sigma,\eta)$ and $\mathcal{G}(\sigma,\eta,k)$ be the probability density functions (PDFs) of ESG and EGG, respectively. Table 1 shows the definitions and basic properties of the distributions. In the table, $r, T, \sigma, \eta \in \mathbb{R}_+$ and $d \in \mathbb{N}_+$. $\Gamma(\cdot)$ is the gamma function. Following the settings in the previous studies (Yang et al., 2020; Li et al., 2022), we set the formal variance to ensure $\mathbb{E}r^2$ is a constant for all the smoothing distributions. We let σ_s and σ_g be the formal variances of EGG and ESG, respectively. The CDFs of the beta distribution $\mathrm{Beta}(\alpha,\alpha)$ and the gamma distribution $\Gamma(\alpha,1)$ are denoted respectively by $\Psi_\alpha(\cdot)$ and $\Lambda_\alpha(\cdot)$. We write $\phi_s(r)$ and $\phi_g(r)$ corresponding to the PDFs of $\mathcal{S}(\sigma,\eta)$ and $\mathcal{G}(\sigma,\eta,k)$ respectively. More details for the distributions are deferred to Appendix A.

5. ESG's certifications: hardly changes with the exponent η

In this section, we provide an analysis of ESG's certifications on randomized smoothing. Overall, for sufficiently large dimensions d, the certified radius offered by ESG has almost nothing to do with the exponent η, despite it being a significant hyperparameter for ESG. In the end, we derive a concise analytic formula that reveals the relationship between sampling probability and certified radius for the ESG distributions, which highly approximates the formula derived by (Cohen et al., 2019) for the Gaussian distribution.

To begin with, we consider the dual problem of Problem (28) for the ESG distributions. Let $\mathcal{P} = \mathcal{S}(\sigma,\eta)$, $p(x)$ be the PDF of ESG, and denote $\mathcal{V}$ for $\{z \mid p(z-\delta)+\nu p(z) < 0\}$ where $\nu \in \mathbb{R}$. Formally, to calculate the certified radius r, we have

$$r = \arg\max_{\|\delta\|_2} \max_{\nu\in\mathbb{R}} \quad \mathbb{P}_{z\sim\mathcal{P}+\delta}\{z\in\mathcal{V}\} \ge \frac{1}{2}, \tag{6a}$$

$$\text{s.t.} \quad \mathbb{P}_{z\sim\mathcal{P}}\{z\in\mathcal{V}\} = A. \tag{6b}$$

The proof of duality is omitted since it was provided in Zhang et al. (2020). Herein, we directly give the solution to this problem in the following theorem and defer the deriva-

tion to Appendix B.1.

Theorem 5.1. *For $\delta \in \mathbb{R}^d$ and $\rho \in \mathbb{R}_+$, letting $\delta = (\rho, 0, \cdots, 0)^T$, we have*

$$\mathbb{P}_{z\sim\mathcal{P}+\delta}\{z \in \mathcal{V}\} = \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\omega_\sharp(u,\nu), \tag{7a}$$

$$\mathbb{P}_{z\sim\mathcal{P}}\{z \in \mathcal{V}\} = \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\omega_\natural(u,\nu), \tag{7b}$$

where

$$\omega_\sharp(u,\nu) = \begin{cases} \Psi_{\frac{d-1}{2}}\left(\dfrac{2^{\frac{2}{\eta}}\sigma_s^2(u+\ln(-\nu))^{\frac{2}{\eta}} - (\sigma_s(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right), & u+\ln(-\nu) \geq 0, \\ 0, & u+\ln(-\nu) < 0. \end{cases} \tag{8a}$$

$$\omega_\natural(u,\nu) = \begin{cases} \Psi_{\frac{d-1}{2}}\left(\dfrac{(\rho+\sigma_s(2u)^{\frac{1}{\eta}})^2 - 2^{\frac{2}{\eta}}\sigma_s^2(u-\ln(-\nu))^{\frac{2}{\eta}}}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right), & u-\ln(-\nu) \geq 0, \\ 1, & u-\ln(-\nu) < 0. \end{cases} \tag{8b}$$

Conventionally, we obtain the certified radius r by known sampling probability A. The procedure can be outlined as a two-layer binary search, the outer one of which searches the value of $\|\delta\|_2$, and the inner one finds the maximum ν for the specified $\|\delta\|_2$. However, if we think about this problem in reverse, the relationship between certified radius ρ and sampling probability A can become very obvious. That is, we do not consider calculating the certified radius r from A, but get A from a given ρ. For ESG, we have the following theorem:

Theorem 5.2. *Let $\omega_\sharp(u,\nu)$ and $\omega_\natural(u,\nu)$ be defined as in Theorem 5.1. Then the sampling probability A can be calculated from the certified radius ρ by*

$$A = \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\omega_\natural(u,\nu), \tag{9a}$$

$$\text{s.t. } \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\omega_\sharp(u,\nu) = \frac{1}{2}. \tag{9b}$$

In fact, Theorem 5.2 is still not succinct enough to demonstrate the computation from ρ to A. To simplify it, we make two assumptions, both of which are very mild.

Assumption 5.3. For the ESG distribution $\mathcal{S}(\sigma,\eta)$ defined on $\mathbb{R}^d$, assume $d \gg \eta$ and $\sigma \in (0,1]$.

This assumption is reasonable in the high-dimensional machine learning setting. For instance, if we add Gaussian noises to ImageNet, we have $d = 150224, \eta = 2$, and σ is usually 0.5 or 1.0. Based on this, it is quite simple to approximate the formal variance σ_s. We see the following lemma:

Lemma 5.4. *Under Assumption 5.3, let d be the dimension, and σ_s be defined as in Table 1, and we have $\sigma_s = \Theta(d^{\frac{1}{2}-\frac{1}{\eta}})$.*

The proof is based on Stirling's formula and we leave it to Appendix B.2. With this property, talking $\sigma_s = (\frac{\eta}{2})^{\frac{1}{\eta}}\sigma d^{\frac{1}{2}-\frac{1}{\eta}}$ will only introduce infinitesimal errors. Next, another assumption we need is:

Assumption 5.5. Let ν be found by Equation (29b), and we assume $\frac{d}{\eta} \gg \ln(-\nu)$.

Appendix B.3 has more on the mildness of the assumption above. It helps us exploit the concentration property of the gamma distributions:

Lemma 5.6. *(Bilateral Concentration of the Gamma Distribution) Let $X \sim \Gamma(\frac{d}{\eta},1)$ be a random variable, where $\eta \in \mathbb{R}_+, d \in \mathbb{N}_+$. Let $\iota = \frac{\eta}{\epsilon^2 d}$, then for any $\iota \in (0,1)$, the following inequality holds:*

$$\mathbb{P}\{(1-\epsilon)\frac{d}{\eta} < X < (1+\epsilon)\frac{d}{\eta}\} \geq 1-\iota. \tag{10}$$

The proof is very similar to that of Lemma D.5. We still take Gaussian noises on ImageNet to illustrate Lemma 5.6. In fact, the Gaussian distribution will lead $u \sim \Gamma(\frac{d}{2},1)$ in Equation (9b), which satisfies $\mathbb{P}\{0.99 \cdot \frac{d}{2} < u < 1.01 \cdot \frac{d}{2}\} > 0.9935$. In other words, Assumption 5.5 essentially guarantees $\mathbb{P}\{u+\ln(-v) < 0\}$ and $\mathbb{P}\{u-\ln(-v) < 0\}$ are almost 0, thus we can omit these branches in Equations (8a) and (8b) when estimating the integral. WLOG, we let u be a constant value, and then let

$$\Psi_{\frac{d-1}{2}}\left(\frac{2^{\frac{2}{\eta}}\sigma_s^2(u+\ln(-\nu))^{\frac{2}{\eta}} - (\rho-\sigma_s(2u)^{\frac{1}{\eta}})^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right) = \frac{1}{2}. \tag{11}$$

Substituting $\sigma_s = (\frac{\eta}{2})^{\frac{1}{\eta}}\sigma d^{\frac{1}{2}-\frac{1}{\eta}}$ and $u = \frac{d}{\eta}$ into the equation above, we get

$$\frac{\sigma^2 d(1+\frac{\eta\ln(-\nu)}{d})^{\frac{2}{\eta}} - (\rho-\sigma\sqrt{d})^2}{4\rho\sigma\sqrt{d}} = \frac{1}{2}. \tag{12}$$

By Assumption 5.5, recalling the equivalent infinitesimal replacement $(1+x)^a \sim 1+ax$ when $x \to 0$, we see

$$\ln(-\nu) \approx \frac{\rho^2}{2\sigma^2}. \tag{13}$$

The solution for $\ln(-\nu)$ also verifies the Assumption 5.5 that $\frac{d}{\eta} \gg \ln(-\nu)$ is mild, considering practically we seldom see ℓ_2 certified radius $\rho > 5$ in RS. WLOG, injecting $\ln(-\nu) = \frac{\rho^2}{2\sigma^2}$ into Equation (9a), we finally obtain

$$A = \Psi_{\frac{d-1}{2}}(\frac{1}{2} + \frac{\rho}{2\sigma\sqrt{d}}). \tag{14}$$

In fact, this estimation is convergent to Cohen et al. (2019)'s formula when d is sufficiently large:

$$\Psi_{\frac{d-1}{2}}(\frac{1}{2} + \frac{\rho}{2\sigma\sqrt{d}}) \xrightarrow{d\to\infty} \Phi(\frac{\rho}{\sigma}), \tag{15}$$

whose differential case is shown by Ryder (2012). Interestingly, though the exponent η is a significant parameter for the derivation on ESG, this estimation (14) is irrelevant to η. Furthermore, our experiments on real-world datasets in Section 7 show that ESG's certifications are highly inert to η for both RS and DSRS, which corroborates our theoretical analysis, and indicates that there exist similar $A(\rho)$ relationships in DSRS. We show proof and more details for this approximation in Appendix B.4.

6. EGG's certifications: significantly improve with the exponent η

It is proved that taking the General Gaussian distribution as the smoothing distribution in DSRS provides an $\Omega(\sqrt{d})$ lower bound for the ℓ_2 certified radius (Li et al., 2022). This bound can be converted to an $\Omega(1)$ lower bound for the ℓ_∞ certified radius (Kumar et al., 2020), which breaks the curse of dimensionality against ℓ_∞ adversaries in high-dimensional settings. Nevertheless, though Li et al. (2022) proposed a theoretical solution to the curse of dimensionality for the first time, the study on the curse of dimensionality is still lacking. In this work, we investigate the further alleviation of the curse of dimensionality on the basis of DSRS, showing that theoretically, EGG can provide tighter lower bounds for ℓ_2 certified radius than the original DSRS. Primarily, we define a (σ, p, η)-concentration property to start our analysis:

Definition 6.1. ((σ, p, η)-Concentration Property) Let $f : \mathbb{R}^d \to \mathcal{Y}$ be an arbitrarily determined base classifier, $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$ be a labeled example. We say f satisfies (σ, p, η)-concentration assumption at (x_0, y_0) if for $p \in (0, 1)$ and T satisfying

$$\mathbb{P}_{z\sim\mathcal{S}(\sigma,\eta)}\{\|z\|_2 \le T\} = p, \tag{16}$$

f satisfies

$$\mathbb{P}_{z\sim\mathcal{S}(\sigma,\eta)}\{f(x_0 + z) = y_0 \mid \|z\|_2 \le T\} = 1. \tag{17}$$

The concentration property essentially defines a partly robust classifier that makes few mistakes on examples perturbed by limited noises. For instance, letting $\eta = 2$, if a classifier satisfies the $(\sigma, p, 2)$-concentration property on example (x_0, y_0), it will predict almost entirely correctly on perturbed examples with Gaussian noise z with $\|z\|_2 < T$. Despite the assumption being seemingly strict, it is satisfied well in the light of observations from Li et al. (2022): some robust classifiers provide almost perfect predictions for some examples from ImageNet when the ℓ_2 length $\|z\|_2 < \sigma\sqrt{d}$, but the accuracy intensely declines when the noise is stronger.

Herein, we first let the base classifier f satisfy the $(\sigma, p, 2)$-concentration property, the original assumption used in Li et al. (2022). Then we show a theorem that some EGG distributions with their corresponding truncated counterparts can certify the ℓ_2 radius with $\Omega(\sqrt{d})$ lower bounds by DSRS. In other words, the current solution to the curse of dimensionality provided by DSRS can be significantly augmented, that all EGG with $\eta \in (0, 2)$ have the potential to break the curse with better lower bounds.

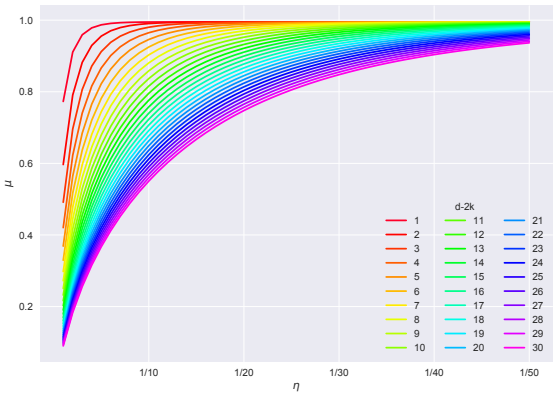


Figure 3. Tight factor μ grows as η shrinks, for most $d - 2k \in [1, 30] \cap \mathbb{N}$.

Theorem 6.2 (EGG with $\eta \in (0, 2)$ can certify $\Omega(\sqrt{d})$ lower bounds)**.** *Let $d \in \mathbb{N}_+$ be a sufficiently large input dimension, $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$ be a labeled example and $f : \mathbb{R}^d \to \mathcal{Y}$ be a base classifier that satisfies the $(\sigma, p, 2)$-concentration property w.r.t. (x_0, y_0). For the DSRS method, let $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ be the smoothing distribution to give a smoothed classifier $\bar{f}_{\mathcal{P}}$, and $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$ be the supplementary distribution with $T = \sigma\sqrt{2\Lambda^{-1}_{\frac{d}{2}}(p)}, d - 2k \in [1, 30] \cap \mathbb{N}$ and $\eta \in \{1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{50}\}$. Then for the smoothed classifier $\bar{f}_{\mathcal{P}}(x)$, the certified ℓ_2 radius satisfies*

$$r_{DSRS} \ge 0.02\sigma\sqrt{d}. \tag{18}$$

We briefly summarize the proof here and leave the details in Appendix D. In essence, Theorem 1 is generalizing Theorem 2 of Li et al. (2022) to EGG. Practically, it is intractable to derive an analytic solution for the certified radius r_{DSRS} for EGG. Therefore, the problem is simplified by introducing the concentration assumption, whereupon we solve Problem (5) only considering the truncated distribution $\mathcal{Q}$. The derivation is very similar to the solution of Problem (5), and the certification of the radius ρ for EGG is finally dependent on the discriminant

$$\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\Psi_{\frac{d-1}{2}}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right) \ge \frac{1}{2}. \tag{19}$$

With this formulation, whether a radius ρ can be certified can be directly judged since the LHS of Equation (19) is a

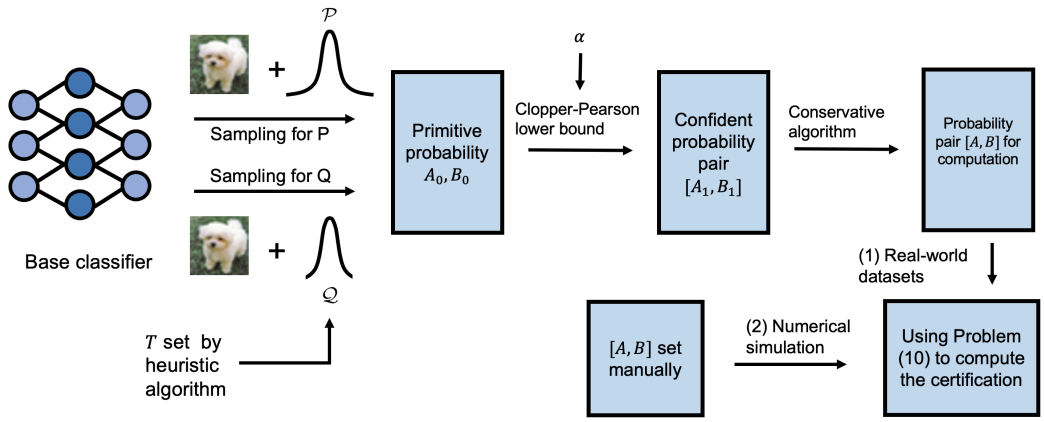


Figure 4. Illustration for experiments.

function of ρ. For Theorem 6.2, we just need to substitute $\rho = 0.02\sigma\sqrt{d}$ into Equation (19), and check whether the inequality holds. If true, then $\rho = 0.02\sigma\sqrt{d}$ is certified, and vice versa. For the detailed derivation, please see the proof of Lemma D.3 in the appendix.

Plus, we explore the effects of EGG under the (σ, p, η)-concentration assumption, where we can construct $\Omega(d^{1/\eta})$ lower bounds for the certified radius. However, though it is formally tighter than $\Omega(\sqrt{d})$, we find these two lower bounds are fundamentally equivalent. See Theorem D.7 in Appendix D.3.

Tighter constant factors of lower bounds. Besides certifying the lower bound radius $0.02\sigma\sqrt{d}$, the proof of Theorem 6.2 naturally contains an approach to determining tight constant factors for each EGG distribution (Algorithm 1). In fact, the value of the LHS of Equation (19) monotonically decreases with ρ, meaning that performing a simple binary search on ρ can provide the accurate certified radius. Therefore, we consider parameterizing the radius ρ into $\mu\sigma\sqrt{d}$, thus binary searching on the constant factor will derive the tight constant factor μ for the $\Omega(\sqrt{d})$ lower bound (Algorithm 1). We report computational results for μ in Figure 3, where for values of $d - 2k$ except 1, the tight constant factor μ increases monotonically as the η decreases. Essentially, these results demonstrate that the solution to the curse of dimensionality provided by Li et al. (2022) (with General Gaussian, namely $\eta = 2$ in EGG) can be improved by choosing smaller $\eta \in (0, 2)$, for most $d - 2k \in [1, 30] \cap \mathbb{N}$.

Overall, the theoretical analysis shows that when the base classifier satisfies the concentration property, EGG with $\eta \in (0, 2)$ brings significant enhancement for the lower bound of the certified radius offered by DSRS (Li et al., 2022). Obviously, it is hard for the realistic model to be perfectly *concentrated*, thus EGG takes different effects under real classifiers. However, what remains unchanged is that EGG also comprehensively improves the certification of DSRS. See Section 7 for details.

7. Experiments

In this section, we report the effects of ESG and EGG distributions on the certified radius. Figure 4 sketches the outline of our experiments. Here, we only focused on real-world datasets and leave details for numerical simulation experiments in Appendix F. For the case of real-world datasets, A and B in Problem (57) are initially from Monte Carlo sampling results on a given base classifier, then reduced to respective Clopper-Pearson confidence intervals (Clopper & Pearson, 1934), and finally determined by a conservative algorithm (Li et al., 2022). The procedure for real-world datasets can also be seen in Algorithm 2 in Appendix E.1.

Experimental setups. All base classifiers used in this work are trained by CIFAR-10 (Krizhevsky et al., 2009) or ImageNet (Russakovsky et al., 2015), taking EGG with $\eta = 2$ as the noise distribution.

We choose NP certification as the baseline since it is the state-of-the-art method for single distribution certification. The sampling distribution for the NP method is $\mathcal{P} = \mathcal{S}(\sigma, \eta)$ for ESG, and $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ for EGG. For fairness, the sampling number N is set to 100000, with the significance level $\alpha = 0.001$.

In the double-sampling process, we set $k = 1530$ and $k = 75260$ for CIFAR-10 and ImageNet, respectively, in consistent with base classifiers. The threshold parameter T for $\mathcal{Q}$ is determined by a heuristic algorithm from Li et al. (2022). Specifically, we set

$$T_{\mathcal{S}} = \sigma_s (2\Lambda_{\frac{d}{\eta}}^{-1}(\kappa))^{\frac{1}{\eta}} = \sigma\sqrt{\frac{d\Gamma(\frac{d}{\eta})}{\Gamma(\frac{d+2}{\eta})}}(\Lambda_{\frac{d}{\eta}}^{-1}(\kappa))^{\frac{1}{\eta}},$$

$$T_{\mathcal{G}} = \sigma_g (2\Lambda_{\frac{d-2k}{\eta}}^{-1}(\kappa))^{\frac{1}{\eta}} = \sigma\sqrt{\frac{d\Gamma(\frac{d-2k}{\eta})}{\Gamma(\frac{d-2k+2}{\eta})}}(\Lambda_{\frac{d-2k}{\eta}}^{-1}(\kappa))^{\frac{1}{\eta}}, \tag{20}$$

where κ is determined by the heuristic algorithm (Li et al.,

Table 2. Certified radius at r for standardly augmented models, certified by ESG under DSRS

Dataset	Method	Certified accuracy at r													
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
CIFAR10	ESG, $\eta = 1.0$	57.6%	42.6%	31.3%	21.5%	15.8%	12.8%	8.6%	6.8%	4.3%	2.3%	1.3%	0.8%	0.3%	0.1%
	ESG, $\eta = 2.0$ (Gaussian)	57.6%	42.6%	31.6%	21.5%	15.8%	12.7%	8.8%	6.8%	4.5%	2.4%	1.3%	0.7%	0.2%	0.2%
	ESG, $\eta = 4.0$	57.6%	42.6%	31.3%	21.5%	15.9%	12.9%	8.6%	6.9%	4.3%	2.4%	1.3%	0.8%	0.2%	0.1%
	ESG, $\eta = 8.0$	57.8%	42.6%	31.6%	21.6%	15.9%	12.9%	8.9%	6.7%	4.2%	2.4%	1.3%	0.9%	0.2%	0.1%
ImageNet	ESG, $\eta = 1.0$	59.6%	51.5%	43.2%	37.9%	33.0%	26.8%	23.1%	21.5%	19.9%	17.4%	13.8%	11.5%	10.3%	7.7%
	ESG, $\eta = 2.0$, (Gaussian)	59.6%	51.6%	43.1%	38.0%	32.9%	26.9%	23.1%	21.5%	19.7%	17.4%	13.6%	11.4%	10.1%	8.3%
	ESG, $\eta = 4.0$	59.6%	51.5%	43.2%	38.0%	32.9%	27.2%	23.1%	21.6%	19.9%	17.2%	13.6%	11.4%	10.2%	8.0%
	ESG, $\eta = 8.0$	59.6%	51.5%	43.2%	38.0%	33.0%	26.8%	23.1%	21.6%	19.7%	17.3%	13.6%	11.5%	10.1%	8.4%

Table 3. Certified radius at r for standardly augmented models, certified by EGG under DSRS

Dataset	Method	Certified accuracy at r													
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
CIFAR10	EGG, $\eta = 0.25$	54.2%	37.6%	23.5%	16.5%	9.4%	4.5%	0.5%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	EGG, $\eta = 0.5$	55.5%	40.4%	25.2%	19.1%	13.4%	8.5%	5.5%	2.0%	0.4%	0.1%	0.0%	0.0%	0.0%	0.0%
	EGG, $\eta = 1.0$	56.3%	41.7%	28.2%	20.0%	15.1%	10.5%	7.1%	4.2%	1.9%	0.9%	0.1%	0.0%	0.0%	0.0%
	DSRS (Li et al., 2022) (EGG, $\eta = 2.0$)	56.7%	42.4%	29.3%	20.2%	15.7%	11.5%	8.0%	5.5%	2.6%	1.5%	0.6%	0.1%	0.0%	0.0%
	EGG, $\eta = 4.0$	57.5%	42.5%	30.0%	20.2%	15.9%	12.2%	8.5%	6.5%	3.4%	1.8%	0.9%	0.4%	0.0%	0.0%
	Ours (EGG, $\eta = 8.0$)	**57.6%**	**42.5%**	**30.9%**	**20.6%**	**15.8%**	**12.3%**	**8.6%**	**6.6%**	**3.7%**	**2.1%**	**1.1%**	**0.5%**	**0.2%**	0.0%
ImageNet	EGG, $\eta = 0.25$	53.8%	41.4%	28.4%	20.1%	7.1%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	EGG, $\eta = 0.5$	54.9%	46.3%	36.4%	26.3%	22.1%	15.2%	8.7%	3.1%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%
	EGG, $\eta = 1.0$	57.0%	47.8%	39.9%	32.8%	24.9%	22.0%	18.5%	13.1%	9.2%	5.0%	2.1%	0.5%	0.0%	0.0%
	DSRS (Li et al., 2022) (EGG, $\eta = 2.0$)	58.4%	48.5%	41.5%	35.2%	28.9%	23.3%	21.3%	18.8%	14.1%	11.1%	8.9%	6.1%	2.2%	1.4%
	EGG, $\eta = 4.0$	58.7%	49.9%	42.6%	36.4%	31.0%	23.9%	22.3%	20.2%	17.3%	13.2%	10.7%	9.2%	6.8%	4.0%
	Ours (EGG, $\eta = 8.0$)	**59.1%**	**50.8%**	**42.9%**	**36.8%**	**31.8%**	**24.6%**	**22.6%**	**20.7%**	**18.9%**	**14.5%**	**11.7%**	**10.1%**	**8.6%**	**5.2%**

2022), a simple function of Monte Carlo sampling probability from $\mathcal{P}$. The sampling numbers N_1, N_2 are 50000, and the significance levels α_1, α_2 are 0.0005 for Monte Carlo sampling, equal for $\mathcal{P}$ and $\mathcal{Q}$. The error bound e for certified radius is set at 1×10^{-6}.

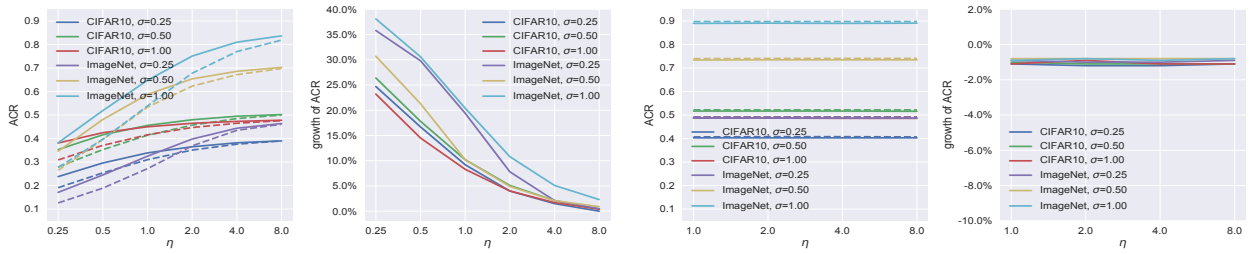
The settings of the exponent η for ESG and EGG are slightly different. We choose $\eta \in \{1.0, 2.0, 4.0, 8.0\}$ as the exponent of ESG distributions. For EGG, we take $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ and $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$. To display the increasing trends more clearly, we choose $\eta \in \{0.25, 0.5, 1.0, 2.0, 4.0, 8.0\}$ for them. For the convenience of comparison, we fix the base classifier for each group of experiments, and all the sampling distributions keep $\mathbb{E}r^2$ the same with the base classifier following the setting of the previous work (Yang et al., 2020; Li et al., 2022).

Evaluation metrics. We consider the ℓ_2 certified radius in all experiments. To evaluate the certified robustness of smoothed classifiers, we take *certified accuracy* $\triangleq CA(r, \mathcal{J})$ at radius r on test dataset $\mathcal{J}$ as the basic metric (Cohen et al., 2019; Zhang et al., 2020; Li et al., 2022). For (x_i, y_i) in $\mathcal{J}$, if the certified radius computed by a certification method (e.g. NP, DSRS) is r_i, and the output for x_i through the smoothed classifier is y_i, we say (x_i, y_i) is certified accurate at radius r_i for the smoothed classifier. On this basis, $CA(r, \mathcal{J})$ is the ratio of examples in $\mathcal{J}$ whose certified radius $r_i \geq r$. Defined on the certified accuracy, *Average Certified Radius* (ACR) (Zhai et al., 2020; Jeong & Shin, 2020) is another main metric that we use to show results of different distributions for real-world datasets. Formally, we have $ACR \triangleq \int_{r \geq 0} CA(r, \mathcal{J}) \cdot \mathrm{d}r$.

Integral methods. The `scipy` package loses precision when calculating integrals for the $\Gamma(a, 1)$ distribution with large parameters (say, $a > 500$) on infinite intervals. To solve this problem, we implement a Linear Numerical Integration (LNI) method to compute the expectations fast and accurately based on Lemma 5.6. With this great property, we can compute the integral for the gamma distribution by only considering $1 - \iota$ total mass. Though primitive, we find the LNI method that uniformly segments the integration interval provides good precision for certifications on CIFAR-10 and ImageNet. For all the ESG experiments, we set the number of segments to 256, and $\iota = 10^{-4}$. We illustrate the effect of the segment number on CIFAR-10 and ImageNet in Figure 6. For EGG distributions, we inherit the integral method from Li et al. (2022), where we compute the integrals on the interval $(0, +\infty)$ by the `scipy` package. The computational methods for ESG and EGG are left to Appendix C.

Experimental results. We only discuss the effect of ESG and EGG in DSRS, since their influence on NP is similar. Table 2 and Table 3 report the maximum certified accuracy among base classifiers with $\sigma \in \{0.25, 0.50, 1.00\}$, which is widely adopted by previous work to show experimental results of certified robustness. To show the universality of our results, we provide experimental results on Consistency (Jeong & Shin, 2020) and SmoothMix (Jeong et al., 2021) models in Appendix H.1. We leave the detailed results on NP and DSRS in Appendix H.2.

The exponent of ESG has little impact on the certification. As shown in Table 2, on both CIFAR10 and ImageNet, the certification provided by ESG distributions is highly insensitive to the alternation of η, which echos our theoretical



(a) (b) (c) (d)

Figure 5. ACR results on real-world datasets. **(a).** ACR monotonically increases with η in EGG. **(b).** The ACR growth gain from DSRS relative to NP shrinks with η in EGG. **(c).** ACR stays almost constant in ESG. **(d).** The ACR growth gain from DSRS remains almost constant in ESG. For (a) and (c), solid lines represent results from DSRS, and dotted lines represent results from NP.

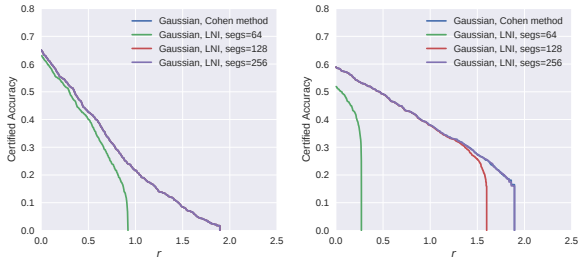


Figure 6. Results for the ℓ_2 certified radius under different numbers of segment settings of LNI, both base clasifiers are standardly augmented by General Gaussian with $\sigma = 0.50$. Left: For CIFAR-10, the curves show segs ≥ 128 is enough for CIFAR-10. Right: on ImageNet. The curves of Cohen's formula and segs = 256 are almost overlapped.

analysis in Section 5. Currently, the mainstream view believes Gaussian is the best distribution to provide the ℓ_2 certified radius for RS, and our results show many ESG can provide the best as well.

The larger the exponent in EGG, the better the certification on real classifiers. Table 3 reveals the phenomenon that certified accuracy at r increases with the η of EGG. On both CIFAR10 and ImageNet, our strategy to use EGG with a larger η (8.0 in the tables) performs obviously better than General Gaussian (EGG with $\eta = 2.0$) used in DSRS (Li et al., 2022). This is different from our theoretical analysis in 6, because real classifiers do not perfectly satisfy the concentration property. Our Figure 10b has more details on this.

Better sampling probability + better probability utilization $\Rightarrow$ better certification. The certified accuracy provided by (DS)RS is decided by two steps: sampling and certifying. For ESG, the stable results under changing η indicate different ESG noise gives almost the same prediction accuracy on classifiers. Moreover, our results show that the larger η in EGG offers both higher sampling accuracy and a better ability to use the probability.

The improvement brought by EGG is finite. For EGG, though the certification improves with η, Figure 5b reveals the growth brought by DSRS relative to NP shrinks with η, despite the fact that certified accuracy provided by DSRS keeps increasing. Furthermore, we see there is likely to be an upper bound for the DSRS certification: Gaussian's certification (in this work, $\eta = 2$ for ESG). Additionally, the right of Figure 5 also implies the convergence. Therefore, chances are high that the growth of EGG's certification is not endless, and it is hard for EGG to challenge ESG's position as the optimal distribution in (DS)RS.

8. Conclusion

We report in this paper that the exponent η in ESG is almost unable to affect the certification in RS and DSRS, from both theoretical and experimental perspectives. We derive the analytic sampling probability-certified radius relation under high-dimensional assumptions, creatively bridging the beta distribution and the normal distribution, and broadening the current optimal distribution in RS from Gaussian to many of ESG. In addition, we find EGG distributions provide significant amelioration on the current solution to the curse of dimensionality on RS, and they can improve the ℓ_2 certified radii of smoothed classifiers on real datasets. The working mechanism of EGG can be quite different, depending on whether the classifier satisfies a concentration property.

Impact Statement

This paper is dedicated to advancing the certified robustness of machine learning, and we do not think any of its societal impacts must be highlighted here.

Acknowledgements

This work is supported in part by the Overseas Research Cooperation Fund of Tsinghua Shenzhen International Graduate School (HW2021013).

References

Alfarra, M., Bibi, A., Khan, N., Torr, P. H. S., and Ghanem, B. Deformrs: Certifying input deformations with randomized smoothing. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pp. 6001–6009. AAAI Press, 2022.

Athalye, A., Carlini, N., and Wagner, D. A. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples. In Dy, J. G. and Krause, A. (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 274–283. PMLR, 2018.

Bansal, A., Chiang, P., Curry, M. J., Jain, R., Wigington, C., Manjunatha, V., Dickerson, J. P., and Goldstein, T. Certified neural network watermarks with randomized smoothing. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvári, C., Niu, G., and Sabato, S. (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 1450–1465. PMLR, 2022.

Blum, A., Dick, T., Manoj, N., and Zhang, H. Random smoothing might be unable to certify $l\infty$ robustness for high-dimensional images. *J. Mach. Learn. Res.*, 21:211:1–211:21, 2020.

Carlini, N. and Wagner, D. Adversarial examples are not easily detected: Bypassing ten detection methods. In *Proceedings of the 10th ACM workshop on artificial intelligence and security*, pp. 3–14, 2017.

Carlini, N., Tramèr, F., Dvijotham, K. D., Rice, L., Sun, M., and Kolter, J. Z. (certified!!) adversarial robustness for free! In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023.

Chernoff, H. and Scheffe, H. A generalization of the neyman-pearson fundamental lemma. *The Annals of Mathematical Statistics*, pp. 213–225, 1952.

Chiang, P., Curry, M. J., Abdelkader, A., Kumar, A., Dickerson, J., and Goldstein, T. Detection as regression: Certified object detection with median smoothing. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Clopper, C. J. and Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934.

Cohen, J., Rosenfeld, E., and Kolter, J. Z. Certified adversarial robustness via randomized smoothing. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pp. 1310–1320. PMLR, 2019.

Cullen, A. C., Montague, P., Liu, S., Erfani, S. M., and Rubinstein, B. I. P. Double bubble, toil and trouble: Enhancing certified robustness through transitivity. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A. (eds.), *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022.

Dvijotham, K. D., Hayes, J., Balle, B., Kolter, J. Z., Qin, C., György, A., Xiao, K., Gowal, S., and Kohli, P. A framework for robustness certification of smoothed classifiers using f-divergences. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020.

Eiras, F., Alfarra, M., Torr, P. H. S., Kumar, M. P., Dokania, P. K., Ghanem, B., and Bibi, A. ANCER: anisotropic certification via sample-wise volume maximization. *Trans. Mach. Learn. Res.*, 2022, 2022.

Fischer, M., Baader, M., and Vechev, M. T. Scalable certified segmentation via randomized smoothing. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3340–3351. PMLR, 2021.

Goodfellow, I. J., Shlens, J., and Szegedy, C. Explaining and harnessing adversarial examples. In Bengio, Y. and LeCun, Y. (eds.), *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

Hao, Z., Ying, C., Dong, Y., Su, H., Song, J., and Zhu, J. Gsmooth: Certified robustness against semantic transformations via generalized randomized smoothing. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvári, C., Niu,

G., and Sabato, S. (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8465–8483. PMLR, 2022.

Jeong, J. and Shin, J. Consistency regularization for certified robustness of smoothed classifiers. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Jeong, J., Park, S., Kim, M., Lee, H., Kim, D., and Shin, J. Smoothmix: Training confidence-calibrated smoothed classifiers for certified robustness. In Ranzato, M., Beygelzimer, A., Dauphin, Y. N., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 30153–30168, 2021.

Jia, J., Wang, B., Cao, X., and Gong, N. Z. Certified robustness of community detection against adversarial structural perturbation via randomized smoothing. In Huang, Y., King, I., Liu, T., and van Steen, M. (eds.), *WWW '20: The Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pp. 2718–2724. ACM / IW3C2, 2020.

Kotz, S. Multivariate distributions at a cross road. In *A Modern Course on Statistical Distributions in Scientific Work: Volume 1—Models and Structures Proceedings of the NATO Advanced Study Institute held at the University of Calgagry, Calgary, Alberta, Canada July 29–August 10, 1974*, pp. 247–270. Springer, 1975.

Krizhevsky, A., Hinton, G., et al. Learning multiple layers of features from tiny images. 2009.

Kumar, A., Levine, A., Goldstein, T., and Feizi, S. Curse of dimensionality on randomized smoothing for certifiable robustness. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5458–5467. PMLR, 2020.

Kurakin, A., Goodfellow, I. J., and Bengio, S. Adversarial machine learning at scale. In *International Conference on Learning Representations*, 2017.

Lécuyer, M., Atlidakis, V., Geambasu, R., Hsu, D., and Jana, S. Certified robustness to adversarial examples with differential privacy. In *2019 IEEE Symposium on Security and Privacy, SP 2019, San Francisco, CA, USA, May 19-23, 2019*, pp. 656–672. IEEE, 2019.

Lee, G., Yuan, Y., Chang, S., and Jaakkola, T. S. Tight certificates of adversarial robustness for randomly smoothed classifiers. In Wallach, H. M., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E. B., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 4911–4922, 2019.

Levine, A. and Feizi, S. (de)randomized smoothing for certifiable defense against patch attacks. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Levine, A. and Feizi, S. Improved, deterministic smoothing for l_1 certified robustness. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 6254–6264. PMLR, 2021.

Li, B., Chen, C., Wang, W., and Carin, L. Certified adversarial robustness with additive noise. In Wallach, H. M., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E. B., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 9459–9469, 2019.

Li, L., Weber, M., Xu, X., Rimanic, L., Kailkhura, B., Xie, T., Zhang, C., and Li, B. TSS: transformation-specific smoothing for robustness certification. In Kim, Y., Kim, J., Vigna, G., and Shi, E. (eds.), *CCS '21: 2021 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, Republic of Korea, November 15 - 19, 2021*, pp. 535–557. ACM, 2021.

Li, L., Zhang, J., Xie, T., and Li, B. Double sampling randomized smoothing. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvári, C., Niu, G., and Sabato, S. (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 13163–13208. PMLR, 2022.

Madry, A., Makelov, A., Schmidt, L., Tsipras, D., and Vladu, A. Towards deep learning models resistant to adversarial attacks. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.

Pautov, M., Tursynbek, N., Munkhoeva, M., Muravev, N., Petiushko, A., and Oseledets, I. V. CC-CERT: A probabilistic approach to certify general robustness of neural networks. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pp. 7975–7983. AAAI Press, 2022.

Raghunathan, A., Steinhardt, J., and Liang, P. Semidefinite relaxations for certifying robustness to adversarial examples. In Bengio, S., Wallach, H. M., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pp. 10900–10910, 2018.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115: 211–252, 2015.

Ryder, R. Theorem: The beta(b, b) distribution converges to the normal distribution when $b \to \infty$, 2012. URL https://www.math.wm.edu/~leemis/chart/UDR/PDFs/BetaNormal.pdf.

Salman, H., Li, J., Razenshteyn, I. P., Zhang, P., Zhang, H., Bubeck, S., and Yang, G. Provably robust deep learning via adversarially trained smoothed classifiers. In Wallach, H. M., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E. B., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 11289–11300, 2019.

Salman, H., Sun, M., Yang, G., Kapoor, A., and Kolter, J. Z. Denoised smoothing: A provable defense for pretrained classifiers. In *Advances in Neural Information Processing Systems*, volume 33, pp. 21945–21957. Curran Associates, Inc., 2020.

Súkeník, P., Kuvshinov, A., and Günnemann, S. Intriguing properties of input-dependent randomized smoothing. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvári, C., Niu, G., and Sabato, S. (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 20697–20743. PMLR, 2022.

Teng, J., Lee, G.-H., and Yuan, Y. ℓ_1 adversarial robustness certificates: a randomized smoothing approach.

Uesato, J., O'Donoghue, B., Kohli, P., and van den Oord, A. Adversarial risk and the dangers of evaluating against weak attacks. In Dy, J. G. and Krause, A. (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 5032–5041. PMLR, 2018.

Wong, E. and Kolter, J. Z. Provable defenses against adversarial examples via the convex outer adversarial polytope. In Dy, J. G. and Krause, A. (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 5283–5292. PMLR, 2018.

Wong, E., Schmidt, F. R., Metzen, J. H., and Kolter, J. Z. Scaling provable adversarial defenses. In Bengio, S., Wallach, H. M., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pp. 8410–8419, 2018.

Wu, Q., Ye, H., Gu, Y., Zhang, H., Wang, L., and He, D. Denoising masked autoencoders help robust classification. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023.

Yang, G., Duan, T., Hu, J. E., Salman, H., Razenshteyn, I. P., and Li, J. Randomized smoothing of all shapes and sizes. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 10693–10705. PMLR, 2020.

Yatsura, M., Sakmann, K., Hua, N. G., Hein, M., and Metzen, J. H. Certified defences against adversarial patch attacks on semantic segmentation. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023.

Zhai, R., Dan, C., He, D., Zhang, H., Gong, B., Ravikumar, P., Hsieh, C., and Wang, L. MACER: attack-free and scalable robust training via maximizing certified radius. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020.

Zhang, D., Ye, M., Gong, C., Zhu, Z., and Liu, Q. Black-box certification with randomized smoothing: A functional optimization based framework. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

Appendix

A. Supplementary for definitions of distributions

A.1. Derivation for PDFs

Let $\phi_s(r), \phi_s(r,T)$ be PDFs of $\mathcal{S}(\sigma,\eta)$ and $\mathcal{S}_t(\sigma,\eta,T)$ respectively. Given $\mathcal{S}(\sigma,\eta) \propto \exp\left(-\frac{r^\eta}{2\sigma_s^\eta}\right)$, we have

$$\phi_s(r) = \frac{\eta}{2}\frac{1}{(2\sigma_s^\eta)^{\frac{d}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d}{\eta})}\exp(-\frac{1}{2}(\frac{r}{\sigma_s})^\eta), \tag{21a}$$

$$\phi_s^{-1}(r) = \sigma_s\left(-2\ln\left(\frac{\Gamma(\frac{d}{\eta})}{\Gamma(\frac{d}{2})}2^{\frac{d}{\eta}+1}\sigma_s^d\pi^{\frac{d}{2}}\frac{r}{\eta}\right)\right)^{\frac{1}{\eta}}, \tag{21b}$$

$$\phi_s(r,T) = \frac{\eta}{2}\frac{1}{(2\sigma_s^\eta)^{\frac{d}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\gamma(\frac{d}{\eta},\frac{T^\eta}{2\sigma_s^\eta})}\exp(-\frac{1}{2}(\frac{r}{\sigma_s})^\eta), \tag{21c}$$

and the ratio constant

$$C_s = \frac{\phi_s(r,T)}{\phi_s(r)} = \frac{\Gamma(\frac{d}{\eta})}{\gamma(\frac{d}{\eta},\frac{T^\eta}{2\sigma_s^\eta})}. \tag{22}$$

Similarly, since $\phi_g(r) \propto r^{-2k}\exp\left(-\frac{r^\eta}{2\sigma_g^\eta}\right)$, we have

$$\phi_g(r) = \frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-2k}{\eta})}r^{-2k}\exp(-\frac{1}{2}(\frac{r}{\sigma_g})^\eta), \tag{23a}$$

$$\phi_g^{-1}(r) = \sigma_g\left(\frac{4k}{\eta}W\left(\frac{\eta}{2k}\left(\frac{\Gamma(\frac{d-2k}{\eta})}{\Gamma(\frac{d}{2})}2^{\frac{d}{\eta}+1}\sigma_g^d\pi^{\frac{d}{2}}\frac{r}{\eta}\right)^{-\frac{\eta}{2k}}\right)\right)^{\frac{1}{\eta}}. \tag{23b}$$

In the equations above, $W(\cdot)$ is the principal branch of the Lambert W function. Let $\phi_g(r,T)$ be PDF of $\mathcal{G}_t(\sigma,\eta,k,T)$, then

$$\phi_g(r,T) = \frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\gamma(\frac{d-2k}{\eta},\frac{T^\eta}{2\sigma_g^\eta})}r^{-2k}\exp(-\frac{1}{2}(\frac{r}{\sigma_g})^\eta), \tag{24}$$

where $\gamma(\cdot)$ is the lower incomplete gamma function. We see

$$C_g = \frac{\phi_g(r,T)}{\phi_g(r)} = \frac{\Gamma(\frac{d-2k}{\eta})}{\gamma(\frac{d-2k}{\eta},\frac{T^\eta}{2\sigma_g^\eta})}. \tag{25}$$

A.2. Visualization for distributions

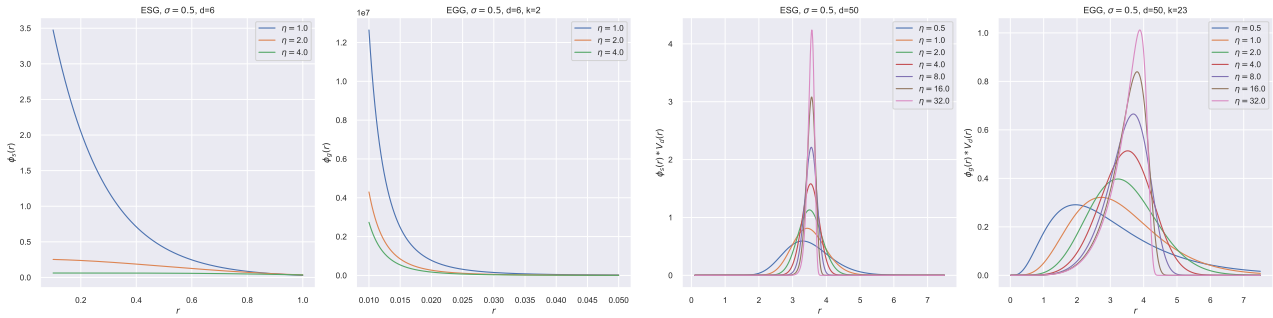
We show the distributions for r in the space by setting

$$y = \phi_s(r)\cdot V_d(r) = \frac{\eta}{2}\frac{1}{(2\sigma_s^\eta)^{\frac{d}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d}{\eta})}\exp(-\frac{1}{2}(\frac{r}{\sigma_s})^\eta)\cdot\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}r^{d-1} \tag{26}$$

for ESG, and

$$y = \phi_g(r)\cdot V_d(r) = \frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-2k}{\eta})}r^{-2k}\exp(-\frac{1}{2}(\frac{r}{\sigma_g})^\eta)\cdot\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}r^{d-1} \tag{27}$$

for EGG. See the definition of $V_d(r)$ in Lemma B.1. We provide a simple demonstration for ESG and EGG in Figure 7. Overall, under high-dimensional settings, the difference between $\phi(r)$ of exponential Gaussian distributions can be huge, we thus suggest using $\phi(r)\cdot V_d(r)$ to assist in observing the properties of the distributions. As shown in Figure 7b, ESG tends to concentrate on a *thin shell* in the space, while the r^{-2k} term in EGG alleviates that property.



(a) Visualization for $\phi(r)$. (b) Visualization for $\phi(r) \cdot V_d(r)$.

Figure 7. Demonstration for ESG and EGG.

B. Supplementary for the theoretical analysis of ESG

B.1. Proof of Theorem 5.1

The ESG distributions inherit the ℓ_2-symmetry of the Gaussian distribution in high-dimensional space, thus the certified radius can be derived by the level-set method (Yang et al., 2020; Li et al., 2022). To begin with, we consider the problem for the Neyman-Pearson certification. Consider example $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$. Let $\tilde{f}_{x_0} : \mathbb{R}^d \to \{0, 1\}$ be a true binary classifier of base classifier $f : \mathbb{R}^d \to \mathcal{Y}$ with respect to input x_0:

$$\begin{aligned} \min_{\tilde{f}_{x_0} \in \mathcal{F}} \quad & \mathbb{E}_{z \sim \mathcal{S}_{(\sigma,\eta)}} \left(\tilde{f}_{x_0}(\delta + z) \right), \\ \text{s.t.} \quad & \mathbb{E}_{z \sim \mathcal{S}_{(\sigma,\eta)}} \left(\tilde{f}_{x_0}(z) \right) = A. \end{aligned} \tag{28}$$

We usually solve the problem above by transforming it into its dual problem:

$$\max_{\nu \in \mathbb{R}} \quad \mathbb{P}_{z \sim \mathcal{S}_{(\sigma,\eta)} + \delta} \{ p(z - \delta) + \nu p(z) < 0 \}, \tag{29a}$$

$$\text{s.t.} \quad \mathbb{P}_{z \sim \mathcal{S}_{(\sigma,\eta)}} \{ p(z - \delta) + \nu p(z) < 0 \} = A. \tag{29b}$$

Unlike the DSRS case, the invariant set in Neyman-pearson certification is $\{z \mid p(z - \delta) + \nu p(z) < 0\}$, and we denote it as $\mathcal{V}$. We have the following lemma to calculate the volume of $d - 1$ dimensional hypersphere.

Lemma B.1. *(Equation (16) in Li et al. (2022), Volume of Hypersphere, restated) Let $r \in \mathbb{R}_+, d \in \mathbb{N}_+$. The volume $V_d(r)$ of $d - 1$ dimensional hypersphere with radius r is*

$$V_d(r) = \frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2} + 1)} r^{d-1}. \tag{30}$$

For all integrals in this paper, WLOG we let d be an even number.

For Equation (29b), we have

$$\begin{aligned} & \mathbb{P}_{z \sim \mathcal{S}_{(\sigma,\eta)}} \{ z \in \mathcal{V} \} \\ = & \int_0^\infty \phi_s(r) V_d(r) dr \cdot \mathbb{P}\{ x \in \mathcal{V} \mid \|x\|_2 = r \} \\ = & \frac{1}{\Gamma(\frac{d}{\eta})} \int_0^\infty u^{\frac{d}{\eta} - 1} \exp(-u) du \cdot \mathbb{P}\{ x \in \mathcal{V} \mid \|x\|_2 = \sigma_s (2u)^{\frac{1}{\eta}} \} \\ = & \mathbb{E}_{u \sim \Gamma(\frac{d}{\eta}, 1)} \mathbb{P}\{ x \in \mathcal{V} \mid \|x\|_2 = \sigma_s (2u)^{\frac{1}{\eta}} \} \end{aligned} \tag{31}$$

Next, we consider $\omega_\natural(u, \nu) = \mathbb{P}\{ p(x - \delta) + \nu p(x) < 0 \mid \|x\|_2 = \sigma_s (2u)^{\frac{1}{\eta}} \}$. Obviously, $\phi_s(r)$ is a monotonically decreasing function with respect to r. Since all PDFs of $\mathcal{S}(\sigma, \eta)$ are bounded functions, we set the upper bound to be U for convenience.

Let $r = 0$, we have

$$U = \frac{\eta}{2} \frac{1}{(2\sigma_s^{\eta})^{\frac{d}{\eta}} \pi^{\frac{d}{2}}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d}{\eta})} \tag{32}$$

for $\phi_s(r)$. Namely, $\forall x \in [0, \infty),\ p(x) \in (0, U]$, where $U < +\infty$ is a constant for any determined distribution. As a result, if for a specific x, we have $-\nu p(x) > U$, the probability $\mathbb{P}\{p(x-\delta) + \nu p(x) < 0 \mid \|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}\}$ will always be 1. Next, we suppose $-\nu p(x) \in (0, U]$, otherwise $-\nu p(x)$ is outside the domain of $\phi_s^{-1}(x)$. When $\|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}$, we have

$$\begin{aligned} & p(x-\delta) + \nu p(x) < 0 \\ \Longleftrightarrow & \phi_s(\|x-\delta\|_2) \leq -\nu \phi_s(\|x\|_2) \\ \Longleftrightarrow & \|x-\delta\|_2 \geq \phi_s^{-1}(-\nu U \exp(-u)). \end{aligned} \tag{33}$$

Then we are ready to solve

$$0 < -\nu U \exp(-u) \leq U, \tag{34}$$

where $u \geq 0$. Since ν is always negative, the left side $0 < -\nu U \exp(-u)$ always holds. For the right side, we notice

$$-\nu U \exp(-u) \leq U \Longleftrightarrow \exp(-u) \leq -\frac{1}{\nu} \Longleftrightarrow u - \ln(-\nu) \geq 0. \tag{35}$$

Now we know for $\nu < 0$,

$$u - \ln(-\nu) \geq 0 \Longleftrightarrow -\nu p(x) \in (0, U], \tag{36}$$

and

$$u - \ln(-\nu) < 0 \Longleftrightarrow -\nu p(x) \in (U, \infty], \tag{37}$$

which means $\mathbb{P}\{p(x-\delta) + \nu p(x) < 0 \mid \|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}\} = 1$ when $u - \ln(-\nu) < 0$. Therefore,

$$\omega_{\natural}(u, \nu) = \begin{cases} \Psi_{\frac{d-1}{2}} \left(\dfrac{(\rho + \sigma_s(2u)^{\frac{1}{\eta}})^2 - \left(\phi_s^{-1}(-\nu \phi_s(\sigma_s(2u)^{\frac{1}{\eta}}))\right)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}} \right) & , u - \ln(-\nu) \geq 0, \\ 1 & , u - \ln(-\nu) < 0. \end{cases} \tag{38}$$

For Equation (29a), we have

$$\begin{aligned} & \mathbb{P}_{z \sim \mathcal{S}_{(\sigma,\eta)}}\{z \in \mathcal{V} + \delta\} \\ = & \int_0^{\infty} \phi_s(r) V_d(r) dr \cdot \mathbb{P}\{x \in \mathcal{V} + \delta \mid \|x\|_2 = r\} \\ = & \frac{1}{\Gamma(\frac{d}{\eta})} \int_0^{\infty} u^{\frac{d}{\eta}-1} \exp(-u) du \cdot \mathbb{P}\{x \in \mathcal{V} + \delta \mid \|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}\} \\ = & \mathbb{E}_{u \sim \Gamma(\frac{d}{\eta}, 1)} \mathbb{P}\{x \in \mathcal{V} + \delta \mid \|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}\} \end{aligned} \tag{39}$$

We write $\omega_{\sharp}$ for $\mathbb{P}\{p(x) + \nu p(x+\delta) < 0 \mid \|x\|_2 = \sigma_g(2u)^{\frac{1}{\eta}}\}$. When $-\frac{1}{\nu} p(x) \in (0, U]$, we have

$$\begin{aligned} & p(x) + \nu p(x+\delta) < 0 \\ \Longleftrightarrow & p(x) \leq -\nu p(x+\delta) \\ \Longleftrightarrow & \|x+\delta\|_2 \leq \phi_s^{-1}(-\frac{1}{\nu} \phi_s(\|x\|_2)). \end{aligned} \tag{40}$$

Injecting $\|x\|_2 = \sigma_s(2u)^{\frac{1}{\eta}}$ into inequalities above, we get

$$-\frac{1}{\nu} p(x) \in (0, U] \Longleftrightarrow u + \ln(-\nu) \geq 0, \tag{41}$$

which is the boundary condition for $\omega_{\sharp}$. Though looks similar, it differs significantly from $\omega_{\natural}$. Let $-\frac{1}{\nu}\phi_s(\|x\|_2) \in (U, +\infty]$, then

$$p(x) + \nu p(x+\delta) < 0 \Longleftrightarrow \phi_s(\|x+\delta\|_2) \geq U, \tag{42}$$

which means $\mathbb{P}\{p(x)+\nu p(x+\delta)<0 \mid \|x\|_2=\sigma_g(2u)^{\frac{1}{\eta}}\}=0$ under $u+\ln(-\nu)<0$. Finally, we have

$$\omega_\sharp(u,\nu)=\begin{cases}\Psi_{\frac{d-1}{2}}\left(\dfrac{\phi_s^{-1}(-\frac{1}{\nu}\phi_s(\sigma_s(2u)^{\frac{1}{\eta}}))^2\}-(\sigma_s(2u)^{\frac{1}{\eta}}-\rho)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right) & ,u+\ln(-\nu)\geq 0,\\ 0 & ,u+\ln(-\nu)<0.\end{cases} \tag{43}$$

Calculation of $\phi_s^{-1}(r)$. Appearing in both $\omega_\natural$ and $\omega_\sharp$ functions above, $\phi_s^{-1}(r)$ is an indispensable value in the system. Here we consider the case of $\omega_\natural$. Let ξ be $\phi_s^{-1}(-\nu\phi_s(\sigma_s(2u)^{\frac{1}{\eta}})$ when $-\nu\phi_s(\sigma_s(2u)^{\frac{1}{\eta}}\in(0,U]$, then

$$\phi_s(\xi)=-\nu\phi_s(\sigma_s(2u)^{\frac{1}{\eta}}). \tag{44}$$

We thus have

$$\begin{aligned}&\frac{\eta}{2}\frac{1}{(2\sigma_s^\eta)^{\frac{d}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d}{\eta})}\exp(-\frac{\xi^\eta}{2\sigma_s^\eta})=-\frac{\nu\eta}{2}\frac{1}{(2\sigma_s^\eta)^{\frac{d}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d}{\eta})}\exp(-u)\\ \iff&\exp(-\frac{1}{2}(\frac{\xi}{\sigma_s})^\eta)=-\nu\exp(-u).\end{aligned} \tag{45}$$

Obviously $\xi\geq 0$, then $\exp(-\frac{1}{2}(\frac{\xi}{\sigma_s})^\eta)\in(0,1]$. When Equation (45) has a solution, we need

$$0<-\nu\exp(-u)\leq 1. \tag{46}$$

Therefore, we get the boundary condition for $\nu<0$:

$$u\geq\ln(-\nu). \tag{47}$$

Under Equation (47), Equation (44) can be solved:

$$\xi=2^{\frac{1}{\eta}}\sigma_s(u-\ln(-\nu))^{\frac{1}{\eta}}. \tag{48}$$

Substituting Equation (48) into Equation (38) gives Equation (8b), and the case of $\omega_\sharp$ can be derived similarly.

B.2. Mildness of Assumption 5.3

We first provide the proof for Lemma 5.4.

Proof. By Stirling's approximation, we see

$$\begin{aligned}\lim_{d\to\infty}\frac{\sigma_s}{d^{\frac{1}{2}-\frac{1}{\eta}}}&=2^{-\frac{1}{\eta}}\sqrt{\frac{d^{\frac{2}{\eta}}\Gamma(\frac{d}{\eta})}{\Gamma(\frac{d+2}{\eta})}\sigma}\\ &=2^{-\frac{1}{\eta}}\sqrt{d^{\frac{2}{\eta}}\sqrt{\frac{d-\eta}{d+2-\eta}}\frac{(\frac{d-\eta}{e\eta})^{\frac{d-\eta}{\eta}}}{(\frac{d+2-\eta}{e\eta})^{\frac{d+2-\eta}{\eta}}}}\\ &=(\frac{\eta}{2})^{\frac{1}{\eta}}\sigma,\end{aligned} \tag{49}$$

which by definition means $\sigma_s=\Theta(d^{\frac{1}{2}-\frac{1}{\eta}})$. □

Lemma 5.4 provides a simple but very accurate approximation for computing integrals in Theorem 5.1. Moreover, when $d\gg\eta$ and σ is small, the constant factor above is rather accurate. Let absolute error be $|\sigma_s-(\frac{\eta}{2})^{\frac{1}{\eta}}\sigma d^{\frac{1}{2}-\frac{1}{\eta}}|$, and relative error be $\frac{|\sigma_s-(\frac{\eta}{2})^{\frac{1}{\eta}}\sigma d^{\frac{1}{2}-\frac{1}{\eta}}|}{\sigma_s}$. Table 4 and Table 5 show the errors for the approximation. From the tables, we observe that under common settings of variance and dimension for randomized smoothing, the approximation shows great accuracy. On CIFAR10, the maximum absolute error is 2.95×10^{-3}, and the maximum relative error is 4.88×10^{-4}. Furthermore, the errors are even lower on ImageNet, as we see that the maximum absolute error is 2.59×10^{-4} and the maximum relative error is no more than 9.99×10^{-6}. It is reasonable that a larger dimension ensures less error for approximation since there are only two sources for the error: one is Stirling's approximation, which demands large d to be precise, and another is that d should be far greater than η, since larger d benefits more since we fix $\eta\in\{1.0,2.0,4.0,8.0\}$ in our work.

Table 4. Errors for the approximation of σ_s, $d = 3072$

σ	$\eta = 0.5$		$\eta = 1.0$		$\eta = 2.0$		$\eta = 4.0$		$\eta = 8.0$	
	AE	RE	AE	RE	AE	RE	AE	RE	AE	RE
0.12	2.15e-11	4.88e-04	1.76e-07	1.63e-04	9.88e-14	8.23e-13	8.65e-05	8.14e-05	3.54e-04	1.22e-04
0.25	4.48e-11	4.88e-04	3.67e-07	1.63e-04	2.06e-13	8.23e-13	1.80e-04	8.14e-05	7.37e-04	1.22e-04
0.50	8.96e-11	4.88e-04	7.34e-07	1.63e-04	4.12e-13	8.23e-13	3.60e-04	8.14e-05	1.47e-03	1.22e-04
1.00	1.79e-10	4.88e-04	1.47e-06	1.63e-04	8.23e-13	8.23e-13	7.21e-04	8.14e-05	2.95e-03	1.22e-04

Table 5. Errors for the approximation of σ_s, $d = 150224$

σ	$\eta = 0.5$		$\eta = 1.0$		$\eta = 2.0$		$\eta = 4.0$		$\eta = 8.0$	
	AE	RE	AE	RE	AE	RE	AE	RE	AE	RE
0.12	1.29e-15	9.99e-06	5.15e-10	3.33e-06	4.46e-12	3.72e-11	4.68e-06	1.66e-06	3.11e-05	2.50e-06
0.25	2.68e-15	9.99e-06	1.07e-09	3.33e-06	9.30e-12	3.72e-11	9.74e-06	1.66e-06	6.48e-05	2.50e-06
0.50	5.36e-15	9.99e-06	2.15e-09	3.33e-06	1.86e-11	3.72e-11	1.95e-05	1.66e-06	1.30e-04	2.50e-06
1.00	1.07e-14	9.99e-06	4.29e-09	3.33e-06	3.72e-11	3.72e-11	3.90e-05	1.66e-06	2.59e-04	2.50e-06

B.3. Mildness of Assumption 5.5

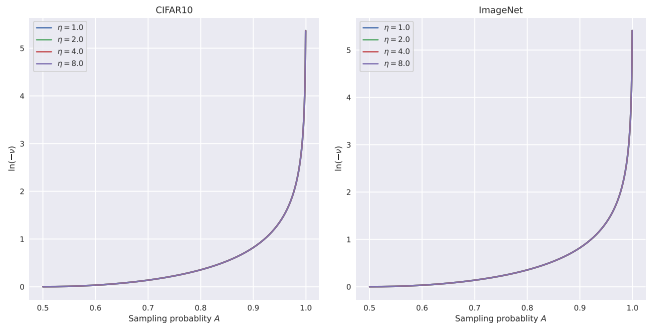


Figure 8. Under commonly used settings for RS, $\frac{d}{\eta} \gg \ln(-\nu)$ almost always holds.

In addition to Equation (13), we show Assumption 5.5 is very mild for practical use in RS computation. In Figure 8, all the $\ln(-\nu)$ are computed by the conventional procedure by binary search without any approximation. For each subfigure, we uniformly select 1000 points in $(0.5, 1)$ to simulate the sampling probability A. The maximum A we show in the figure is 0.9995, which corresponds to $\ln(-\nu) \approx 5.4$ for the listed ESG distributions. Intuitively, the value of $\ln(-\nu)$ will increase exponentially when $A \to 1^-$, but it is more than enough for the practical use of the RS framework (the confident lower bound of A can hardly ever reach 1), especially for ImageNet.

B.4. Extensions for Equation (14)

Equation (14) is an accurate approximation for Cohen et al. (2019)'s formula not only for little ρ that appears in the RS certification. For this, we consider

$$A = \Psi_{\frac{d-1}{2}}(\frac{1}{2} + \frac{x}{2\sqrt{d}}), \tag{50}$$

where $x \in (0, \sqrt{d})$. Here we provide a simple proof for the convergence in Equation (14).

Theorem B.2. *(Convergence of the Beta(d, d) distribution's CDF) Let $\Psi_d(x)$ be the CDF of the beta(d, d) distribution and $\Phi(x)$ be the CDF of the standard normal distribution, then for $x \in (-\sqrt{2d}, \sqrt{2d})$, $\lim_{d\to\infty} \Psi_d(\frac{1}{2} + \frac{x}{2\sqrt{2d}}) = \Phi(x)$, with an $O(1/\sqrt{d})$ error bound .*

Proof. Let $\psi_d(x)$ be the PDF of the $Beta(d, d)$ distribution and $\phi(x)$ be the PDF of the standard normal distribution. By Ryder (2012), we have

$$\psi_d(x) = \phi(x)(1 + O(\frac{1}{d})). \tag{51}$$

WLOG, we consider appropriate real numbers $c_1, c_2 \in \mathbb{R}_+$, such that

$$(1 - \frac{c_1}{d})\phi(x) < \psi_d(x) < (1 + \frac{c_2}{d})\phi(x). \tag{52}$$

Then we get

$$\lim_{d\to+\infty}\int_{-\sqrt{2d}}^{x}(1-\frac{c_1}{d})\phi(t)dt < \lim_{d\to+\infty}\int_{-\sqrt{2d}}^{x}\psi_d(t)dt < \lim_{d\to+\infty}\int_{-\sqrt{2d}}^{x}(1+\frac{c_2}{d})\phi(t)dt. \tag{53}$$

By $|\phi(x)| < \frac{1}{\sqrt{2\pi}}$, we have

$$\lim_{d\to+\infty}(\int_{-\sqrt{2d}}^{x}\phi(t)dt - \frac{(x+\sqrt{2d})c_1}{\sqrt{2\pi}d}) < \lim_{d\to+\infty}\int_{-\sqrt{2d}}^{x}\psi_d(t)dt < \lim_{d\to+\infty}(\int_{-\sqrt{2d}}^{x}\phi(t)dt + \frac{(x+\sqrt{2d})c_2}{\sqrt{2\pi}d}). \tag{54}$$

Thus for $x \in (-\sqrt{2d}, \sqrt{2d})$, we have

$$\Phi(x) - \lim_{d\to+\infty}\frac{2c_1}{\sqrt{\pi d}} < \lim_{d\to+\infty}\int_{-\sqrt{2d}}^{x}\psi_d(t)dt < \Phi(x) + \lim_{d\to+\infty}\frac{2c_2}{\sqrt{\pi d}}, \tag{55}$$

which by definition means

$$\lim_{d\to\infty}\Psi_d(\frac{1}{2}+\frac{x}{2\sqrt{2d}}) = \Phi(x), \tag{56}$$

and the error bound is $O(1/\sqrt{d})$. □

In addition, we show $\Psi_{\frac{d-1}{2}}(\frac{1}{2}+\frac{x}{2\sqrt{d}})$ and $\Phi(x)$ in the same figure for $d \in [10^1, 10^2, 10^3, 10^4, 10^5, 10^6]$ in Figure 9. From the figure, only when $d = 10$ can we perceive the errors with eyes. Furthermore, to see the error more clearly, we uniformly select 100000 points in the interval $(0, \sqrt{d})$, and report the maximum absolute error and the maximum relative error among all the points in Table 6, where we see both errors match the $O(1/d)$ error bound in Ryder (2012), slightly tighter than our proved $O(1/\sqrt{d})$ bound.

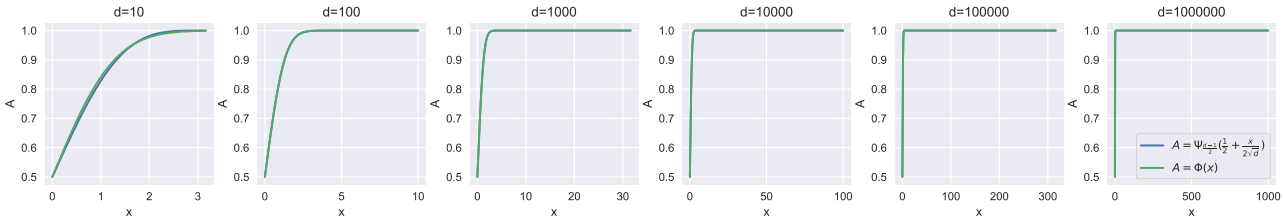


Figure 9. Comparison of $\Psi_{\frac{d-1}{2}}(\frac{1}{2}+\frac{x}{2\sqrt{d}})$ and $\Phi(x)$ for different d.

Table 6. Maximum absolute errors and maximum relative errors in 100000 points in $(0, \sqrt{d})$ for different d

d	10^1	10^2	10^3	10^4	10^5	10^6
AE	1.46e-2	1.38e-3	1.38e-4	1.38e-5	1.38e-6	1.38e-7
RE	1.93e-2	1.83e-3	1.82e-4	1.82e-5	1.82e-6	1.82e-7

C. Computational methods for ESG and EGG in DSRS

To evaluate the performance of ESG/EGG on real-world datasets, we consider solving Problem (57), the strong dual problem of Problem (5):

$$\begin{aligned} \max_{\nu_1,\nu_2\in\mathbb{R}} \quad & \mathbb{P}_{z\sim\mathcal{P}+\delta}\{p(z-\delta)+\nu_1 p(z)+\nu_2 q(z)<0\}, \\ \text{s.t.} \quad & \mathbb{P}_{z\sim\mathcal{P}}\{p(z-\delta)+\nu_1 p(z)+\nu_2 q(z)<0\} = A, \\ & \mathbb{P}_{z\sim\mathcal{Q}}\{p(z-\delta)+\nu_1 p(z)+\nu_2 q(z)<0\} = B. \end{aligned} \tag{57}$$

We do not elaborate on the solution process since previous works (Yang et al., 2020; Li et al., 2022) have done well. Herein, we directly give solutions to Problem (57) for ESG and EGG. Overall, the case of EGG is a relatively straightforward generalization of General Gaussian, while the derivation for ESG includes nontrivial branches. When taking ESG as the smoothing distribution and TESG as the supplementary distribution, we solve Problem (57) by the following theorem:

Theorem C.1 (Integral form of Problem (57) with ESG)**.** *Let $\mathcal{P} = \mathcal{S}(\sigma, \eta)$ with PDF $p(\cdot)$, $\mathcal{Q} = \mathcal{S}_t(\sigma, \eta, T)$ with PDF $q(\cdot)$ and $C_s = \frac{\Gamma(\frac{d}{\eta})}{\gamma(\frac{d}{\eta}, \frac{T^\eta}{2\sigma_s^\eta})}$, where $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function. Let $\mathcal{W} \triangleq \{z \mid p(z-\delta) + \nu_1 p(z) + \nu_2 q(z) < 0\}$. Then*

$$
\begin{aligned}
\mathbb{P}_{z\sim\mathcal{P}}\{z \in \mathcal{W}\} &= \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)} \begin{cases} \omega_1(u,\nu_1), & u \geq \dfrac{T^\eta}{2\sigma_s^\eta}, \\ \omega_1(u,\nu_1 + C_s\nu_2), & u < \dfrac{T^\eta}{2\sigma_s^\eta}, \end{cases} \\
\mathbb{P}_{z\sim\mathcal{Q}}\{z \in \mathcal{W}\} &= C_s \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\, \omega_1(u,\nu_1 + C_s\nu_2) \cdot \mathbb{1}_{u \leq \frac{T^\eta}{2\sigma_s^\eta}}, \\
\mathbb{P}_{z\sim\mathcal{P}+\delta}\{z \in \mathcal{W}\} &= \begin{cases} \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)}\, \omega_2(u), & \nu_1 \geq 0, \\ \mathbb{E}_{u\sim\Gamma(\frac{d}{\eta},1)} \{\omega_2(u) + \omega_3(u)\}, & \nu_1 < 0, \end{cases}
\end{aligned}
\tag{58}
$$

where

$$
\begin{aligned}
\omega_1(u,\nu) &= \begin{cases} \Psi_{\frac{d-1}{2}}\left(\dfrac{(\rho + \sigma_s(2u)^{\frac{1}{\eta}})^2 - 2^{\frac{2}{\eta}}\sigma_s^2(u - \ln(-\nu))^{\frac{2}{\eta}}}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right) & , u - \ln(-\nu) \geq 0, \\ 1 & , u - \ln(-\nu) < 0, \end{cases} \\
\omega_2(u) &= \begin{cases} \Psi_{\frac{d-1}{2}}\left(\dfrac{\min\{T^2, 2^{\frac{2}{\eta}}\sigma_s^2(u + \ln(-\nu_1 - C_s\nu_2))^{\frac{2}{\eta}}\} - (\sigma_s(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right) & , u + \ln(-(\nu_1 + C_s\nu_2)) \geq 0, \\ 0 & , u + \ln(-(\nu_1 + C_s\nu_2)) < 0, \end{cases} \\
\omega_3(u) &= \begin{cases} \max\left\{\Psi_{\frac{d-1}{2}}\left(\dfrac{2^{\frac{2}{\eta}}\sigma_s^2(u + \ln(-\nu_1))^{\frac{2}{\eta}} - (\sigma_s(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right) - \Psi_{\frac{d-1}{2}}\left(\dfrac{T^2 - (\sigma_s(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_s(2u)^{\frac{1}{\eta}}}\right), 0\right\} & , u + \ln(-\nu_1) \geq 0, \\ 0 & , u + \ln(-\nu_1) < 0. \end{cases}
\end{aligned}
\tag{59}
$$

Similarly, We have the solution to Problem (57) for EGG as follows:

Theorem C.2 (Integral form of Problem (57) with EGG)**.** *Let $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ with PDF $p(\cdot)$, $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$ with PDF $q(\cdot)$ and $C_g = \frac{\Gamma(\frac{(d-2k)}{\eta})}{\gamma(\frac{d-2k}{\eta}, \frac{T^\eta}{2\sigma_g^\eta})}$, where $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function. Let $\mathcal{W} \triangleq \{z \mid p(z-\delta) + \nu_1 p(z) + \nu_2 q(z) < 0\}$. Then*

$$
\begin{aligned}
\mathbb{P}_{z\sim\mathcal{P}}\{z \in \mathcal{W}\} &= \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)} \begin{cases} \omega_1(u,\nu_1), & u \geq \dfrac{T^\eta}{2\sigma_g^\eta}, \\ \omega_1(u,\nu_1 + C_g\nu_2), & u < \dfrac{T^\eta}{2\sigma_g^\eta}, \end{cases} \\
\mathbb{P}_{z\sim\mathcal{Q}}\{z \in \mathcal{W}\} &= C_g \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\, \omega_1(u,\nu_1 + C_g\nu_2) \cdot \mathbb{1}_{u \leq \frac{T^\eta}{2\sigma_g^\eta}}, \\
\mathbb{P}_{z\sim\mathcal{P}+\delta}\{z \in \mathcal{W}\} &= \begin{cases} \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\, \omega_2(u), & \nu_1 \geq 0, \\ \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)} \{\omega_2(u) + \omega_3(u)\}, & \nu_1 < 0, \end{cases}
\end{aligned}
\tag{60}
$$

where

$$
\begin{aligned}
\omega_1(u,\nu) &= \Psi_{\frac{d-1}{2}}\left(\frac{(\rho + \sigma_g(2u)^{\frac{1}{\eta}})^2 - \left(\frac{4kW(\frac{\eta u}{2k}(-\nu)^{\frac{\eta}{2k}}\exp(\frac{\eta u}{2k}))\sigma_g^\eta}{\eta}\right)^{\frac{2}{\eta}}}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right), \\
\omega_2(u) &= \Psi_{\frac{d-1}{2}}\left(\frac{\min\{T^2, \sigma_g^2(\frac{4kW(\frac{\eta u}{2k}(-\nu_1 - C_g\nu_2)^{\frac{\eta}{2k}}\exp(\frac{\eta u}{2k}))}{\eta})^{\frac{2}{\eta}}\} - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right), \\
\omega_3(u) &= \Psi_{\frac{d-1}{2}}\left(\frac{\sigma_g^2(\frac{4kW(\frac{\eta u}{2k}(-\nu_1)^{\frac{\eta}{2k}}\exp(\frac{\eta u}{2k}))}{\eta})^{\frac{2}{\eta}} - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right) - \Psi_{\frac{d-1}{2}}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right).
\end{aligned}
\tag{61}
$$

Compared to Theorem C.2, we emphasize the following differences: (1) $p(\cdot)$ represents different PDFs in the two theorems; (2) there are more branches for ω functions in Theorem C.1, which originates from different properties of the logarithmic function and the Lambert W function.

D. Proof of Theorem 6.2 and Theorem D.7

We put proofs of Theorem 6.2 and Theorem D.7 together since they share common thinking. Overall, this section includes 3 parts. We first introduce lemmas based mainly on the concentration properties of beta and gamma distributions, then derive the solution for a lower bound of Problem (5). Finally, we prove Theorem 6.2 and Theorem D.7 respectively, on the basis of those introduced lemmas. Both proofs are essentially generalizations for appendix F.3 in Li et al. (2022).

D.1. Preliminaries

We show lemmas for proofs in this section. In a nutshell, Lemma D.1 offers the probability that the mass of ESG is within T; Lemma D.3 gives a solution for a lower bound of Problem (5); Lemma D.4 and Lemma D.5 reveal concentration properties of beta and gamma distributions; Lemma D.6 proves monotonicity of a function that appears in the proof of Theorem D.7.

Lemma D.1. *For a random variable $z \sim \mathcal{S}(\sigma, \eta)$ and a determined threshold $T \in \mathbb{R}_+$,*

$$\mathbb{P}\{\|z\|_2 \leq T\} = \Lambda_{\frac{d}{\eta}}(\frac{T^\eta}{2\sigma_s^\eta}). \tag{62}$$

Proof. We have ϕ_s, the PDF of z from Equation (21a) and $V_d(r)$ from Lemma B.1, then

$$\begin{aligned}\mathbb{P}\{\|z\|_2 \leq T\} &= \int_0^T \phi_s(r)\exp(-\frac{r^\eta}{2\sigma_s^\eta})\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}r^{d-1}dr \\ &= \frac{1}{\Gamma(\frac{d}{\eta})}\int_0^{\frac{T^\eta}{2\sigma_s^\eta}} t^{\frac{d}{\eta}-1}\exp(-t)dt \\ &= \Lambda_{\frac{d}{\eta}}(\frac{T^\eta}{2\sigma_s^\eta}).\end{aligned} \tag{63}$$

□

Lemma D.2. *Given a base classifier $f : \mathbb{R}^d \to \mathcal{Y}$ satisfies (σ, p, η)-concentration property at input $(x_0, y_0) \in \mathbb{R}^d$. For the supplementary distribution $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$, where $T = \sigma_s(2\Lambda^{-1}_{\frac{d}{\eta}}(p))^{\frac{1}{\eta}}, \eta \in \mathbb{R}_+$ and $2d - k \in [1, 30] \cap \mathbb{N}$, we have*

$$\mathbb{P}_{z\sim\mathcal{Q}}\{f(x_0 + z) = y_0\} = 1. \tag{64}$$

Proof. If f satisfies the (σ, p, η)-concentration property, when p is fixed, we have

$$\mathbb{P}_{z\sim S(\sigma,\eta)}\{f(x_0 + z) = y_0 \mid \|z\|_2 \leq T\} = 1 \tag{65}$$

for $T = \sigma_s(2\Lambda^{-1}_{\frac{d}{\eta}}(p))^{\frac{1}{\eta}}$ from Definition 6.1 and Lemma D.1. Notice that though Equation (65) is defined by $S(\sigma_s, \eta)$, $f(x_0 + z) = y_0$ holds for almost all $\|z\|_2 \leq T$ since distribution $\mathcal{S}$ has positive density almost everywhere. Consider the case $z \sim \mathcal{G}_t(\sigma, \eta, k, T)$; we thereby have

$$\mathbb{P}_{z\sim\mathcal{G}_t(\sigma,\eta,k,T)}\{f(x_0 + z) \neq y_0 \mid 0 < \|z\|_2 \leq T\} = 0. \tag{66}$$

By Equation (23a), we have

$$\begin{aligned}&\mathbb{P}_{z\sim\mathcal{G}_t(\sigma,\eta,k,T)}\{z = 0\} \\ =&\mathbb{P}_{z\sim\mathcal{G}(\sigma,\eta,k)}\{z = 0 \mid \|z\|_2 \leq T\} \\ \leq&\lim_{t\to 0}\mathbb{P}_{z\sim\mathcal{G}(\sigma,\eta,k)}\{\|z\|_2 \leq t\} \\ =&\lim_{t\to 0}\int_0^t \frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-2k}{\eta})}r^{-2k}\exp(-\frac{1}{2}(\frac{r}{\sigma_g})^\eta)\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}r^{d-1}dr \\ =&\frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-2k}{\eta})}\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}\lim_{t\to 0}\int_0^t \exp(-\frac{1}{2}(\frac{r}{\sigma_g})^\eta)r^{d-2k-1}dr \\ =&0.\end{aligned} \tag{67}$$

Therefore, we see

$$\mathbb{P}_{z\sim\mathcal{G}_t(\sigma,\eta,k,T)}\{f(x_0+z)=y_0\}=\mathbb{P}_{z\sim\mathcal{G}(\sigma,\eta,k)}\{f(x_0+z)=y_0 \mid 0\le\|z\|_2\le T\}=1, \tag{68}$$

which concludes the proof. □

Here we give a solution for a lower bound of Problem (5). The lower bound found with the help of Lemma D.2 can be solved by the level set method, sharing the same thinking with proofs of Theorem C.1 and C.2. We refer the readers to previous papers (Yang et al., 2020; Li et al., 2022) for more details.

Lemma D.3. *Under the setting of Lemma D.2, we let*

$$\begin{aligned} R=\max\quad & \rho, \\ \text{s.t.}\quad & \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\Psi_{\frac{d-1}{2}}\left(\frac{T^2-(\sigma_g(2u)^{\frac{1}{\eta}}-\rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right)\ge\frac{1}{2}. \end{aligned} \tag{69}$$

If R_D is the tightest ℓ_2 certified radius when $\mathcal{P}=\mathcal{G}(\sigma,\eta,k)$ is the smoothing distribution by DSRS, then $R_D\ge R$.

Proof. According to Lemma D.2, we define $\mathcal{Q}=\mathcal{G}_t(\sigma,\eta,k,T)$ as the supplementary distribution for $\mathcal{P}$. Then the Problem (5) can be simplified as

$$\begin{aligned} &\min_{\tilde{f}_{x_0}\in\mathcal{F}}\quad \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(\delta+z)\right), \\ &\text{s.t.}\quad \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(z)\right)=A,\ \mathbb{E}_{z\sim\mathcal{Q}}\left(\tilde{f}_{x_0}(z)\right)=B \\ \overset{(a)}{\ge}&\min_{\tilde{f}_{x_0}\in\mathcal{F}}\quad \mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}_{x_0}(\delta+z)\right), \\ &\text{s.t.}\quad \mathbb{E}_{z\sim\mathcal{Q}}\left(\tilde{f}_{x_0}(z)\right)=1, \end{aligned} \tag{70}$$

where (a) is because the subjection $\mathbb{E}_{z\sim\mathcal{P}}[\tilde{f}_{x_0}(z)]=A$ offers extra information outside $\{x \mid \|x-x_0\|_2\le T\}$ in $\mathbb{R}^n$, where $p(x)\ge 0$. It is obvious that the equality holds if $A=0$. For the simplified problem, we have the worst function $\tilde{f}^*_{x_0}(t)$:

$$\tilde{f}^*_{x_0}(t)=\begin{cases}1, \|t\|_2\le T,\\ 0, \|t\|_2>T.\end{cases} \tag{71}$$

With the worst classifier, a lower bound for Problem (5) can finally be written as

$$\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \tag{72}$$

for a fixed δ. Similar to Theorem C.1 and Theorem C.2, the final simplified problem is solved by the level set method (Yang et al., 2020). We have

$$\begin{aligned} &\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right)\\ =&\mathbb{E}_{\epsilon\sim\mathcal{P}}\{\|\delta+x\|_2\le T\}\\ =&\int_0^\infty\frac{\eta}{2}\frac{1}{(2\sigma_g^\eta)^{\frac{d-2k}{\eta}}\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-2k}{\eta})}r^{d-2k-1}\exp[-\frac{1}{2}(\frac{r}{\sigma_g})^\eta]\frac{d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}dr\mathbb{P}\{\|x+\delta\|_2\le T \mid \|x\|_2=r\}\\ =&\frac{1}{\Gamma(\frac{d-2k}{\eta})}\int_0^\infty u^{\frac{d-2k}{\eta}-1}\exp(-u)du\mathbb{P}\{\|x+\delta\|_2\le T \mid \|x\|_2=\sigma_g(2u)^{\frac{1}{\eta}}\}. \end{aligned} \tag{73}$$

As $\|x+\delta\|_2\le T\iff x_1\le\frac{T^2-\rho^2-\sigma_g^2(2u)^{\frac{2}{\eta}}}{2\rho}$, we have

$$\frac{1+\frac{x_1}{\sigma_g(2u)^{\frac{1}{\eta}}}}{2}\sim\text{Beta}(\frac{d-1}{2},\frac{d-1}{2}) \tag{74}$$

by Lemma I.23 of (Yang et al., 2020). Thus, we get

$$\mathbb{P}\{\|x+\delta\|_2 \le T \mid \|x\|_2 = \sigma_g(2u)^{\frac{1}{\eta}}\} = \Psi_{\frac{d-1}{2}}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right). \tag{75}$$

Combining Equation (73) and Equation (75), we finally get

$$\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) = \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\Psi_{\frac{d-1}{2}}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right). \tag{76}$$

If we can ensure that $\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \ge 0.5$ for $\delta = (\rho, 0, 0, \cdots, 0)^T$, $\|\delta\|_2$ will be qualified as a ℓ_2 certified radius due to ℓ_2-symmetry in the R^n space (Zhang et al., 2020). As R is the solution to the lower-bound problem of Problem (5), we have $R_D \ge R$, which concludes the proof. □

The next two lemmas describe the concentration of the gamma distribution and the beta distribution, both are required for the lower estimation of Problem (5).

Lemma D.4. *(Concentration of the beta distribution) Let $\tau \in (\frac{1}{2}, 1)$, $\theta \in (0,1)$. Then there exist $d_0 \in \mathbb{N}_+$, for any $d \ge d_0$,*

$$\Psi_{\frac{d-1}{2}}(\tau) \ge \theta. \tag{77}$$

Proof. Let $X \sim \mathrm{Beta}(\frac{d-1}{2}, \frac{d-1}{2})$. By property of the beta distribution, we have $\mathbb{E}X = \frac{1}{2}$, $\mathbb{D}X = \frac{1}{4d}$, and

$$\mathbb{P}\{X > \tau\} = \mathbb{P}\{X - \frac{1}{2} > \tau - \frac{1}{2}\} \le \mathbb{P}\{|X - \frac{1}{2}| \ge \tau - \frac{1}{2}\} \le \frac{1}{4d(\tau - \frac{1}{2})^2}. \tag{78}$$

We then have

$$\Psi_{\frac{d-1}{2}}(\tau) = 1 - \mathbb{P}\{X > \tau\} \ge 1 - \frac{1}{4d(\tau - \frac{1}{2})^2}. \tag{79}$$

Let $1 - \frac{1}{4d(\tau-\frac{1}{2})^2} \ge \theta$, then $d \ge \frac{1}{4(1-\theta)(\tau-\frac{1}{2})^2}$. Picking $d_0 = \lceil \frac{1}{4(1-\theta)(\tau-\frac{1}{2})^2} \rceil$ concludes the proof. □

Lemma D.5. *(Unilateral concentration of the gamma distribution) Let $p \in (0,1)$, $d \in \mathbb{N}_+$, $\eta \in \mathbb{R}_+$ and $\beta \in (0,1)$. Then there exist $d_0 \in \mathbb{N}_+$, for any $d \ge d_0$,*

$$\Lambda_{\frac{d}{\eta}}\left(\frac{\beta d}{\eta}\right) \le p. \tag{80}$$

Proof. Let $X \sim \Gamma(\frac{d}{\eta}, 1)$, then $\mathbb{E}X = \frac{d}{\eta}$, $\mathbb{D}X = \frac{d}{\eta}$,

$$\Lambda_{\frac{d}{\eta}}\left(\frac{\beta d}{\eta}\right) = \mathbb{P}\left\{X \le \frac{\beta d}{\eta}\right\} \le \mathbb{P}\left\{|X - \frac{d}{\eta}| \ge \frac{(1-\beta)d}{\eta}\right\} \le \frac{\eta}{(1-\beta)^2 d}. \tag{81}$$

If $\frac{\eta}{(1-\beta)^2 d} \le p$, then $d \ge \frac{\eta}{(1-\beta)^2 p}$. Let $d_0 = \lceil \frac{\eta}{(1-\beta)^2 p} \rceil$, for any $d \ge d_0$, we have $\frac{\eta}{(1-\beta)^2 d} \le \frac{\eta}{(1-\beta)^2 d_0} \le p$, which concludes the proof. □

Remark. Both Lemma D.4 and Lemma D.5 illustrate that random variables following beta and gamma distributions are highly concentrated towards their respective expectations in the high-dimensional setting.

The next lemma is required in the proof of Theorem D.7.

Lemma D.6. *Let $x \in \mathbb{N}_+$, $\eta = \frac{1}{n}, n \in \mathbb{N}_+$, $g(x) = \frac{x}{\left(\prod_{i=1}^{\frac{2}{\eta}}(\frac{x+2}{\eta} - i)\right)^{\frac{\eta}{2}}}$. Then $g(x)$ is a non-decreasing function with respect to x.*

Proof. Obviously $g(x) > 0$. Let $h(x) = \ln g(x)$, then

$$\frac{\mathrm{d}h(x)}{\mathrm{d}x} \geq 0 \Longleftrightarrow \frac{1}{g(x)}\frac{\mathrm{d}g(x)}{\mathrm{d}x} \geq 0 \Longleftrightarrow \frac{\mathrm{d}g(x)}{\mathrm{d}x} \geq 0, \tag{82}$$

Therefore,

$$\frac{\mathrm{d}h(x)}{\mathrm{d}x} = \frac{1}{x} - \frac{\eta}{2}\sum_{i=1}^{\frac{2}{\eta}} \frac{1}{x+2-i\eta} \geq \frac{1}{x} - \frac{\eta}{2}\sum_{i=1}^{\frac{2}{\eta}} \frac{1}{x} = 0. \tag{83}$$

Thus, for all $x \in \mathbb{N}_+$, we have $\frac{\mathrm{d}g(x)}{\mathrm{d}x} \geq 0$, which concludes the proof. □

D.2. Proof of Theorem 6.2

Proof. We let the smoothing distribution $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ and the supplementary distribution $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$. In addition, we suppose the base classifier satisfies $(\sigma, p, 2)$-concentration property, which implies the base classifier is highly robust for restricted Gaussian noises. For the convenience of future discussions, we parameterize 0.02 as μ, and the worst classifier $\tilde{f}^*_{x_0}$ is defined the same as that in Lemma D.2.

We see the condition $\eta = 2$ simplifies some lemmas above. Let $\eta = 2$ in Lemma D.1, we get $\mathbb{P}_{z\sim\mathcal{S}(\sigma,2)}\{\|z\|_2 \leq T\} = \Lambda_{\frac{d}{2}}(\frac{T^2}{2\sigma_s^2})$, which means

$$T = \sigma\sqrt{2\Lambda^{-1}_{\frac{d}{2}}(p)}. \tag{84}$$

By definition of $(\sigma, p, 2)$-concentration and Lemma D.2, we have $\mathbb{P}_{z\sim\mathcal{G}_t(\sigma,\eta,k,T)}\{f(x_0+z) = y_0\} = 1$, thus we can find a lower bound to estimate Problem (5) by Lemma D.3. Let $\eta = 2$ in Lemma D.4, we see $\Lambda_{\frac{d}{2}}(\frac{\beta d}{2}) \leq p$. We have thereby

$$T \geq \sigma\sqrt{\beta d}. \tag{85}$$

Equation (85) will be used to find the lower bound of Problem (72). We then consider the solution of the lower-bound Problem (72) in Lemma D.3. We have

$$\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}} \geq \tau \Longleftrightarrow \sigma_g(2u)^{\frac{2}{\eta}} - (2-4\tau)\rho\sigma_g(2u)^{\frac{1}{\eta}} + \rho^2 - T^2 \leq 0. \tag{86}$$

Notice Equation (86) is a one-variable quadratic inequality with respect to $\sigma_g(2u)^{\frac{1}{\eta}}$. Let $\rho = \mu\sigma\sqrt{d}$ where the constant $\mu \in \mathbb{R}_+$. When the discriminant Δ for Equation (86) is positive, the solution for it is

$$\sigma_g(2u)^{\frac{1}{\eta}} \in \left[0, (1-2\tau)\rho + \sqrt{T^2 + (4\tau^2 - 4\tau)\rho^2}\right) \Longleftrightarrow 0 \leq u < \frac{1}{2}\left(\frac{(1-2\tau)\rho + \sqrt{T^2 + (4\tau^2-4\tau)\rho^2}}{\sigma_g}\right)^{\eta}. \tag{87}$$

Now we are ready to show the minimization for $\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right)$ for base classifier satisfies $(\sigma, p, 2)$-concentration property. We have

$$\begin{aligned}
&\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right)\\
=&\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\Psi_{\frac{d-1}{2}}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}}\right)\\
\geq&\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(\frac{T^2 - (\sigma_g(2u)^{\frac{1}{\eta}} - \rho)^2}{4\rho\sigma_g(2u)^{\frac{1}{\eta}}} \geq \tau\right)\\
\geq&\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(u < \frac{1}{2}\left(\frac{(1-2\tau)\rho + \sqrt{T^2 + (4\tau^2-4\tau)\rho^2}}{\sigma_g}\right)^{\eta}\right)
\end{aligned}$$

$$\overset{(a)}{\geq} \theta \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)} \mathbb{I}\left(u < \frac{1}{2}\left(\frac{(1-2\tau)\mu\sigma\sqrt{d} + \sqrt{\sigma^2\beta d + (4\tau^2-4\tau)\mu^2\sigma^2 d}}{\sigma_g}\right)^{\eta}\right)$$
$$\overset{(b)}{=} \theta \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)} \mathbb{I}\left(u < \left(\sqrt{\frac{\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d-2k}{\eta})}}\left((1-2\tau)\mu + \sqrt{\beta + (4\tau^2-4\tau)\mu^2}\right)\right)^{\eta}\right), \tag{88}$$

where $\mathbb{I}(\cdot)$ is the indicator function, (a) is by $\rho = \mu\sigma\sqrt{d}$ and Equation (85); (b) is because

$$\sigma_g = 2^{-\frac{1}{\eta}}\sqrt{\frac{d\Gamma(\frac{d-2k}{\eta})}{\Gamma(\frac{d-2k+2}{\eta})}}\sigma \tag{89}$$

by definition. We write

$$m = \left(\sqrt{\frac{\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d-2k}{\eta})}}\left((1-2\tau)\mu + \sqrt{\beta + (4\tau^2-4\tau)\mu^2}\right)\right)^{\eta}, \tag{90}$$

to get

$$\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \geq \theta \mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)} \mathbb{1}_{u<m} = \theta\Lambda_{\frac{d-2k}{\eta}}(m). \tag{91}$$

Let $\theta = 0.999, \beta = 0.99, \tau = 0.6,\ \mu = 0.02$ (Li et al., 2022). We show the value of $\Lambda_{\frac{d-2k}{\eta}}(m)$ when $d-2k \in [1,30]\cap\mathbb{Z}$ and $\eta = \frac{1}{n}, n \in [1,50]\cap\mathbb{Z}$ in Table 7. Observing that there's no value greater than $\frac{1}{2\theta}$, we have $\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \geq \frac{1}{2}$, which concludes the proof. □

Algorithm 1: Algorithm for finding tight μ for the $\Omega(\sqrt{d})$ lower bound

Input: input dimension d, hyperparameters k, β, τ, θ, exponent η, error limitation e
Output: tight constant μ_l for a specified EGG

```
μ_l, μ_r ← 0, 1
while μ_r − μ_l > e do
    μ_m ← (μ_r + μ_l)/2
    m_m ← θΛ_{(d−2k)/η}( ( √(Γ((d−2k+2)/η)/Γ((d−2k)/η)) ((1−2τ)μ_m + √(β + (4τ²−4τ)μ_m²)) )^η )
    if m_m > 1/2 then
        μ_l ← μ_m
    else
        μ_r ← μ_m
    end
end
μ ← μ_l
return μ
```

Remark. We exhaust the cases for η since the analytic solution to Problem (72) includes intractable gamma function terms. The proof above can easily be generalized to other $\eta \in \mathbb{R}_+$. e.g., we have tried the sequence of $\eta \in [0.02, 1]$ increasing by 0.001, no value greater than $\frac{1}{2\theta}$ is observed, meaning these ηs are all qualified to provide $\Omega(\sqrt{d})$ lower bounds for the ℓ_2 certified radius under the setting of Theorem 6.2. We have also shown results for $\eta \in [2, 10]\cap\mathbb{N}$ in Table 7, where the boundary value for $\frac{1}{2\theta}$ is marked red. It is remarkable that $\Lambda_{\frac{d-2k}{\eta}}(m)$ decreases significantly when $\eta > 2$, which is in line with Figure 2 (left) though we are only considering the lower bound for the certified radius. In addition to the fixed constant 0.02, the tight constant factor μ for each EGG distribution can be computed by Algorithm 1.

D.3. Theorem D.7

The following theorem introduces $d^{1/\eta}$ into the lower bound using (σ, p, η)-concentration assumption, which can be proved to be almost equivalent to Theorem 6.2.

Theorem D.7. *Let $d \in \mathbb{N}_+$ be a sufficiently large input dimension, $(x_0, y_0) \in \mathbb{R}^d \times \mathcal{Y}$ be a labeled example and $f : \mathbb{R}^d \to \mathcal{Y}$ be a base classifier which satisfies (σ, p, η)-concentration property w.r.t. (x_0, y_0). For the DSRS method, let $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ be the smoothing distribution to give a smoothed classifier $\bar{f}_{\mathcal{P}}$, and $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$ be the supplementary distribution with $T = \sigma_s\sqrt{2\Lambda^{-1}_{\frac{d}{\eta}}(p)}, d - 2k \in [1, 30] \cap \mathbb{N}$ and $\eta \in \{1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{50}\}$. Then for the smoothed classifier $\bar{f}_{\mathcal{P}}(x)$ the certified ℓ_2 radius*

$$r_\eta \geq 0.02\sigma_s d^{\frac{1}{\eta}}, \tag{92}$$

where σ_s is the formal variance of $\mathcal{S}(\sigma, \eta)$. When σ_s is converted to σ keeping $\mathbb{E}r^2$ a constant, we still have

$$r_\eta = \Omega(\sqrt{d}). \tag{93}$$

Proof. In this section, we first prove that a base classifier that satisfies a certain concentration property can certify $\Theta(d^{\frac{1}{\eta}})$ ℓ_2 radii given the smoothing distribution $\mathcal{P} = \mathcal{G}(\sigma, \eta, k)$ and the supplementary distribution $\mathcal{Q} = \mathcal{G}_t(\sigma, \eta, k, T)$. Then, by converting σ_s to σ, we derive a $\Theta(\sqrt{d})$ lower bound for the certified radius. Like Appendix D.2, we find the lower bound for Problem (5) by Lemma D.3.

In Lemma D.1, when η is an arbitrarily positive real number, we have

$$p = \Lambda_{\frac{d}{\eta}}\left(\frac{T^\eta}{2\sigma_s^\eta}\right) \Longleftrightarrow T = \sigma_s(2\Lambda^{-1}_{\frac{d}{\eta}}(p))^{\frac{1}{\eta}}. \tag{94}$$

We thereby obtain

$$T \geq (\frac{2\beta}{\eta})^{\frac{1}{\eta}}\sigma_s d^{\frac{1}{\eta}} \tag{95}$$

by Lemma D.5. Then we have Equation (86) and Equation (87) the same as in Appendix D.2. Let $\rho = \zeta\sigma_s d^{\frac{1}{\eta}}$ where $\zeta \in \mathbb{R}_+$. Let $\eta = \frac{1}{n}$, $n \in \mathbb{N}_+, \forall d \geq \tilde{d}$, where $\tilde{d}$ is a sufficiently large real integer which satisfies Lemma D.4 and Lemma D.5. We then have the estimation

$$\begin{aligned}
&\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta + z)\right)\\
&\overset{(a)}{\geq}\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(u < \frac{1}{2}\left(\frac{(1-2\tau)\rho + \sqrt{T^2 + (4\tau^2 - 4\tau)\rho^2}}{\sigma_g}\right)^\eta\right)\\
&\overset{(b)}{=}\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(u < \frac{1}{2}\left(\sqrt{\frac{\Gamma(\frac{d}{\eta})\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d+2}{\eta})\Gamma(\frac{d-2k}{\eta})}}\frac{(1-2\tau)\rho + \sqrt{T^2 + (4\tau^2 - 4\tau)\rho^2}}{\sigma_s}\right)^\eta\right)\\
&\overset{(c)}{\geq}\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(u < \frac{1}{2}\left(\sqrt{\frac{\Gamma(\frac{d}{\eta})\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d+2}{\eta})\Gamma(\frac{d-2k}{\eta})}}\frac{(1-2\tau)\zeta\sigma_s d^{\frac{1}{\eta}} + \sqrt{(\frac{2\beta}{\eta})^{\frac{2}{\eta}}\sigma_s^2 d^{\frac{2}{\eta}} + (4\tau^2 - 4\tau)(\zeta\sigma_s d^{\frac{1}{\eta}})^2}}{\sigma_s}\right)^\eta\right)\\
&\overset{(d)}{\geq}\theta\mathbb{E}_{u\sim\Gamma(\frac{d-2k}{\eta},1)}\mathbb{I}\left(u < \frac{\tilde{d}}{2\left(\prod_{i=1}^{\frac{2}{\eta}}(\frac{\tilde{d}+2}{\eta} - i)\right)^{\frac{\eta}{2}}}\left(\sqrt{\frac{\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d-2k}{\eta})}}\left((1-2\tau)\zeta + \sqrt{(\frac{2\beta}{\eta})^{\frac{2}{\eta}} + (4\tau^2 - 4\tau)\zeta^2}\right)\right)^\eta\right).
\end{aligned} \tag{96}$$

In the equations above: (a) solve the inequality with respect to u in the indicator function, whose solution is shown in Equation (87); (b) for a constant $\mathbb{E}r^2$, we have

$$\sigma_g = \sqrt{\frac{\Gamma(\frac{d-2k}{\eta})\Gamma(\frac{d+2}{\eta})}{\Gamma(\frac{d}{\eta})\Gamma(\frac{d-2k+2}{\eta})}}\sigma_s; \tag{97}$$

(c) by Lemma D.5 and $\rho = \zeta\sigma_s d^{\frac{1}{\eta}}$; (d) by Lemma D.6.

We write

$$m = \frac{\tilde{d}}{2\left(\prod_{i=1}^{\frac{2}{\eta}}(\frac{\tilde{d}+2}{\eta}-i)\right)^{\frac{\eta}{2}}}\left(\sqrt{\frac{\Gamma(\frac{d-2k+2}{\eta})}{\Gamma(\frac{d-2k}{\eta})}}\left((1-2\tau)\zeta+\sqrt{(\frac{2\beta}{\eta})^{\frac{2}{\eta}}+(4\tau^2-4\tau)\zeta^2}\right)\right)^{\eta}, \tag{98}$$

and then we have

$$\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \geq \Lambda_{\frac{d-2k}{\eta}}(m) \tag{99}$$

by Equation (96). Notice the lower estimation is slightly different from Appendix D.2 since there is a $\tilde{d}$ in the expression. Pick $p = 0.5,\ \tilde{d} = 25000, \theta = 0.999, \beta = 0.99, \tau = 0.6,\ \zeta = 0.02$ (Li et al., 2022). We show the value of $\Lambda_{\frac{d-2k}{\eta}}(m)$ when $d-2k \in [1,30]\cap\mathbb{Z}$ and $\eta = \frac{1}{n}, n \in [1,50]\cap\mathbb{Z}$ in Table 8. As a result, all values in Table 8 are greater than $\frac{1}{2\theta} \approx 0.5005$, which means

$$\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \geq \theta\cdot\frac{1}{2\theta} = \frac{1}{2}. \tag{100}$$

Recalling that our goal here is to check whether $\mathbb{E}_{z\sim\mathcal{P}}\left(\tilde{f}^*_{x_0}(\delta+z)\right) \geq \frac{1}{2}$ holds for some determined ρ, we have finished the proof that $R_\eta \geq \rho = \zeta\sigma_s d^{\frac{1}{\eta}}$. Superficially, we get an $\Omega(d^{\frac{1}{\eta}})$ bound for the ℓ_2 certified radius, which seems tighter than the $\Omega(\sqrt{d})$ one. But by Lemma 5.4, it is essentially equivalent to the $\Omega(\sqrt{d})$ lower bound; we thus remark that Theorem 6.2 and Theorem D.7 are different manifestations of the same fact. □

Remark. Different from Theorem 6.2, this proof can not be generalized directly to $\eta > 2$ due to the property of factorial. In essence, the (σ, p, η)-concentration assumption is slightly less strict for base classifiers than the $(\sigma, p, 2)$-concentration assumption.

E. Supplementary for experimental methods

E.1. DSRS algorithm

Taking EGG as an example, the DSRS algorithm used in this work is shown in Algorithm 2.

E.2. Computational overhead

All of our experiments on real-world datasets are composed of sampling and certification, which are finished with 4 NVIDIA RTX 3080 GPUs and CPUs. The most computationally intensive procedure is sampling. For $\sigma = 0.50$ base classifiers, it takes about 5s, 200s to sample 50000 times under noise distributions on CIFAR10, ImageNet with one GPU. Given that we uniformly pick 1000 data points from each dataset, one sampling procedure takes around 1 hour, 1 day for 50000 noises on CIFAR10, ImageNet respectively with one GPU. Naturally, the sampling time almost doubles for 100000 noises.

The computation for certification only relies on CPUs. The running time for certification includes NP certification and DSRS certification, and is basically constant for different datasets. Usually, it takes 1-2 days to complete. The overall computational time for standard DSRS certification is strictly larger than that for pure NP certification, with respect to a specific number of samples. For instance, if we compute a certificate for 100000 noises, we need one NP certification for 100000 noises, while we need one NP certification for 50000 noises and one DSRS certification for the remaining 50000 noises to finish the DSRS computation. Generally, the computational time for standard DSRS certification is one to two times that of pure NP certification.

E.3. Transferability of the incremental effect

We show results for real-world datasets on different base classifiers in this section. In Table 9, we observe that $\eta = 8.0$ shows better performance than General Gaussian ($\eta = 2.0$) overall. This alludes to some defects in the current training method for General Gaussian because intuitively, if the models are trained by General Gaussian, then sampling with General Gaussian should have provided the best certified results among all η, but we find the incremental effect with η still exists in General-Gaussian-augmented models. We name the base classifiers by {training method}-{distribution} in experimental results for ACR. In all the experiments in this subsection, we set the training method $\in$ {StdAug (standard augmentation),

Table 7. Value of $\Lambda_{\frac{d-2k}{\eta}}(m)$ for Theorem 6.2.

$\eta \backslash d-2k$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
10	0.584	0.515	0.487	0.473	0.465	0.459	0.456	0.453	0.450	0.448	0.447	0.445	0.444	0.442	0.441	0.440	0.439	0.438	0.437	0.436	0.435	0.434	0.433	0.432	0.431	0.430	0.429	0.428	0.427	0.426
9	0.586	0.519	0.492	0.478	0.471	0.465	0.462	0.459	0.456	0.454	0.453	0.451	0.450	0.448	0.447	0.446	0.445	0.443	0.442	0.441	0.440	0.439	0.438	0.437	0.436	0.435	0.435	0.434	0.433	0.432
8	0.590	0.524	0.498	0.485	0.478	0.472	0.469	0.466	0.463	0.461	0.459	0.458	0.456	0.455	0.454	0.452	0.451	0.450	0.449	0.448	0.447	0.446	0.445	0.444	0.443	0.442	0.441	0.440	0.439	0.438
7	0.594	0.531	0.506	0.493	0.486	0.481	0.477	0.474	0.471	0.469	0.467	0.465	0.464	0.462	0.461	0.460	0.458	0.457	0.456	0.455	0.454	0.453	0.452	0.451	0.450	0.449	0.448	0.447	0.446	0.445
6	0.600	0.539	0.515	0.503	0.495	0.490	0.486	0.483	0.480	0.478	0.476	0.474	0.472	0.471	0.469	0.468	0.467	0.465	0.464	0.463	0.462	0.461	0.460	0.459	0.458	0.457	0.456	0.455	0.454	0.453
5	0.609	0.551	0.528	0.515	0.507	0.502	0.497	0.494	0.491	0.489	0.486	0.484	0.482	0.481	0.479	0.478	0.476	0.475	0.474	0.472	0.471	0.470	0.469	0.468	0.467	0.466	0.465	0.464	0.463	0.462
4	0.622	0.567	0.544	0.531	0.522	0.516	0.512	0.508	0.504	0.502	0.499	0.497	0.495	0.493	0.491	0.489	0.488	0.486	0.485	0.484	0.482	0.481	0.480	0.479	0.478	0.477	0.476	0.475	0.474	0.473
3	0.643	0.589	0.566	0.552	0.543	0.536	0.530	0.526	0.522	0.519	0.516	0.513	0.511	0.508	0.506	0.505	0.503	0.501	0.500	0.498	0.497	0.496	0.494	0.493	0.492	0.491	0.490	0.489	0.488	0.487
2	0.678	0.625	0.600	0.584	0.573	0.564	0.558	0.552	0.547	0.543	0.540	0.537	0.534	0.531	0.529	0.526	0.524	0.522	0.521	0.519	0.517	0.516	0.514	0.513	0.512	0.510	0.509	0.508	0.507	0.506
1	0.754	0.697	0.666	0.646	0.631	0.619	0.610	0.602	0.596	0.590	0.585	0.581	0.577	0.573	0.570	0.567	0.564	0.561	0.559	0.557	0.555	0.553	0.551	0.549	0.547	0.546	0.544	0.543	0.541	0.540
1/2	0.841	0.782	0.745	0.720	0.701	0.685	0.673	0.662	0.654	0.646	0.639	0.633	0.628	0.623	0.618	0.614	0.610	0.607	0.604	0.601	0.598	0.595	0.593	0.590	0.588	0.586	0.584	0.582	0.580	0.578
1/3	0.891	0.835	0.797	0.769	0.747	0.730	0.716	0.704	0.693	0.684	0.676	0.669	0.663	0.657	0.652	0.647	0.642	0.638	0.634	0.631	0.628	0.624	0.621	0.619	0.616	0.613	0.611	0.609	0.607	0.605
1/4	0.924	0.872	0.834	0.805	0.783	0.764	0.749	0.736	0.724	0.714	0.706	0.698	0.691	0.684	0.678	0.673	0.668	0.663	0.659	0.655	0.651	0.648	0.644	0.641	0.638	0.635	0.633	0.630	0.628	0.625
1/5	0.946	0.899	0.863	0.834	0.811	0.792	0.776	0.762	0.750	0.739	0.730	0.721	0.714	0.707	0.700	0.695	0.689	0.684	0.679	0.675	0.671	0.667	0.664	0.660	0.657	0.654	0.651	0.648	0.645	0.643
1/6	0.961	0.920	0.885	0.858	0.835	0.815	0.799	0.784	0.772	0.761	0.751	0.742	0.734	0.727	0.720	0.714	0.708	0.702	0.697	0.693	0.688	0.684	0.680	0.677	0.673	0.670	0.667	0.664	0.661	0.658
1/7	0.972	0.936	0.904	0.877	0.854	0.835	0.818	0.804	0.791	0.779	0.769	0.760	0.752	0.744	0.737	0.730	0.724	0.719	0.713	0.708	0.704	0.699	0.695	0.691	0.688	0.684	0.681	0.678	0.675	0.672
1/8	0.979	0.948	0.919	0.893	0.871	0.852	0.835	0.821	0.808	0.796	0.785	0.776	0.767	0.759	0.752	0.745	0.739	0.733	0.728	0.723	0.718	0.713	0.709	0.705	0.701	0.697	0.694	0.691	0.687	0.684
1/9	0.985	0.958	0.931	0.907	0.886	0.867	0.850	0.836	0.822	0.811	0.800	0.790	0.782	0.773	0.766	0.759	0.752	0.746	0.741	0.735	0.730	0.726	0.721	0.717	0.713	0.709	0.706	0.702	0.699	0.696
1/10	0.989	0.966	0.942	0.919	0.898	0.880	0.863	0.849	0.836	0.824	0.813	0.803	0.794	0.786	0.779	0.771	0.765	0.759	0.753	0.747	0.742	0.737	0.733	0.728	0.724	0.720	0.717	0.713	0.710	0.706
1/11	0.992	0.972	0.950	0.929	0.909	0.891	0.875	0.861	0.848	0.836	0.825	0.815	0.806	0.798	0.790	0.783	0.776	0.770	0.764	0.758	0.753	0.748	0.743	0.739	0.735	0.731	0.727	0.723	0.720	0.716
1/12	0.994	0.977	0.957	0.937	0.919	0.902	0.886	0.872	0.859	0.847	0.836	0.826	0.817	0.809	0.801	0.794	0.787	0.780	0.774	0.769	0.763	0.758	0.753	0.749	0.744	0.740	0.736	0.733	0.729	0.725
1/13	0.995	0.982	0.964	0.945	0.927	0.911	0.895	0.882	0.869	0.857	0.846	0.836	0.827	0.819	0.811	0.803	0.797	0.790	0.784	0.778	0.773	0.768	0.763	0.758	0.754	0.749	0.745	0.741	0.738	0.734
1/14	0.997	0.985	0.969	0.952	0.935	0.919	0.904	0.890	0.878	0.866	0.856	0.846	0.837	0.828	0.820	0.813	0.806	0.799	0.793	0.787	0.782	0.776	0.771	0.767	0.762	0.758	0.754	0.750	0.746	0.742
1/15	0.997	0.988	0.973	0.957	0.941	0.926	0.912	0.899	0.886	0.875	0.864	0.854	0.845	0.837	0.829	0.821	0.814	0.808	0.801	0.796	0.790	0.785	0.780	0.775	0.770	0.766	0.762	0.758	0.754	0.750
1/16	0.998	0.990	0.977	0.962	0.947	0.933	0.919	0.906	0.894	0.883	0.872	0.862	0.853	0.845	0.837	0.829	0.822	0.816	0.809	0.803	0.798	0.792	0.787	0.783	0.778	0.773	0.769	0.765	0.761	0.758
1/17	0.999	0.992	0.980	0.967	0.953	0.939	0.925	0.913	0.901	0.890	0.880	0.870	0.861	0.852	0.844	0.837	0.830	0.823	0.817	0.811	0.805	0.800	0.795	0.790	0.785	0.781	0.776	0.772	0.768	0.765
1/18	0.999	0.993	0.983	0.970	0.957	0.944	0.931	0.919	0.907	0.897	0.886	0.877	0.868	0.860	0.852	0.844	0.837	0.830	0.824	0.818	0.812	0.807	0.802	0.797	0.792	0.788	0.783	0.779	0.775	0.771
1/19	0.999	0.994	0.985	0.974	0.961	0.949	0.937	0.925	0.914	0.903	0.893	0.883	0.875	0.866	0.858	0.851	0.844	0.837	0.831	0.825	0.819	0.814	0.808	0.803	0.799	0.794	0.790	0.786	0.782	0.778
1/20	0.999	0.995	0.987	0.977	0.965	0.953	0.941	0.930	0.919	0.909	0.899	0.890	0.881	0.873	0.865	0.857	0.850	0.844	0.837	0.831	0.826	0.820	0.815	0.810	0.805	0.800	0.796	0.792	0.788	0.784
1/21	1.000	0.996	0.989	0.979	0.969	0.957	0.946	0.935	0.924	0.914	0.904	0.895	0.887	0.878	0.871	0.863	0.856	0.850	0.843	0.837	0.832	0.826	0.821	0.816	0.811	0.806	0.802	0.798	0.794	0.790
1/22	1.000	0.997	0.990	0.982	0.972	0.961	0.950	0.939	0.929	0.919	0.910	0.901	0.892	0.884	0.876	0.869	0.862	0.855	0.849	0.843	0.837	0.832	0.827	0.822	0.817	0.812	0.808	0.804	0.799	0.795
1/23	1.000	0.997	0.992	0.984	0.974	0.964	0.954	0.944	0.934	0.924	0.915	0.906	0.897	0.889	0.882	0.874	0.868	0.861	0.855	0.849	0.843	0.837	0.832	0.827	0.822	0.818	0.813	0.809	0.805	0.801
1/24	1.000	0.998	0.993	0.986	0.977	0.967	0.957	0.947	0.938	0.928	0.919	0.911	0.902	0.894	0.887	0.880	0.873	0.866	0.860	0.854	0.848	0.843	0.838	0.833	0.828	0.823	0.819	0.814	0.810	0.806
1/25	1.000	0.998	0.994	0.987	0.979	0.970	0.960	0.951	0.942	0.932	0.924	0.915	0.907	0.899	0.892	0.885	0.878	0.871	0.865	0.859	0.853	0.848	0.843	0.838	0.833	0.828	0.824	0.819	0.815	0.811
1/26	1.000	0.998	0.995	0.989	0.981	0.972	0.963	0.954	0.945	0.936	0.928	0.919	0.911	0.904	0.896	0.889	0.882	0.876	0.870	0.864	0.858	0.853	0.848	0.842	0.838	0.833	0.829	0.824	0.820	0.816
1/27	1.000	0.999	0.995	0.990	0.983	0.975	0.966	0.957	0.948	0.940	0.931	0.923	0.915	0.908	0.901	0.894	0.887	0.881	0.874	0.869	0.863	0.857	0.852	0.847	0.842	0.838	0.833	0.829	0.825	0.821
1/28	1.000	0.999	0.996	0.991	0.984	0.977	0.969	0.960	0.952	0.943	0.935	0.927	0.919	0.912	0.905	0.898	0.891	0.885	0.879	0.873	0.867	0.862	0.857	0.852	0.847	0.842	0.838	0.833	0.829	0.825
1/29	1.000	0.999	0.996	0.992	0.986	0.979	0.971	0.963	0.955	0.946	0.938	0.931	0.923	0.916	0.909	0.902	0.895	0.889	0.883	0.877	0.872	0.866	0.861	0.856	0.851	0.847	0.842	0.838	0.834	0.830
1/30	1.000	0.999	0.997	0.993	0.987	0.980	0.973	0.965	0.957	0.949	0.942	0.934	0.927	0.919	0.912	0.906	0.899	0.893	0.887	0.881	0.876	0.870	0.865	0.860	0.855	0.851	0.846	0.842	0.838	0.834
1/31	1.000	0.999	0.997	0.994	0.988	0.982	0.975	0.967	0.960	0.952	0.945	0.937	0.930	0.923	0.916	0.909	0.903	0.897	0.891	0.885	0.880	0.874	0.869	0.864	0.860	0.855	0.850	0.846	0.842	0.838
1/32	1.000	1.000	0.998	0.994	0.989	0.983	0.977	0.970	0.962	0.955	0.947	0.940	0.933	0.926	0.919	0.913	0.907	0.901	0.895	0.889	0.884	0.878	0.873	0.868	0.863	0.859	0.854	0.850	0.846	0.842
1/33	1.000	1.000	0.998	0.995	0.990	0.985	0.978	0.972	0.964	0.957	0.950	0.943	0.936	0.929	0.923	0.916	0.910	0.904	0.898	0.893	0.887	0.882	0.877	0.872	0.867	0.863	0.858	0.854	0.850	0.846
1/34	1.000	1.000	0.998	0.995	0.991	0.986	0.980	0.973	0.967	0.960	0.953	0.946	0.939	0.932	0.926	0.920	0.913	0.907	0.902	0.896	0.891	0.885	0.880	0.876	0.871	0.866	0.862	0.858	0.853	0.849
1/35	1.000	1.000	0.998	0.996	0.992	0.987	0.981	0.975	0.969	0.962	0.955	0.948	0.942	0.935	0.929	0.923	0.917	0.911	0.905	0.899	0.894	0.889	0.884	0.879	0.874	0.870	0.865	0.861	0.857	0.853
1/36	1.000	1.000	0.999	0.996	0.993	0.988	0.983	0.977	0.970	0.964	0.957	0.951	0.944	0.938	0.932	0.926	0.920	0.914	0.908	0.903	0.897	0.892	0.887	0.882	0.878	0.873	0.869	0.865	0.860	0.856
1/37	1.000	1.000	0.999	0.997	0.993	0.989	0.984	0.978	0.972	0.966	0.959	0.953	0.947	0.940	0.934	0.928	0.922	0.917	0.911	0.906	0.901	0.895	0.890	0.886	0.881	0.877	0.872	0.868	0.864	0.860
1/38	1.000	1.000	0.999	0.997	0.994	0.990	0.985	0.980	0.974	0.968	0.962	0.955	0.949	0.943	0.937	0.931	0.925	0.920	0.914	0.909	0.904	0.899	0.894	0.889	0.884	0.880	0.875	0.871	0.867	0.863
1/39	1.000	1.000	0.999	0.997	0.995	0.991	0.986	0.981	0.975	0.969	0.963	0.957	0.951	0.945	0.939	0.934	0.928	0.922	0.917	0.912	0.906	0.901	0.897	0.892	0.887	0.883	0.879	0.874	0.870	0.866
1/40	1.000	1.000	0.999	0.998	0.995	0.991	0.987	0.982	0.977	0.971	0.965	0.959	0.953	0.948	0.942	0.936	0.930	0.925	0.920	0.914	0.909	0.904	0.900	0.895	0.890	0.886	0.882	0.877	0.873	0.869
1/41	1.000	1.000	0.999	0.998	0.995	0.992	0.988	0.983	0.978	0.973	0.967	0.961	0.955	0.950	0.944	0.938	0.933	0.928	0.922	0.917	0.912	0.907	0.902	0.898	0.893	0.889	0.885	0.880	0.876	0.872
1/42	1.000	1.000	0.999	0.998	0.996	0.993	0.989	0.984	0.979	0.974	0.969	0.963	0.957	0.952	0.946	0.941	0.935	0.930	0.925	0.920	0.915	0.910	0.905	0.900	0.896	0.892	0.887	0.883	0.879	0.875
1/43	1.000	1.000	0.999	0.998	0.996	0.993	0.990	0.985	0.981	0.975	0.970	0.965	0.959	0.954	0.948	0.943	0.938	0.932	0.927	0.922	0.917	0.912	0.908	0.903	0.899	0.894	0.890	0.886	0.882	0.878
1/44	1.000	1.000	1.000	0.998	0.997	0.994	0.990	0.986	0.982	0.977	0.972	0.966	0.961	0.956	0.950	0.945	0.940	0.935	0.930	0.925	0.920	0.915	0.910	0.906	0.901	0.897	0.893	0.889	0.885	0.881
1/45	1.000	1.000	1.000	0.999	0.997	0.994	0.991	0.987	0.983	0.978	0.973	0.968	0.963	0.957	0.952	0.947	0.942	0.937	0.932	0.927	0.922	0.917	0.913	0.908	0.904	0.900	0.895	0.891	0.887	0.884
1/46	1.000	1.000	1.000	0.999	0.997	0.995	0.992	0.988	0.984	0.979	0.974	0.969	0.964	0.959	0.954	0.949	0.944	0.939	0.934	0.929	0.924	0.920	0.915	0.911	0.906	0.902	0.898	0.894	0.890	0.886
1/47	1.000	1.000	1.000	0.999	0.997	0.995	0.992	0.989	0.985	0.980	0.976	0.971	0.966	0.961	0.956	0.951	0.946	0.941	0.936	0.931	0.927	0.922	0.917	0.913	0.909	0.905	0.900	0.896	0.893	0.889
1/48	1.000	1.000	1.000	0.999	0.998	0.996	0.993	0.989	0.985	0.981	0.977	0.972	0.967	0.962	0.957	0.953	0.948	0.943	0.938	0.933	0.929	0.924	0.920	0.915	0.911	0.907	0.903	0.899	0.895	0.891
1/49	1.000	1.000	1.000	0.999	0.998	0.996	0.993	0.990	0.986	0.982	0.978	0.973	0.969	0.964	0.959	0.954	0.949	0.945	0.940	0.935	0.931	0.926	0.922	0.918	0.913	0.909	0.905	0.901	0.897	0.894
1/50	1.000	1.000	1.000	0.999	0.998	0.996	0.994	0.991	0.987	0.983	0.979	0.975	0.970	0.965	0.961	0.956	0.951	0.947	0.942	0.937	0.933	0.928	0.924	0.920	0.916	0.911	0.907	0.904	0.900	0.896

Table 8. Value of $\Lambda_{\frac{d-2k}{\eta}}(m)$ for Theorem D.7.

$\eta \backslash d-2k$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	0.753	0.696	0.665	0.644	0.628	0.617	0.607	0.599	0.592	0.586	0.581	0.577	0.572	0.569	0.565	0.562	0.559	0.556	0.554	0.552	0.549	0.547	0.545	0.543	0.541	0.540	0.538	0.537	0.535	0.534
1/2	0.838	0.778	0.741	0.714	0.694	0.678	0.665	0.654	0.645	0.636	0.629	0.623	0.617	0.611	0.606	0.602	0.598	0.594	0.590	0.587	0.584	0.581	0.578	0.575	0.572	0.570	0.568	0.565	0.563	0.561
1/3	0.889	0.830	0.790	0.761	0.738	0.720	0.704	0.691	0.680	0.670	0.661	0.654	0.646	0.640	0.634	0.628	0.623	0.619	0.614	0.610	0.606	0.602	0.599	0.596	0.592	0.589	0.586	0.584	0.581	0.579
1/4	0.921	0.866	0.827	0.796	0.772	0.752	0.735	0.721	0.708	0.697	0.687	0.678	0.670	0.663	0.656	0.650	0.644	0.638	0.633	0.629	0.624	0.620	0.616	0.612	0.608	0.605	0.602	0.598	0.595	0.592
1/5	0.943	0.894	0.855	0.824	0.799	0.778	0.760	0.745	0.732	0.720	0.709	0.699	0.690	0.682	0.675	0.668	0.661	0.655	0.650	0.645	0.640	0.635	0.630	0.626	0.622	0.618	0.614	0.611	0.608	0.604
1/6	0.959	0.915	0.878	0.847	0.822	0.801	0.782	0.766	0.752	0.739	0.728	0.717	0.708	0.699	0.691	0.684	0.677	0.670	0.664	0.659	0.653	0.648	0.643	0.639	0.634	0.630	0.626	0.622	0.618	0.615
1/7	0.970	0.931	0.896	0.867	0.841	0.820	0.801	0.784	0.770	0.756	0.745	0.734	0.724	0.715	0.706	0.698	0.691	0.684	0.677	0.671	0.666	0.660	0.655	0.650	0.645	0.641	0.636	0.632	0.628	0.624
1/8	0.978	0.944	0.911	0.883	0.858	0.836	0.817	0.801	0.785	0.772	0.760	0.748	0.738	0.728	0.719	0.711	0.703	0.696	0.689	0.683	0.677	0.671	0.666	0.660	0.655	0.651	0.646	0.642	0.637	0.633
1/9	0.983	0.954	0.924	0.897	0.873	0.851	0.832	0.815	0.800	0.786	0.773	0.762	0.751	0.741	0.732	0.723	0.715	0.708	0.700	0.694	0.687	0.681	0.675	0.670	0.665	0.660	0.655	0.650	0.646	0.641
1/10	0.988	0.962	0.935	0.909	0.885	0.864	0.845	0.828	0.813	0.799	0.786	0.774	0.763	0.753	0.743	0.734	0.726	0.718	0.711	0.704	0.697	0.691	0.685	0.679	0.673	0.668	0.663	0.658	0.654	0.649
1/11	0.991	0.969	0.944	0.919	0.897	0.876	0.857	0.840	0.824	0.810	0.797	0.785	0.774	0.763	0.754	0.744	0.736	0.728	0.720	0.713	0.706	0.699	0.693	0.687	0.682	0.676	0.671	0.666	0.661	0.656
1/12	0.993	0.974	0.952	0.928	0.907	0.886	0.868	0.851	0.835	0.821	0.808	0.795	0.784	0.773	0.763	0.754	0.745	0.737	0.729	0.722	0.715	0.708	0.701	0.695	0.689	0.684	0.678	0.673	0.668	0.663
1/13	0.995	0.979	0.958	0.936	0.915	0.896	0.878	0.861	0.845	0.831	0.817	0.805	0.794	0.783	0.773	0.763	0.754	0.746	0.738	0.730	0.723	0.716	0.709	0.703	0.697	0.691	0.685	0.680	0.675	0.670
1/14	0.996	0.983	0.964	0.943	0.923	0.904	0.886	0.870	0.854	0.840	0.827	0.814	0.803	0.792	0.781	0.772	0.762	0.754	0.746	0.738	0.730	0.723	0.716	0.710	0.704	0.698	0.692	0.686	0.681	0.676
1/15	0.997	0.986	0.969	0.950	0.930	0.912	0.895	0.878	0.863	0.849	0.835	0.823	0.811	0.800	0.789	0.780	0.770	0.762	0.753	0.745	0.738	0.730	0.723	0.717	0.710	0.704	0.698	0.692	0.687	0.682
1/16	0.998	0.988	0.973	0.955	0.937	0.919	0.902	0.886	0.871	0.856	0.843	0.831	0.819	0.808	0.797	0.787	0.778	0.769	0.760	0.752	0.744	0.737	0.730	0.723	0.717	0.710	0.704	0.698	0.693	0.687
1/17	0.998	0.990	0.976	0.960	0.943	0.925	0.909	0.893	0.878	0.864	0.851	0.838	0.826	0.815	0.804	0.794	0.785	0.776	0.767	0.759	0.751	0.744	0.736	0.729	0.723	0.716	0.710	0.704	0.698	0.693
1/18	0.999	0.992	0.979	0.964	0.948	0.931	0.915	0.900	0.885	0.871	0.858	0.845	0.833	0.822	0.811	0.801	0.792	0.783	0.774	0.765	0.757	0.750	0.742	0.735	0.729	0.722	0.716	0.710	0.704	0.698
1/19	0.999	0.993	0.982	0.968	0.952	0.937	0.921	0.906	0.891	0.877	0.864	0.852	0.840	0.829	0.818	0.808	0.798	0.789	0.780	0.772	0.764	0.756	0.748	0.741	0.734	0.728	0.721	0.715	0.709	0.703
1/20	0.999	0.994	0.984	0.971	0.957	0.941	0.926	0.911	0.897	0.884	0.870	0.858	0.846	0.835	0.824	0.814	0.804	0.795	0.786	0.778	0.769	0.761	0.754	0.747	0.740	0.733	0.726	0.720	0.714	0.708
1/21	0.999	0.995	0.986	0.974	0.960	0.946	0.931	0.917	0.903	0.889	0.876	0.864	0.852	0.841	0.830	0.820	0.810	0.801	0.792	0.783	0.775	0.767	0.759	0.752	0.745	0.738	0.731	0.725	0.719	0.713
1/22	1.000	0.996	0.988	0.977	0.964	0.950	0.936	0.922	0.908	0.895	0.882	0.870	0.858	0.847	0.836	0.826	0.816	0.806	0.797	0.789	0.780	0.772	0.765	0.757	0.750	0.743	0.736	0.730	0.723	0.717
1/23	1.000	0.997	0.990	0.979	0.967	0.954	0.940	0.926	0.913	0.900	0.887	0.875	0.863	0.852	0.841	0.831	0.821	0.812	0.803	0.794	0.786	0.777	0.770	0.762	0.755	0.748	0.741	0.734	0.728	0.722
1/24	1.000	0.997	0.991	0.981	0.970	0.957	0.944	0.931	0.917	0.905	0.892	0.880	0.869	0.857	0.847	0.836	0.826	0.817	0.808	0.799	0.791	0.782	0.774	0.767	0.759	0.752	0.745	0.739	0.732	0.726
1/25	1.000	0.998	0.992	0.983	0.972	0.960	0.948	0.935	0.922	0.909	0.897	0.885	0.873	0.862	0.852	0.841	0.831	0.822	0.813	0.804	0.795	0.787	0.779	0.772	0.764	0.757	0.750	0.743	0.737	0.730
1/26	1.000	0.998	0.993	0.985	0.975	0.963	0.951	0.938	0.926	0.913	0.901	0.890	0.878	0.867	0.856	0.846	0.836	0.827	0.818	0.809	0.800	0.792	0.784	0.776	0.769	0.761	0.754	0.747	0.741	0.734
1/27	1.000	0.998	0.994	0.986	0.977	0.966	0.954	0.942	0.930	0.918	0.906	0.894	0.883	0.872	0.861	0.851	0.841	0.831	0.822	0.813	0.805	0.796	0.788	0.780	0.773	0.765	0.758	0.751	0.745	0.738
1/28	1.000	0.999	0.995	0.988	0.979	0.968	0.957	0.945	0.933	0.921	0.910	0.898	0.887	0.876	0.865	0.855	0.845	0.836	0.827	0.818	0.809	0.801	0.792	0.785	0.777	0.770	0.762	0.755	0.749	0.742
1/29	1.000	0.999	0.995	0.989	0.981	0.971	0.960	0.948	0.937	0.925	0.913	0.902	0.891	0.880	0.870	0.860	0.850	0.840	0.831	0.822	0.813	0.805	0.797	0.789	0.781	0.774	0.766	0.759	0.753	0.746
1/30	1.000	0.999	0.996	0.990	0.982	0.973	0.962	0.951	0.940	0.929	0.917	0.906	0.895	0.884	0.874	0.864	0.854	0.844	0.835	0.826	0.817	0.809	0.801	0.793	0.785	0.778	0.770	0.763	0.756	0.750
1/31	1.000	0.999	0.996	0.991	0.984	0.975	0.965	0.954	0.943	0.932	0.921	0.910	0.899	0.888	0.878	0.868	0.858	0.848	0.839	0.830	0.821	0.813	0.805	0.797	0.789	0.781	0.774	0.767	0.760	0.753
1/32	1.000	0.999	0.997	0.992	0.985	0.977	0.967	0.957	0.946	0.935	0.924	0.913	0.902	0.892	0.882	0.871	0.862	0.852	0.843	0.834	0.825	0.817	0.808	0.800	0.793	0.785	0.778	0.770	0.763	0.757
1/33	1.000	0.999	0.997	0.993	0.986	0.978	0.969	0.959	0.949	0.938	0.927	0.916	0.906	0.895	0.885	0.875	0.865	0.856	0.847	0.838	0.829	0.820	0.812	0.804	0.796	0.789	0.781	0.774	0.767	0.760
1/34	1.000	1.000	0.998	0.993	0.987	0.980	0.971	0.961	0.951	0.941	0.930	0.920	0.909	0.899	0.889	0.879	0.869	0.860	0.850	0.841	0.833	0.824	0.816	0.808	0.800	0.792	0.785	0.777	0.770	0.763
1/35	1.000	1.000	0.998	0.994	0.988	0.981	0.973	0.963	0.954	0.943	0.933	0.923	0.912	0.902	0.892	0.882	0.873	0.863	0.854	0.845	0.836	0.828	0.819	0.811	0.803	0.796	0.788	0.781	0.774	0.767
1/36	1.000	1.000	0.998	0.995	0.989	0.982	0.974	0.965	0.956	0.946	0.936	0.926	0.915	0.905	0.895	0.885	0.876	0.866	0.857	0.848	0.840	0.831	0.823	0.815	0.807	0.799	0.791	0.784	0.777	0.770
1/37	1.000	1.000	0.998	0.995	0.990	0.984	0.976	0.967	0.958	0.948	0.938	0.928	0.918	0.908	0.898	0.889	0.879	0.870	0.861	0.852	0.843	0.834	0.826	0.818	0.810	0.802	0.795	0.787	0.780	0.773
1/38	1.000	1.000	0.999	0.996	0.991	0.985	0.977	0.969	0.960	0.951	0.941	0.931	0.921	0.911	0.901	0.892	0.882	0.873	0.864	0.855	0.846	0.838	0.829	0.821	0.813	0.805	0.798	0.790	0.783	0.776
1/39	1.000	1.000	0.999	0.996	0.992	0.986	0.979	0.971	0.962	0.953	0.943	0.934	0.924	0.914	0.904	0.895	0.885	0.876	0.867	0.858	0.849	0.841	0.832	0.824	0.816	0.809	0.801	0.793	0.786	0.779
1/40	1.000	1.000	0.999	0.996	0.992	0.987	0.980	0.972	0.964	0.955	0.946	0.936	0.926	0.917	0.907	0.898	0.888	0.879	0.870	0.861	0.852	0.844	0.836	0.827	0.819	0.812	0.804	0.796	0.789	0.782
1/41	1.000	1.000	0.999	0.997	0.993	0.988	0.981	0.974	0.966	0.957	0.948	0.938	0.929	0.919	0.910	0.901	0.891	0.882	0.873	0.864	0.855	0.847	0.839	0.830	0.822	0.815	0.807	0.799	0.792	0.785
1/42	1.000	1.000	0.999	0.997	0.994	0.989	0.982	0.975	0.967	0.959	0.950	0.941	0.931	0.922	0.913	0.903	0.894	0.885	0.876	0.867	0.858	0.850	0.842	0.833	0.825	0.817	0.810	0.802	0.795	0.788
1/43	1.000	1.000	0.999	0.997	0.994	0.989	0.983	0.977	0.969	0.961	0.952	0.943	0.934	0.924	0.915	0.906	0.897	0.888	0.879	0.870	0.861	0.853	0.844	0.836	0.828	0.820	0.813	0.805	0.798	0.790
1/44	1.000	1.000	0.999	0.998	0.995	0.990	0.984	0.978	0.970	0.962	0.954	0.945	0.936	0.927	0.918	0.908	0.899	0.890	0.881	0.873	0.864	0.856	0.847	0.839	0.831	0.823	0.815	0.808	0.800	0.793
1/45	1.000	1.000	0.999	0.998	0.995	0.991	0.985	0.979	0.972	0.964	0.956	0.947	0.938	0.929	0.920	0.911	0.902	0.893	0.884	0.875	0.867	0.858	0.850	0.842	0.834	0.826	0.818	0.810	0.803	0.796
1/46	1.000	1.000	0.999	0.998	0.995	0.991	0.986	0.980	0.973	0.965	0.957	0.949	0.940	0.931	0.922	0.913	0.904	0.895	0.887	0.878	0.869	0.861	0.853	0.844	0.836	0.828	0.821	0.813	0.806	0.798
1/47	1.000	1.000	1.000	0.998	0.996	0.992	0.987	0.981	0.974	0.967	0.959	0.951	0.942	0.933	0.924	0.916	0.907	0.898	0.889	0.881	0.872	0.864	0.855	0.847	0.839	0.831	0.823	0.816	0.808	0.801
1/48	1.000	1.000	1.000	0.998	0.996	0.993	0.988	0.982	0.976	0.968	0.961	0.952	0.944	0.935	0.927	0.918	0.909	0.900	0.892	0.883	0.874	0.866	0.858	0.850	0.842	0.834	0.826	0.818	0.811	0.803
1/49	1.000	1.000	1.000	0.999	0.996	0.993	0.989	0.983	0.977	0.970	0.962	0.954	0.946	0.937	0.929	0.920	0.911	0.903	0.894	0.885	0.877	0.869	0.860	0.852	0.844	0.836	0.828	0.821	0.813	0.806
1/50	1.000	1.000	1.000	0.999	0.997	0.994	0.989	0.984	0.978	0.971	0.964	0.956	0.948	0.939	0.931	0.922	0.914	0.905	0.896	0.888	0.879	0.871	0.863	0.855	0.847	0.839	0.831	0.823	0.816	0.808

Algorithm 2: Standard algorithm for Double Sampling Randomized Smoothing by Exponential General Gaussian (EGG) distributions on real-world datasets

Input: base classifier f, example x, formal variance σ, exponent η, hyperparameter k, significance level α, error bound for certified radius e, heuristic algorithm H, conservative algorithm C, DualBinarySearch algorithm D (Algorithms H, C, D from Li et al. (2022)).

Output: certified radius r

Initialize the smoothing distribution as EGG: $\mathcal{P} \leftarrow \mathcal{G}(\sigma, \eta, k)$
$A_1 \leftarrow$ SamplingUnderNoise $(f, x, \mathcal{P}, \alpha)$
/* A_1 is the Clopper-Pearson lower bound for the sampling result (Clopper & Pearson, 1934; Cohen et al., 2019) */
$T \leftarrow H(A_1)$ /* T is the hyperparameter for the truncated distribution $\mathcal{Q}$ */
Initialize the supplementary distribution as TEGG: $\mathcal{Q} \leftarrow \mathcal{G}_t(\sigma, \eta, k, T)$
$B_1 \leftarrow$ SamplingUnderNoise $(f, x, \mathcal{Q}, \alpha)$
$A, B \leftarrow C(A_1, B_1)$ /* A, B for Problem (57) are determined by algorithm C */
$r_l \leftarrow 0, r_r \leftarrow I$
/* Initialization for the binary search on r, where I is big enough */
while $r_r - r_l > e$ **do**
  $r_m \leftarrow (r_r + r_l)/2$
  $p_m \leftarrow D(A, B)$ /* Problem (57) can be solved by D with given A, B */
  if $p_m > 1/2$ **then**
    $r_l \leftarrow r_m$
  else
    $r_r \leftarrow r_m$
  end
end
$r \leftarrow r_l$
return r

Table 9. The incremental effect with η on different base classifiers.

Dataset	Model	η	Certified Accuracy at r										ACR
			0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	
CIFAR-10	StdAug-GS	2.0((Li et al., 2022), the SOTA)	56.9%	48.0%	40.9%	33.5%	26.4%	19.5%	13.8%	10.6%	7.4%	3.3%	0.437
		8.0	58.1%	50.8%	42.8%	35.9%	30.6%	23.2%	18.1%	13.7%	10.1%	7.3%	0.489
	StdAug-GGS	2.0((Li et al., 2022), the SOTA)	58.3%	51.7%	44.5%	38.1%	29.3%	22.7%	17.7%	13.1%	8.3%	3.8%	0.480
		8.0	58.2%	52.1%	44.4%	38.8%	30.9%	24.1%	19.1%	14.4%	10.8%	6.8%	0.502
	Consistency-GGS	2.0((Li et al., 2022), the SOTA)	52.0%	48.8%	44.7%	42.1%	38.5%	36.0%	32.9%	28.6%	24.0%	19.6%	0.618
		8.0	51.7%	48.6%	44.7%	42.2%	38.8%	36.1%	34.1%	29.9%	26.8%	22.3%	0.650
	SmoothMix-GGS	2.0((Li et al., 2022), the SOTA)	55.8%	52.3%	49.1%	45.3%	41.7%	37.7%	34.6%	30.0%	26.3%	21.2%	0.662
		8.0	55.5%	52.1%	49.1%	45.6%	42.0%	38.2%	35.2%	31.7%	27.8%	24.3%	0.695
ImageNet	StdAug-GGS	2.0((Li et al., 2022), the SOTA)	56.7%	52.8%	49.4%	45.7%	41.5%	38.1%	34.0%	30.6%	25.4%	19.4%	0.654
		8.0	56.7%	52.9%	49.8%	46.3%	42.9%	39.3%	35.8%	32.7%	29.1%	24.6%	0.703

Consistency, SmoothMix}, the training distribution $\in$ {GS (Gaussian), GGS (General Gaussian)}. We use $\sigma = 0.50$ for all the training and sampling distributions.

F. Numerical simulation

We also conduct a numerical simulation to explore the effects of EGG distributions on currently unattainable A, B pairs from Problem (57). We only show simulative experiments for the EGG distribution since we observe great monotonicity for certified robustness (and/or certified radius) w.r.t. η from EGG, which does not occur in ESG. We consider two cases: $B = 1$ and $B < 1$ for B in Problem (57).

Results and analyses. (1) $B = 1$. This is actually an ideal case since it is impossible to train such a base classifier. Here we still present the results, as they demonstrate the theoretical performance of EGG distributions. From Figure 2 (left), we observe there is a monotonically decreasing tendency in the certified radius w.r.t. η, which seems contradictory to the results on real-world datasets. This is partly because the concentration assumption (i.e., $B = 1$) is more friendly to the smaller

Table 10. Numerical simulation for EGG distributions (metric: certified radius).

η	A	0.6				0.7				0.8		
	B	0.6	0.7	0.8	0.9	0.6	0.7	0.8	0.9	0.7	0.8	0.9
0.5		0.188	0.194	0.216	0.273	0.408	0.391	0.407	0.471	0.678	0.632	0.675
1.0		0.218	0.225	0.251	0.320	0.471	0.451	0.470	0.546	0.778	0.726	0.776
2.0		0.234	0.242	0.271	0.346	0.506	0.485	0.505	0.589	0.836	0.779	0.833
4.0		0.243	0.251	0.281	0.360	0.525	0.502	0.524	0.611	0.867	0.807	0.863
8.0		0.247	0.255	0.286	0.367	0.534	0.511	0.533	0.622	0.882	0.821	0.878
16.0		0.249	0.257	0.288	0.370	0.538	0.515	0.537	0.627	0.889	0.827	0.885
32.0		0.250	0.258	0.289	0.371	0.540	0.517	0.539	0.629	0.893	0.830	0.887
64.0		0.250	0.258	0.290	0.371	0.541	0.518	0.539	0.629	0.894	0.831	0.888
Increase (64.0 to 2.0)		6.8%	6.6%	7.0%	7.2%	6.9%	6.8%	6.7%	6.8%	6.9%	6.7%	6.6%

η. In fact, the major mass of the EGG distribution with smaller η gathers near 0. This makes smaller η more sensitive to relaxation of concentration assumption (see Figure 10b), which leads to worse performance on real-world datasets compared to larger η. (2) $B < 1$. As shown in Table 10, the monotonically increasing tendency w.r.t. η continues to exist, manifesting the essential superiority of the large η EGG distribution in the DSRS framework. We also notice that the increase is not endless, since $\eta = 64.0$ shows a marginal increment to that of $\eta = 32.0$ (see also Figure 2 (right)), which may imply some convergence due to the extremely slow growth. We do not show results for larger η due to floating-point limitations and low necessity based on our observation. For all the distributions used for numerical simulation, we set $\sigma = 1.0$ and $k = \frac{d}{2} - 5$ for fair comparison.

F.1. $B = 1$ settings

We need to modify Theorem C.2 slightly to get certified results. We see

$$\begin{aligned} B = 1 &\iff C_g \mathbb{E}_{u \sim \Gamma(\frac{d-2k}{\eta}, 1)} \, \omega_1(u, \nu_1 + C_g \nu_2) \cdot \mathbb{1}_{u \leq \frac{T^\eta}{2\sigma_g^\eta}} = 1 \\ &\iff \nu_1 + C_g \nu_2 \to -\infty, \end{aligned} \tag{101}$$

which makes

$$\omega_2(u) = \Psi_{\frac{d-1}{2}} \left(\frac{T^2 - (t-\rho)^2}{4\rho t} \right). \tag{102}$$

Injecting $\omega_2(u)$ into Theorem C.2, we can compute the certified radius as in general cases. In $B = 1$ experiments, T is set to $\sqrt{2\Lambda_{\frac{d}{2}}(0.5)}$ due to the assumption of the $(1, 0.5, 2)$-concentration property.

F.2. $B < 1$ settings

Theorem C.2 can be directly used in computing the certified radius for this case. The values set for A and B are not completely random as there is an inherent constraint (Li et al., 2022):

$$\begin{cases} \dfrac{B}{C_g} \leq A \leq 1 - \dfrac{1-B}{C_g}, \\ 0 \leq B \leq 1. \end{cases} \tag{103}$$

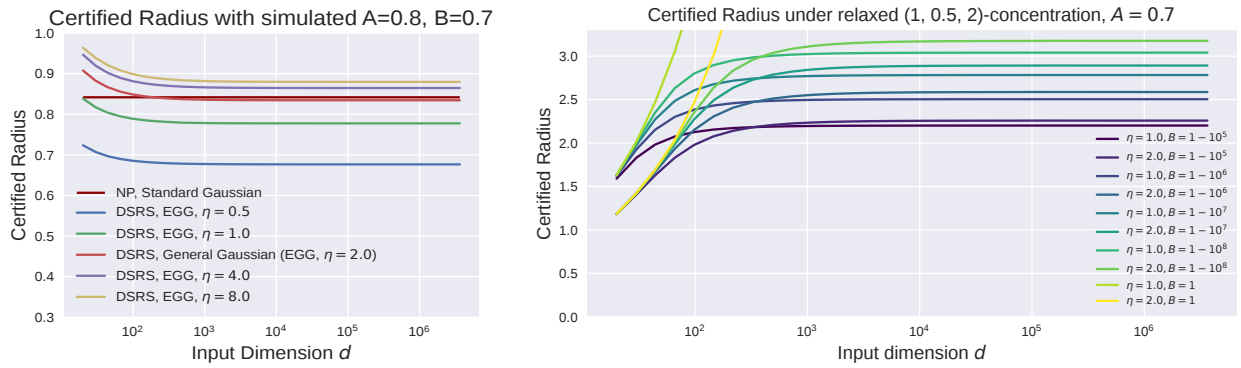
This constraint should also be considered when setting A in case $B = 1$. For each EGG distribution, we set $C_g = 2$, meaning $T = \sigma_g (2\Lambda^{-1}_{\frac{d-2k}{\eta}}(0.5))^{\frac{1}{\eta}}$ for Figure 2 (right) and Table 10.

F.3. The effect of dimension

The results in Table 10 almost do not change on large d. That is, our results that certified radius increases monotonically with η are general for common datasets like CIFAR-10 and ImageNet. See Figure 10a.

F.4. Different sensitivity to relaxation for η

We observe contrary monotonicity for certified radius $w.r.t.$ η under $B = 1$ and $B < 1$. A direct reason is that small η in EGG is more susceptible to relaxation of $B = 1$. We set $T = \sqrt{2\Lambda_{\frac{d}{2}}(0.5)}$ for $B < 1$ curves in this figure for fair



(a) Certified radius vs. d, $A = 0.8, B = 0.7$. (b) Certified radius with $\eta = 1$ vs. $\eta = 2$ under relaxed concentration assumption.

Figure 10. Results for numerical simulation.

Table 11. Maximum certified accuracy w.r.t. σ, Consistency models

Dataset	Method	Certified accuracy at r													
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
CIFAR10	EGG, $\eta = 1.0$	62.1%	50.7%	38.2%	33.1%	24.2%	19.2%	16.7%	14.3%	11.3%	9.2%	6.6%	4.1%	1.4%	0.0%
	EGG, $\eta = 2.0$	62.5%	52.0%	38.5%	34.4%	27.4%	20.6%	17.0%	14.7%	12.7%	10.5%	8.5%	6.3%	3.9%	2.5%
	EGG, $\eta = 4.0$	62.5%	52.2%	39.1%	35.4%	28.3%	21.1%	17.5%	15.3%	13.0%	10.9%	9.2%	7.0%	5.2%	3.1%
	EGG, $\eta = 8.0$	62.5%	52.6%	40.4%	35.3%	28.6%	22.3%	17.6%	15.5%	13.2%	11.3%	9.6%	7.8%	5.8%	3.9%
CIFAR10	ESG, $\eta = 1.0$	62.6%	52.9%	41.6%	35.5%	29.3%	23.7%	17.7%	15.8%	13.7%	11.8%	10.0%	8.8%	6.8%	4.6%
	ESG, $\eta = 2.0$	62.7%	53.0%	41.4%	35.5%	29.4%	24.0%	17.7%	16.0%	13.8%	11.9%	10.1%	8.7%	6.7%	4.4%
	ESG, $\eta = 4.0$	62.7%	52.9%	41.6%	35.5%	29.5%	23.8%	17.8%	15.7%	13.9%	11.9%	10.2%	8.8%	6.7%	4.8%
	ESG, $\eta = 8.0$	62.7%	52.9%	41.7%	35.5%	29.1%	23.8%	17.6%	15.8%	13.6%	11.8%	10.1%	8.3%	6.7%	4.8%

comparison with $B = 1$ ones. Figure 10b demonstrates that the smaller η performs better when $B = 1$, while performs worse when $B < 1$ than the larger η. This indicates that the smaller η suffers more from the relaxation of the concentration assumption than the larger η.

G. Comparison to other work

Li et al. (2022) is one of the cornerstones of this work. To the best of our knowledge, we are the first to generalize the DSRS certification systematically. For NP certification, though many attempts in the community have investigated the interrelationship between the smoothing distribution and certified radius, nobody has shown results as we do. Yang et al. (2020) trained models for each smoothing distribution, while we fixed the base classifier for the convenience of comparison. They also did not fully consider the distributions we use. Kumar et al. (2020) showed similar results to us for ESG in their Figure 5, but they only showed sampling results due to the lack of computing method, and they also trained base models for each distribution like Yang et al. (2020). Zhang et al. (2020) showed results for $\eta = 2$ EGG distribution under a fixed-model setting, but they didn't consider $\eta \neq 2$ cases.

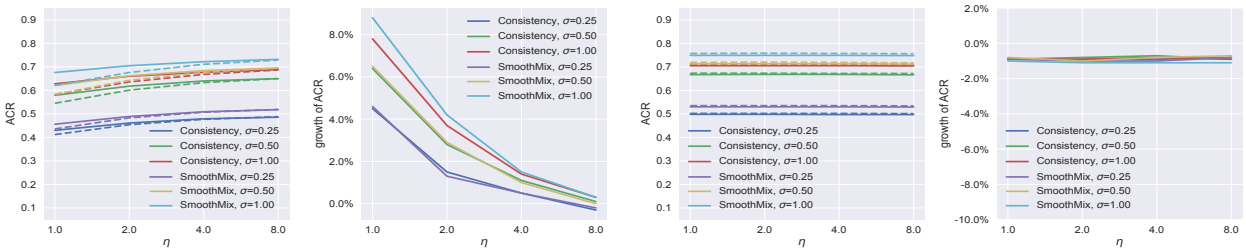
H. Supplementary for experimental results

H.1. Certified radius at r for Consistency (Jeong & Shin, 2020) and SmoothMix (Jeong et al., 2021) models, maximum results

Table 11 and Table 12 show the experimental results for certified accuracy at radius r. Each data is the **maximum** one among base classifiers with $\sigma \in \{0.25, 0.50, 1.00\}$. All the base classifiers in both tables are trained under the General Gaussian distribution, by Consistency (Jeong & Shin, 2020) and SmoothMix (Jeong et al., 2021) respectively. From the tables, we can see the rule observed on classifiers augmented standardly continues to exist, that certified accuracy increases monotonically with the η of EGG, and stays almost constant with the η of ESG.

Table 12. Maximum certified accuracy w.r.t. σ, SmoothMix models

Dataset	Method	Certified accuracy at r													
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
CIFAR10	EGG, $\eta = 1.0$	63.7%	53.8%	40.9%	34.3%	26.6%	21.1%	17.0%	14.2%	10.5%	7.7%	4.0%	1.5%	0.1%	0.0%
	EGG, $\eta = 2.0$	64.5%	55.0%	41.7%	35.6%	28.9%	21.3%	18.0%	15.2%	12.3%	9.7%	6.4%	3.7%	1.4%	0.4%
	EGG, $\eta = 4.0$	64.4%	55.5%	43.0%	35.9%	29.5%	23.4%	18.3%	15.8%	12.8%	10.2%	7.7%	4.6%	2.1%	0.9%
	EGG, $\eta = 8.0$	64.7%	55.7%	43.9%	36.2%	30.1%	24.3%	18.6%	15.8%	13.2%	10.6%	8.0%	5.4%	2.7%	1.3%
CIFAR10	ESG, $\eta = 1.0$	64.6%	56.5%	46.5%	36.6%	31.3%	25.6%	19.1%	16.2%	13.4%	11.4%	8.9%	6.3%	4.0%	1.8%
	ESG, $\eta = 2.0$	64.7%	56.3%	45.8%	36.6%	31.2%	25.6%	18.9%	16.3%	13.4%	11.6%	8.8%	6.2%	4.0%	1.7%
	ESG, $\eta = 4.0$	64.6%	56.3%	46.0%	36.5%	31.0%	25.6%	19.0%	16.2%	13.6%	11.5%	9.2%	6.0%	3.9%	1.8%
	ESG, $\eta = 8.0$	64.6%	56.1%	46.5%	36.4%	31.1%	25.5%	18.9%	16.3%	13.5%	11.3%	9.0%	6.2%	4.1%	1.6%



(a) (b) (c) (d)

Figure 11. ACR results on Consistency and SmoothMix models. **(a).** ACR monotonically increases with η in EGG. **(b).** The ACR growth gain from DSRS relative to NP shrinks with η in EGG. **(c).** ACR stays almost constant in ESG. **(d).** The ACR growth gain from DSRS remains almost constant in ESG. For (a) and (c), solid lines represent results from DSRS, and dotted lines represent results from NP.

H.2. Full experimental results for certifications

We show full experimental results in this section, where the maximum results for certified accuracy in Table 2, Table 3, Table 11 and Table 12 originated from. All the base classifiers in this section are trained under the EGG distribution with $\eta = 2$ (the General Gaussian distribution used in DSRS (Li et al., 2022)). Our base classifiers are only affected by the dataset and the formal variance. For example, in Table 13, all data under $\sigma = 0.25$ use the same base classifier, no matter what η is. In addition, for $\sigma = 0.25$, the base classifiers for EGG and ESG are the same, which guarantees fair comparisons between distinctive distributions for certification.

Table 13. Full experimental results for certified accuracy, standard augmentation, EGG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	0.25	NP	41.4%	0.5%												
		DSRS	52.6%	7.9%												
		(Growth)	11.2%	7.4%												
	0.5	NP	51.6%	14.5%	0.1%											
		DSRS	55.5%	29.5%												
		(Growth)	3.9%	15.0%												
	1.0	NP	55.6%	29.1%	5.5%											
		DSRS	56.3%	36.5%	9.3%											
		(Growth)	0.7%	7.4%	3.8%											
	2.0	NP	56.2%	35.7%	13.4%											
		DSRS	56.7%	38.4%	16.9%											
		(Growth)	0.5%	2.7%	3.5%											
	4.0	NP	57.3%	38.5%	18.5%											
		DSRS	57.5%	39.3%	20.0%											
		(Growth)	0.2%	0.8%	1.5%											
	8.0	NP	57.5%	39.2%	21.2%											
		DSRS	57.6%	40.1%	22.0%											
		(Growth)	0.1%	0.9%	0.8%											
0.50	0.25	NP	50.9%	23.7%	3.9%	0.1%										
		DSRS	54.2%	37.6%	16.4%	2.0%										
		(Growth)	3.3%	13.9%	12.5%	1.9%										
	0.5	NP	53.0%	34.6%	16.6%	4.5%	0.3%									
		DSRS	53.6%	40.4%	25.2%	12.9%	2.5%									
		(Growth)	0.6%	5.8%	8.6%	8.4%	2.2%									
	1.0	NP	53.7%	39.8%	23.1%	12.4%	3.6%	0.7%								
		DSRS	54.1%	41.7%	28.2%	17.3%	9.2%	1.8%								
		(Growth)	0.4%	1.9%	5.1%	4.9%	5.6%	1.1%								
	2.0	NP	53.8%	41.2%	27.9%	17.0%	8.9%	2.9%	0.1%							
		DSRS	54.0%	42.4%	29.3%	19.4%	11.6%	3.8%	0.6%							
		(Growth)	0.2%	1.2%	1.4%	2.4%	2.7%	0.9%	0.5%							
	4.0	NP	53.9%	42.2%	29.7%	19.0%	11.3%	4.7%	1.5%							
		DSRS	54.0%	42.5%	30.0%	20.2%	12.8%	5.8%	1.5%							
		(Growth)	0.1%	0.3%	0.3%	1.2%	1.5%	1.1%	0.0%							
	8.0	NP	54.0%	42.6%	30.0%	20.0%	12.7%	6.8%	2.3%							
		DSRS	54.2%	42.5%	30.9%	20.6%	13.4%	6.8%	1.8%							
		(Growth)	0.2%	-0.1%	0.9%	0.6%	0.7%	0.0%	-0.5%							
1.00	0.25	NP	40.6%	28.7%	18.0%	9.0%	3.1%	0.4%								
		DSRS	41.8%	32.6%	23.5%	16.5%	9.4%	4.5%	0.5%	0.1%						
		(Growth)	1.2%	3.9%	5.5%	7.5%	6.3%	4.1%								
	0.5	NP	40.9%	31.8%	21.8%	15.7%	9.1%	4.9%	1.6%	0.2%	0.1%					
		DSRS	41.1%	33.6%	24.7%	19.1%	13.4%	8.5%	5.5%	2.0%	0.4%	0.1%				
		(Growth)	0.2%	1.8%	2.9%	3.4%	4.3%	3.6%	3.9%	1.8%	0.3%					
	1.0	NP	40.4%	32.4%	24.0%	17.7%	12.6%	8.5%	5.0%	2.2%	0.5%	0.1%	0.1%			
		DSRS	40.2%	33.2%	25.5%	20.0%	15.1%	10.5%	7.1%	4.2%	1.9%	0.9%	0.1%			
		(Growth)	-0.2%	0.8%	1.5%	2.3%	2.5%	2.0%	2.1%	2.0%	1.4%	0.8%	0.0%			
	2.0	NP	40.2%	32.6%	24.6%	18.9%	15.0%	10.1%	7.4%	4.0%	2.0%	0.7%	0.1%	0.1%		
		DSRS	40.2%	32.8%	25.5%	20.2%	15.7%	11.5%	8.0%	5.5%	2.6%	1.5%	0.6%	0.1%		
		(Growth)	0.0%	0.2%	0.9%	1.3%	0.7%	1.4%	0.6%	1.5%	0.6%	0.8%	0.5%	0.0%		
	4.0	NP	40.0%	32.7%	25.2%	19.7%	15.5%	11.4%	8.2%	5.3%	3.0%	1.5%	0.6%	0.1%	0.1%	
		DSRS	39.6%	32.7%	25.7%	20.2%	15.9%	12.2%	8.5%	6.5%	3.4%	1.8%	0.9%	0.4%		
		(Growth)	-0.4%	0.0%	0.5%	0.5%	0.4%	0.8%	0.3%	1.2%	0.4%	0.3%	0.3%	0.3%		
	8.0	NP	39.7%	32.5%	25.5%	20.2%	15.7%	12.1%	8.5%	6.3%	3.4%	2.0%	0.9%	0.5%	0.2%	
		DSRS	39.5%	32.6%	25.5%	20.2%	15.8%	12.3%	8.6%	6.6%	3.7%	2.1%	1.1%	0.5%	0.2%	
		(Growth)	-0.2%	0.1%	0.0%	0.0%	0.1%	0.2%	0.1%	0.3%	0.3%	0.1%	0.2%	0.0%	0.0%	

Table 14. Full experimental results for certified accuracy, standard augmentation, EGG, ImageNet

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	0.25	NP	9.9%													
		DSRS	45.8%													
		(Growth)	35.9%													
	0.5	NP	44.6%	0.6%												
		DSRS	54.9%	9.7%												
		(Growth)	10.3%	9.1%												
	1.0	NP	54.0%	21.1%	0.6%											
		DSRS	57.0%	39.3%	1.4%											
		(Growth)	3.0%	18.2%	0.8%											
	2.0	NP	57.1%	41.7%	17.6%											
		DSRS	58.4%	47.9%	24.1%											
		(Growth)	1.3%	6.2%	6.5%											
	4.0	NP	58.4%	48.0%	33.6%											
		DSRS	58.7%	49.9%	36.2%											
		(Growth)	0.3%	1.9%	2.6%											
	8.0	NP	58.7%	50.0%	37.8%											
		DSRS	59.1%	50.8%	38.7%											
		(Growth)	0.4%	0.8%	0.9%											
0.50	0.25	NP	49.7%	26.7%	0.1%											
		DSRS	53.8%	41.4%	14.8%											
		(Growth)	4.1%	14.7%	14.7%											
	0.5	NP	52.1%	40.6%	26.4%	8.3%	0.1%									
		DSRS	54.4%	46.3%	36.4%	22.5%	2.9%									
		(Growth)	2.3%	5.7%	10.0%	14.2%	2.8%									
	1.0	NP	52.8%	44.8%	35.7%	26.2%	16.1%	6.3%								
		DSRS	54.2%	47.8%	39.9%	32.8%	22.9%	8.9%								
		(Growth)	1.4%	3.0%	4.2%	6.6%	6.8%	2.6%								
	2.0	NP	53.4%	47.0%	39.4%	33.3%	24.5%	17.4%	8.4%							
		DSRS	53.8%	48.5%	41.5%	35.2%	28.9%	19.4%	11.3%							
		(Growth)	0.4%	1.5%	2.1%	1.9%	4.4%	2.0%	2.9%							
	4.0	NP	53.3%	47.7%	41.3%	35.2%	29.3%	21.3%	14.0%							
		DSRS	53.6%	48.6%	42.6%	36.4%	31.0%	22.8%	14.4%							
		(Growth)	0.3%	0.9%	1.3%	1.2%	1.7%	1.5%	0.4%							
	8.0	NP	53.3%	48.3%	42.2%	36.5%	31.1%	24.0%	16.9%							
		DSRS	53.6%	48.8%	42.9%	36.8%	31.8%	24.6%	16.5%							
		(Growth)	0.3%	0.5%	0.7%	0.3%	0.7%	0.6%	-0.4%							
1.00	0.25	NP	38.8%	29.8%	15.7%	3.3%										
		DSRS	41.0%	35.3%	28.4%	20.1%	7.1%	0.8%								
		(Growth)	2.2%	5.5%	12.7%	16.8%										
	0.5	NP	40.3%	34.1%	26.4%	19.0%	10.7%	4.1%	1.4%	0.1%						
		DSRS	40.9%	37.0%	31.6%	26.3%	22.1%	15.2%	8.7%	3.1%	0.8%					
		(Growth)	0.6%	2.9%	5.2%	7.3%	11.4%	11.1%	7.3%	3.0%						
	1.0	NP	41.7%	36.4%	30.7%	25.5%	20.9%	15.3%	10.5%	6.3%	3.3%	1.7%	0.7%	0.2%		
		DSRS	42.0%	37.9%	33.4%	29.3%	24.9%	22.0%	18.5%	13.1%	9.2%	5.0%	2.1%	0.5%		
		(Growth)	0.3%	1.5%	2.7%	3.8%	4.0%	6.7%	8.0%	6.8%	5.9%	3.3%	1.4%	0.3%		
	2.0	NP	42.5%	37.2%	32.8%	29.2%	24.7%	21.4%	17.4%	13.8%	10.1%	7.8%	5.5%	3.3%	2.2%	1.1%
		DSRS	42.9%	38.2%	34.3%	30.2%	26.8%	23.3%	21.3%	18.8%	14.1%	11.1%	8.9%	6.1%	2.2%	1.4%
		(Growth)	0.4%	1.0%	1.5%	1.0%	2.1%	1.9%	3.9%	5.0%	4.0%	3.3%	3.4%	2.8%	0.0%	0.3%
	4.0	NP	42.2%	38.0%	33.8%	30.8%	26.1%	23.3%	21.1%	18.2%	13.9%	11.2%	9.3%	8.1%	6.0%	4.2%
		DSRS	42.5%	38.5%	34.6%	31.3%	27.5%	23.9%	22.3%	20.2%	17.3%	13.2%	10.7%	9.2%	6.8%	4.0%
		(Growth)	0.3%	0.5%	0.8%	0.5%	1.4%	0.6%	1.2%	2.0%	3.4%	2.0%	1.4%	1.1%	0.8%	-0.2%
	8.0	NP	42.4%	38.1%	34.5%	31.1%	26.9%	24.0%	22.2%	19.9%	16.8%	12.8%	11.0%	9.5%	8.0%	6.1%
		DSRS	42.4%	38.4%	35.0%	31.4%	28.0%	24.4%	22.6%	20.7%	18.9%	14.5%	11.7%	10.1%	8.6%	5.2%
		(Growth)	0.0%	0.3%	0.5%	0.3%	1.1%	0.4%	0.4%	0.8%	2.1%	1.7%	0.7%	0.6%	0.6%	-0.9%

Table 15. Full experimental results for certified accuracy, standard augmentation, ESG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	57.8%	40.7%	25.6%											
		DSRS	57.6%	40.7%	25.1%											
		(Growth)	-0.2%	0.0%	-0.5%											
	2.0	NP	57.8%	40.8%	25.7%											
		DSRS	57.6%	40.6%	25.1%											
		(Growth)	-0.2%	-0.2%	-0.6%											
	4.0	NP	57.9%	40.9%	26.0%											
		DSRS	57.6%	40.6%	25.0%											
		(Growth)	-0.3%	-0.3%	-1.0%											
	8.0	NP	57.9%	40.8%	25.6%											
		DSRS	57.8%	40.6%	24.9%											
		(Growth)	-0.1%	-0.2%	-0.7%											
0.50	1.0	NP	54.3%	42.7%	31.7%	21.8%	14.0%	8.7%	3.5%							
		DSRS	54.2%	42.6%	31.3%	21.5%	14.1%	8.3%	2.7%							
		(Growth)	-0.1%	-0.1%	-0.4%	-0.3%	0.1%	-0.4%	-0.8%							
	2.0	NP	54.3%	42.6%	31.6%	21.7%	14.0%	8.5%	3.6%							
		DSRS	54.3%	42.6%	31.6%	21.5%	13.9%	8.7%	2.8%							
		(Growth)	0.0%	0.0%	0.0%	-0.2%	-0.1%	0.2%	-0.8%							
	4.0	NP	54.3%	42.7%	31.5%	21.6%	14.3%	8.6%	3.6%							
		DSRS	54.3%	42.6%	31.3%	21.5%	13.9%	8.2%	3.0%							
		(Growth)	0.0%	-0.1%	-0.2%	-0.1%	-0.4%	-0.4%	-0.6%							
	8.0	NP	54.3%	42.6%	31.7%	21.7%	14.1%	8.7%	3.5%							
		DSRS	54.4%	42.6%	31.6%	21.6%	14.0%	8.1%	2.4%							
		(Growth)	0.1%	0.0%	-0.1%	-0.1%	-0.1%	-0.6%	-1.1%							
1.00	1.0	NP	39.6%	32.6%	26.0%	20.5%	15.9%	13.0%	9.2%	7.0%	4.5%	2.5%	1.5%	0.9%	0.5%	0.2%
		DSRS	39.6%	32.5%	25.7%	20.4%	15.8%	12.8%	8.6%	6.8%	4.3%	2.3%	1.3%	0.8%	0.3%	0.1%
		(Growth)	0.0%	-0.1%	-0.3%	-0.1%	-0.1%	-0.2%	-0.6%	-0.2%	-0.2%	-0.2%	-0.2%	-0.1%	-0.2%	-0.1%
	2.0	NP	39.6%	32.6%	25.9%	20.4%	15.9%	13.0%	9.2%	7.0%	4.6%	2.5%	1.3%	0.8%	0.3%	0.2%
		DSRS	39.5%	32.6%	25.8%	20.4%	15.8%	12.7%	8.8%	6.8%	4.5%	2.4%	1.3%	0.7%	0.2%	0.2%
		(Growth)	-0.1%	0.0%	-0.1%	0.0%	-0.1%	-0.3%	-0.4%	-0.2%	-0.1%	-0.1%	0.0%	-0.1%	-0.1%	0.0%
	4.0	NP	39.7%	32.6%	25.9%	20.4%	15.9%	12.9%	9.0%	7.0%	4.6%	2.5%	1.5%	0.8%	0.3%	0.2%
		DSRS	39.5%	32.6%	25.8%	20.3%	15.9%	12.9%	8.6%	6.9%	4.3%	2.4%	1.3%	0.8%	0.2%	0.1%
		(Growth)	-0.2%	0.0%	-0.1%	-0.1%	0.0%	0.0%	-0.4%	-0.1%	-0.3%	-0.1%	-0.2%	0.0%	-0.1%	-0.1%
	8.0	NP	39.6%	32.6%	25.8%	20.3%	15.9%	12.9%	9.2%	7.0%	4.5%	2.5%	1.3%	0.8%	0.3%	0.2%
		DSRS	39.7%	32.4%	25.8%	20.4%	15.9%	12.9%	8.9%	6.7%	4.2%	2.4%	1.3%	0.9%	0.2%	0.1%
		(Growth)	0.1%	-0.2%	0.0%	0.1%	0.0%	0.0%	-0.3%	-0.3%	-0.3%	-0.1%	0.0%	0.1%	-0.1%	-0.1%

Table 16. Full experimental results for certified accuracy, standard augmentation, ESG, ImageNet

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	59.6%	51.6%	42.2%											
		DSRS	59.6%	51.5%	41.6%											
		(Growth)	0.0%	-0.1%	-0.6%											
	2.0	NP	59.6%	51.7%	41.9%											
		DSRS	59.6%	51.6%	41.8%											
		(Growth)	0.0%	-0.1%	-0.1%											
	4.0	NP	59.6%	51.7%	42.0%											
		DSRS	59.6%	51.5%	41.9%											
		(Growth)	0.0%	-0.2%	-0.1%											
	8.0	NP	59.6%	51.6%	42.2%											
		DSRS	59.6%	51.5%	41.5%											
		(Growth)	0.0%	-0.1%	-0.7%											
0.50	1.0	NP	53.6%	49.3%	43.2%	38.2%	33.0%	27.2%	21.0%							
		DSRS	53.6%	49.1%	43.2%	37.9%	33.0%	26.8%	19.1%							
		(Growth)	0.0%	-0.2%	0.0%	-0.3%	0.0%	-0.4%	-1.9%							
	2.0	NP	53.7%	49.2%	43.2%	38.1%	33.1%	27.3%	20.7%							
		DSRS	53.6%	49.2%	43.1%	38.0%	32.9%	26.9%	19.2%							
		(Growth)	-0.1%	0.0%	-0.1%	-0.1%	-0.2%	-0.4%	-1.5%							
	4.0	NP	53.6%	49.2%	43.2%	38.1%	33.0%	27.5%	20.8%							
		DSRS	53.6%	49.2%	43.2%	38.0%	32.9%	27.2%	19.4%							
		(Growth)	0.0%	0.0%	0.0%	-0.1%	-0.1%	-0.3%	-1.4%							
	8.0	NP	53.7%	49.2%	43.2%	38.1%	33.1%	27.4%	21.0%							
		DSRS	53.6%	49.1%	43.2%	38.0%	33.0%	26.8%	19.4%							
		(Growth)	-0.1%	-0.1%	0.0%	-0.1%	-0.1%	-0.6%	-1.6%							
1.00	1.0	NP	42.7%	39.1%	35.4%	32.0%	29.5%	25.3%	23.1%	21.6%	20.0%	17.6%	13.9%	11.7%	10.5%	9.1%
		DSRS	42.6%	38.8%	35.3%	31.9%	28.9%	25.3%	23.1%	21.5%	19.9%	17.4%	13.8%	11.5%	10.3%	7.7%
		(Growth)	-0.1%	-0.3%	-0.1%	-0.1%	-0.6%	0.0%	0.0%	-0.1%	-0.1%	-0.2%	-0.1%	-0.2%	-0.2%	-1.4%
	2.0	NP	42.6%	39.1%	35.3%	32.1%	29.2%	25.3%	23.2%	21.6%	20.0%	17.6%	14.1%	11.7%	10.5%	9.1%
		DSRS	42.5%	39.0%	35.2%	31.7%	29.0%	25.2%	23.1%	21.5%	19.7%	17.4%	13.6%	11.4%	10.1%	8.3%
		(Growth)	-0.1%	-0.1%	-0.1%	-0.4%	-0.2%	-0.1%	-0.1%	-0.1%	-0.3%	-0.2%	-0.5%	-0.3%	-0.4%	-0.8%
	4.0	NP	42.6%	39.1%	35.4%	32.0%	29.1%	25.3%	23.1%	21.6%	19.9%	17.9%	14.5%	11.7%	10.5%	9.1%
		DSRS	42.5%	38.8%	35.2%	31.9%	29.3%	25.3%	23.1%	21.6%	19.9%	17.2%	13.6%	11.4%	10.2%	8.0%
		(Growth)	-0.1%	-0.3%	-0.2%	-0.1%	0.2%	0.0%	0.0%	0.0%	0.0%	-0.7%	-0.9%	-0.3%	-0.3%	-1.1%
	8.0	NP	42.7%	39.0%	35.3%	31.9%	29.4%	25.3%	23.1%	21.7%	20.0%	17.5%	13.9%	11.5%	10.4%	9.3%
		DSRS	42.5%	38.8%	35.3%	32.0%	29.0%	25.2%	23.1%	21.6%	19.7%	17.3%	13.6%	11.5%	10.1%	8.4%
		(Growth)	-0.2%	-0.2%	0.0%	0.1%	-0.4%	-0.1%	0.0%	-0.1%	-0.3%	-0.2%	-0.3%	0.0%	-0.3%	-0.9%

Table 17. Full experimental results for certified accuracy, Consistency, EGG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	61.4%	47.0%	25.2%											
		DSRS	62.1%	50.7%	30.1%											
		(Growth)	0.7%	3.7%	4.9%											
	2.0	NP	61.8%	50.7%	35.1%											
		DSRS	62.5%	52.0%	37.2%											
		(Growth)	0.7%	1.3%	2.1%											
	4.0	NP	62.3%	51.7%	38.2%											
		DSRS	62.5%	52.2%	39.1%											
		(Growth)	0.2%	0.5%	0.9%											
	8.0	NP	62.5%	52.2%	40.2%											
		DSRS	62.5%	52.6%	40.4%											
		(Growth)	0.0%	0.4%	0.2%											
0.50	1.0	NP	49.2%	43.1%	36.3%	28.7%	19.6%	11.6%								
		DSRS	49.5%	43.7%	38.2%	33.1%	24.2%	15.4%								
		(Growth)	0.3%	0.6%	1.9%	4.4%	4.6%	3.8%								
	2.0	NP	49.3%	43.8%	37.8%	32.3%	23.5%	18.0%	9.6%							
		DSRS	49.4%	44.1%	38.5%	34.4%	27.4%	19.6%	11.5%							
		(Growth)	0.1%	0.3%	0.7%	2.1%	3.9%	1.6%	1.9%							
	4.0	NP	49.3%	44.0%	38.3%	34.1%	27.4%	20.4%	14.6%							
		DSRS	49.3%	44.0%	38.7%	35.4%	28.3%	21.1%	14.1%							
		(Growth)	0.0%	0.0%	0.4%	1.3%	0.9%	0.7%	-0.5%							
	8.0	NP	49.3%	44.1%	38.7%	35.0%	28.4%	21.9%	15.8%							
		DSRS	49.3%	44.1%	38.8%	35.3%	28.6%	22.3%	15.2%							
		(Growth)	0.0%	0.0%	0.1%	0.3%	0.2%	0.4%	-0.6%							
1.00	1.0	NP	37.2%	32.5%	29.4%	25.2%	21.8%	17.6%	14.4%	11.7%	8.8%	6.3%	4.5%	2.6%	1.8%	
		DSRS	37.4%	33.0%	29.5%	26.2%	22.8%	19.2%	16.7%	14.3%	11.3%	9.2%	6.6%	4.1%	1.4%	
		(Growth)	0.2%	0.5%	0.1%	1.0%	1.0%	1.6%	2.3%	2.6%	2.5%	2.9%	2.1%	1.5%	-0.4%	
	2.0	NP	37.2%	32.6%	29.7%	25.9%	22.4%	19.0%	16.3%	13.9%	11.3%	8.9%	7.2%	5.1%	3.5%	2.2%
		DSRS	37.1%	32.3%	29.8%	26.5%	23.0%	20.6%	17.0%	14.7%	12.7%	10.5%	8.5%	6.3%	3.9%	2.5%
		(Growth)	-0.1%	-0.3%	0.1%	0.6%	0.6%	1.6%	0.7%	0.8%	1.4%	1.6%	1.3%	1.2%	0.4%	0.3%
	4.0	NP	37.1%	32.5%	29.8%	26.2%	22.7%	20.3%	16.9%	14.9%	12.4%	10.6%	8.7%	6.7%	5.1%	3.2%
		DSRS	37.0%	32.4%	29.8%	26.7%	23.1%	20.9%	17.5%	15.3%	13.0%	10.9%	9.2%	7.0%	5.2%	3.1%
		(Growth)	-0.1%	-0.1%	0.0%	0.5%	0.4%	0.6%	0.6%	0.4%	0.6%	0.3%	0.5%	0.3%	0.1%	-0.1%
	8.0	NP	37.1%	32.5%	29.9%	26.4%	23.0%	20.7%	17.2%	15.2%	13.2%	10.9%	9.6%	7.5%	5.9%	4.2%
		DSRS	36.7%	32.5%	29.8%	26.6%	23.2%	20.9%	17.6%	15.5%	13.2%	11.3%	9.6%	7.8%	5.8%	3.9%
		(Growth)	-0.4%	0.0%	-0.1%	0.2%	0.2%	0.2%	0.4%	0.3%	0.0%	0.4%	0.0%	0.3%	-0.1%	-0.3%

Table 18. Full experimental results for certified accuracy, Consistency, ESG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	62.7%	52.9%	41.8%											
		DSRS	62.6%	52.9%	41.6%											
		(Growth)	-0.1%	0.0%	-0.2%											
	2.0	NP	62.7%	53.0%	42.1%											
		DSRS	62.7%	53.0%	41.4%											
		(Growth)	0.0%	0.0%	-0.7%											
	4.0	NP	62.7%	53.0%	42.0%											
		DSRS	62.7%	52.9%	41.6%											
		(Growth)	0.0%	-0.1%	-0.4%											
	8.0	NP	62.7%	53.0%	42.0%											
		DSRS	62.7%	52.9%	41.7%											
		(Growth)	0.0%	-0.1%	-0.3%											
0.50	1.0	NP	49.3%	44.1%	39.2%	35.5%	29.7%	24.3%	18.7%							
		DSRS	49.3%	44.1%	38.9%	35.5%	29.3%	23.7%	16.9%							
		(Growth)	0.0%	0.0%	-0.3%	0.0%	-0.4%	-0.6%	-1.8%							
	2.0	NP	49.3%	44.1%	39.2%	35.5%	29.7%	24.1%	18.7%							
		DSRS	49.3%	44.1%	38.9%	35.5%	29.4%	24.0%	17.1%							
		(Growth)	0.0%	0.0%	-0.3%	0.0%	-0.3%	-0.1%	-1.6%							
	4.0	NP	49.3%	44.1%	39.1%	35.5%	29.9%	24.1%	18.6%							
		DSRS	49.3%	44.0%	39.0%	35.5%	29.5%	23.8%	17.6%							
		(Growth)	0.0%	-0.1%	-0.1%	0.0%	-0.4%	-0.3%	-1.0%							
	8.0	NP	49.3%	44.1%	39.3%	35.5%	29.6%	24.2%	19.0%							
		DSRS	49.3%	44.1%	38.9%	35.5%	29.1%	23.8%	17.3%							
		(Growth)	0.0%	0.0%	-0.4%	0.0%	-0.5%	-0.4%	-1.7%							
1.00	1.0	NP	36.7%	32.6%	30.0%	26.9%	23.4%	21.0%	18.0%	16.0%	14.0%	12.2%	10.4%	8.8%	7.0%	5.4%
		DSRS	36.5%	32.3%	29.8%	27.0%	23.2%	20.9%	17.7%	15.8%	13.7%	11.8%	10.0%	8.8%	6.8%	4.6%
		(Growth)	-0.2%	-0.3%	-0.2%	0.1%	-0.2%	-0.1%	-0.3%	-0.2%	-0.3%	-0.4%	-0.4%	0.0%	-0.2%	-0.8%
	2.0	NP	36.8%	32.5%	29.9%	27.1%	23.3%	21.0%	18.1%	16.0%	14.0%	12.2%	10.3%	8.8%	7.1%	5.5%
		DSRS	36.6%	32.3%	29.9%	26.8%	23.3%	21.0%	17.7%	16.0%	13.8%	11.9%	10.1%	8.7%	6.7%	4.4%
		(Growth)	-0.2%	-0.2%	0.0%	-0.3%	0.0%	0.0%	-0.4%	0.0%	-0.2%	-0.3%	-0.2%	-0.1%	-0.4%	-1.1%
	4.0	NP	36.6%	32.5%	29.9%	27.0%	23.2%	21.0%	18.0%	16.1%	14.0%	12.2%	10.3%	9.0%	7.2%	5.3%
		DSRS	36.6%	32.3%	29.8%	26.9%	23.3%	20.9%	17.8%	15.7%	13.9%	11.9%	10.2%	8.8%	6.7%	4.8%
		(Growth)	0.0%	-0.2%	-0.1%	-0.1%	0.1%	-0.1%	-0.2%	-0.4%	-0.1%	-0.3%	-0.1%	-0.2%	-0.5%	-0.5%
	8.0	NP	36.7%	32.6%	29.8%	27.0%	23.3%	21.0%	17.9%	16.1%	13.8%	12.2%	10.3%	9.0%	7.1%	5.1%
		DSRS	36.6%	32.1%	29.8%	26.9%	23.2%	20.9%	17.6%	15.8%	13.6%	11.8%	10.1%	8.3%	6.7%	4.8%
		(Growth)	-0.1%	-0.5%	0.0%	-0.1%	-0.1%	-0.1%	-0.3%	-0.3%	-0.2%	-0.4%	-0.2%	-0.7%	-0.4%	-0.3%

Table 19. Full experimental results for certified accuracy, SmoothMix, EGG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	63.3%	49.8%	27.4%											
		DSRS	63.7%	53.8%	32.4%											
		(Growth)	0.4%	4.0%	5.0%											
	2.0	NP	63.8%	53.3%	38.1%											
		DSRS	64.5%	55.0%	40.8%											
		(Growth)	0.7%	1.7%	2.7%											
	4.0	NP	64.2%	55.1%	42.0%											
		DSRS	64.4%	55.5%	43.0%											
		(Growth)	0.2%	0.4%	1.0%											
	8.0	NP	64.6%	55.5%	43.7%											
		DSRS	64.7%	55.7%	43.9%											
		(Growth)	0.1%	0.2%	0.2%											
0.50	1.0	NP	53.0%	46.7%	38.8%	30.1%	22.2%	12.7%								
		DSRS	53.3%	47.7%	40.9%	34.3%	26.6%	16.5%								
		(Growth)	0.3%	1.0%	2.1%	4.2%	4.4%	3.8%								
	2.0	NP	53.2%	47.5%	40.3%	34.0%	26.6%	19.4%	9.6%							
		DSRS	53.3%	48.1%	41.7%	35.6%	28.9%	21.2%	11.4%							
		(Growth)	0.1%	0.6%	1.4%	1.6%	2.3%	1.8%	1.8%							
	4.0	NP	53.3%	48.0%	41.3%	35.1%	29.0%	22.6%	15.5%							
		DSRS	53.3%	48.1%	41.9%	35.9%	29.5%	23.4%	14.3%							
		(Growth)	0.0%	0.1%	0.6%	0.8%	0.5%	0.8%	-1.2%							
	8.0	NP	53.3%	48.2%	41.6%	35.8%	29.5%	23.9%	17.5%							
		DSRS	53.3%	48.3%	42.0%	36.2%	30.1%	24.3%	15.6%							
		(Growth)	0.0%	0.1%	0.4%	0.4%	0.6%	0.4%	-1.9%							
1.00	1.0	NP	43.3%	39.3%	33.2%	27.9%	22.7%	18.2%	15.1%	10.9%	7.4%	3.5%	1.6%	0.8%	0.1%	
		DSRS	43.5%	39.6%	34.5%	29.1%	24.8%	21.1%	17.0%	14.2%	10.5%	7.7%	4.0%	1.5%	0.1%	
		(Growth)	0.2%	0.3%	1.3%	1.2%	2.1%	2.9%	1.9%	3.3%	3.1%	4.2%	2.4%	0.7%	0.0%	
	2.0	NP	43.2%	39.4%	33.9%	29.2%	24.1%	20.5%	16.8%	14.1%	10.4%	7.9%	4.8%	2.0%	1.3%	0.3%
		DSRS	43.1%	39.6%	34.3%	29.3%	24.8%	21.3%	18.0%	15.2%	12.3%	9.7%	6.4%	3.7%	1.4%	0.4%
		(Growth)	-0.1%	0.2%	0.4%	0.1%	0.7%	0.8%	1.2%	1.1%	1.9%	1.8%	1.6%	1.7%	0.1%	0.1%
	4.0	NP	43.1%	39.5%	34.3%	29.6%	24.5%	21.3%	17.7%	15.0%	12.2%	9.6%	6.7%	4.0%	2.1%	1.3%
		DSRS	42.8%	39.6%	34.4%	29.4%	25.2%	21.8%	18.3%	15.8%	12.8%	10.2%	7.7%	4.6%	2.1%	0.9%
		(Growth)	-0.3%	0.1%	0.1%	-0.2%	0.7%	0.5%	0.6%	0.8%	0.6%	0.6%	1.0%	0.6%	0.0%	-0.4%
	8.0	NP	43.1%	39.5%	34.1%	29.8%	24.9%	21.7%	18.4%	15.6%	12.9%	10.1%	8.0%	5.6%	3.2%	1.6%
		DSRS	42.9%	39.5%	34.3%	29.8%	25.2%	21.9%	18.6%	15.8%	13.2%	10.6%	8.0%	5.4%	2.7%	1.3%
		(Growth)	-0.2%	0.0%	0.2%	0.0%	0.3%	0.2%	0.2%	0.2%	0.3%	0.5%	0.0%	-0.2%	-0.5%	-0.3%

Table 20. Full experimental results for certified accuracy, SmoothMix, ESG, CIFAR10

σ	η	Certification method	Certified accuracy at r													
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50
0.25	1.0	NP	64.8%	56.5%	46.7%											
		DSRS	64.6%	56.5%	46.5%											
		(Growth)	-0.2%	0.0%	-0.2%											
	2.0	NP	64.8%	56.5%	46.7%											
		DSRS	64.7%	56.3%	45.8%											
		(Growth)	-0.1%	-0.2%	-0.9%											
	4.0	NP	64.8%	56.4%	46.8%											
		DSRS	64.6%	56.3%	46.0%											
		(Growth)	-0.2%	-0.1%	-0.8%											
	8.0	NP	64.8%	56.5%	46.9%											
		DSRS	64.6%	56.1%	46.5%											
		(Growth)	-0.2%	-0.4%	-0.4%											
0.50	1.0	NP	53.3%	48.3%	42.1%	36.7%	31.6%	26.2%	20.1%							
		DSRS	53.3%	48.3%	42.0%	36.6%	31.3%	25.6%	18.1%							
		(Growth)	0.0%	0.0%	-0.1%	-0.1%	-0.3%	-0.6%	-2.0%							
	2.0	NP	53.2%	48.3%	42.2%	36.7%	31.7%	26.0%	20.4%							
		DSRS	53.3%	48.3%	42.0%	36.6%	31.2%	25.6%	17.6%							
		(Growth)	0.1%	0.0%	-0.2%	-0.1%	-0.5%	-0.4%	-2.8%							
	4.0	NP	53.3%	48.3%	42.1%	36.6%	31.6%	26.1%	20.3%							
		DSRS	53.1%	48.3%	42.0%	36.5%	31.0%	25.6%	18.3%							
		(Growth)	-0.2%	0.0%	-0.1%	-0.1%	-0.6%	-0.5%	-2.0%							
	8.0	NP	53.3%	48.3%	42.0%	36.5%	31.6%	26.3%	20.2%							
		DSRS	53.3%	48.3%	42.0%	36.4%	31.1%	25.5%	18.5%							
		(Growth)	0.0%	0.0%	0.0%	-0.1%	-0.5%	-0.8%	-1.7%							
1.00	1.0	NP	43.1%	39.6%	34.4%	30.2%	25.4%	22.2%	19.0%	16.3%	13.7%	11.6%	9.1%	6.7%	4.9%	2.3%
		DSRS	42.8%	39.5%	34.4%	29.9%	25.3%	22.0%	19.1%	16.2%	13.4%	11.4%	8.9%	6.3%	4.0%	1.8%
		(Growth)	-0.3%	-0.1%	0.0%	-0.3%	-0.1%	-0.2%	0.1%	-0.1%	-0.3%	-0.2%	-0.2%	-0.4%	-0.9%	-0.5%
	2.0	NP	43.0%	39.7%	34.4%	30.0%	25.5%	22.5%	19.0%	16.2%	13.6%	11.5%	9.3%	6.5%	4.6%	2.1%
		DSRS	42.9%	39.4%	34.3%	29.8%	25.3%	22.1%	18.9%	16.3%	13.4%	11.6%	8.8%	6.2%	4.0%	1.7%
		(Growth)	-0.1%	-0.3%	-0.1%	-0.2%	-0.2%	-0.4%	-0.1%	0.1%	-0.2%	0.1%	-0.5%	-0.3%	-0.6%	-0.4%
	4.0	NP	43.0%	39.7%	34.6%	30.1%	25.5%	22.4%	19.0%	16.3%	13.7%	11.6%	9.4%	6.6%	4.8%	2.4%
		DSRS	42.9%	39.5%	34.3%	29.8%	25.4%	22.1%	19.0%	16.2%	13.6%	11.5%	9.2%	6.0%	3.9%	1.8%
		(Growth)	-0.1%	-0.2%	-0.3%	-0.3%	-0.1%	-0.3%	0.0%	-0.1%	-0.1%	-0.1%	-0.2%	-0.6%	-0.9%	-0.6%
	8.0	NP	43.0%	39.7%	34.3%	30.0%	25.5%	22.3%	19.1%	16.2%	13.5%	11.5%	9.2%	6.5%	4.7%	2.5%
		DSRS	42.7%	39.4%	34.3%	29.8%	25.3%	22.3%	18.9%	16.3%	13.5%	11.3%	9.0%	6.2%	4.1%	1.6%
		(Growth)	-0.3%	-0.3%	0.0%	-0.2%	-0.2%	0.0%	-0.2%	0.1%	0.0%	-0.2%	-0.2%	-0.3%	-0.6%	-0.9%

H.3. Supplemental figures for StdAug-GGS models

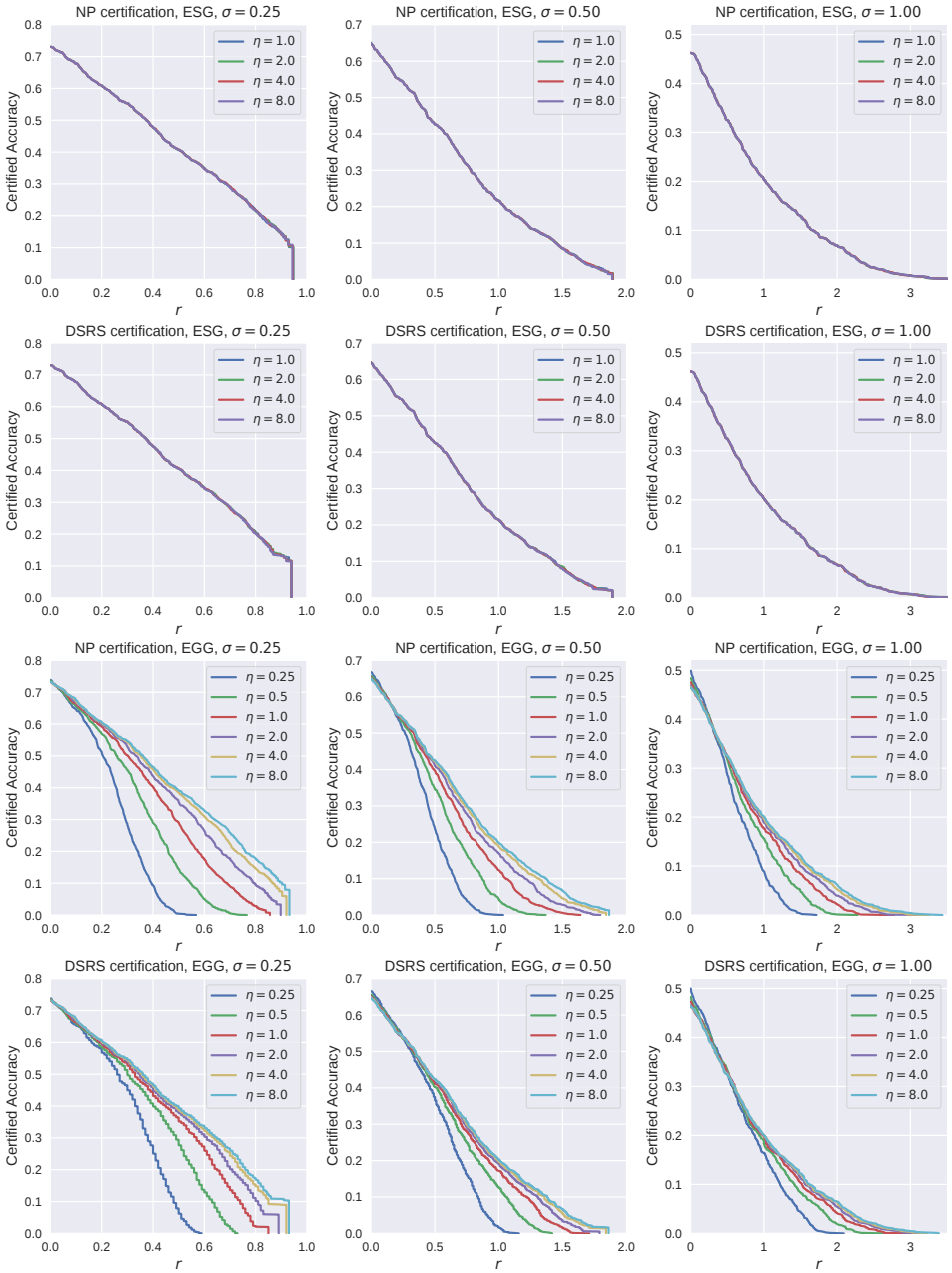


Figure 12. Certified accuracy of ℓ_2 for standardly augmented models, on CIFAR-10 by General Gaussian, $k = 1530$.

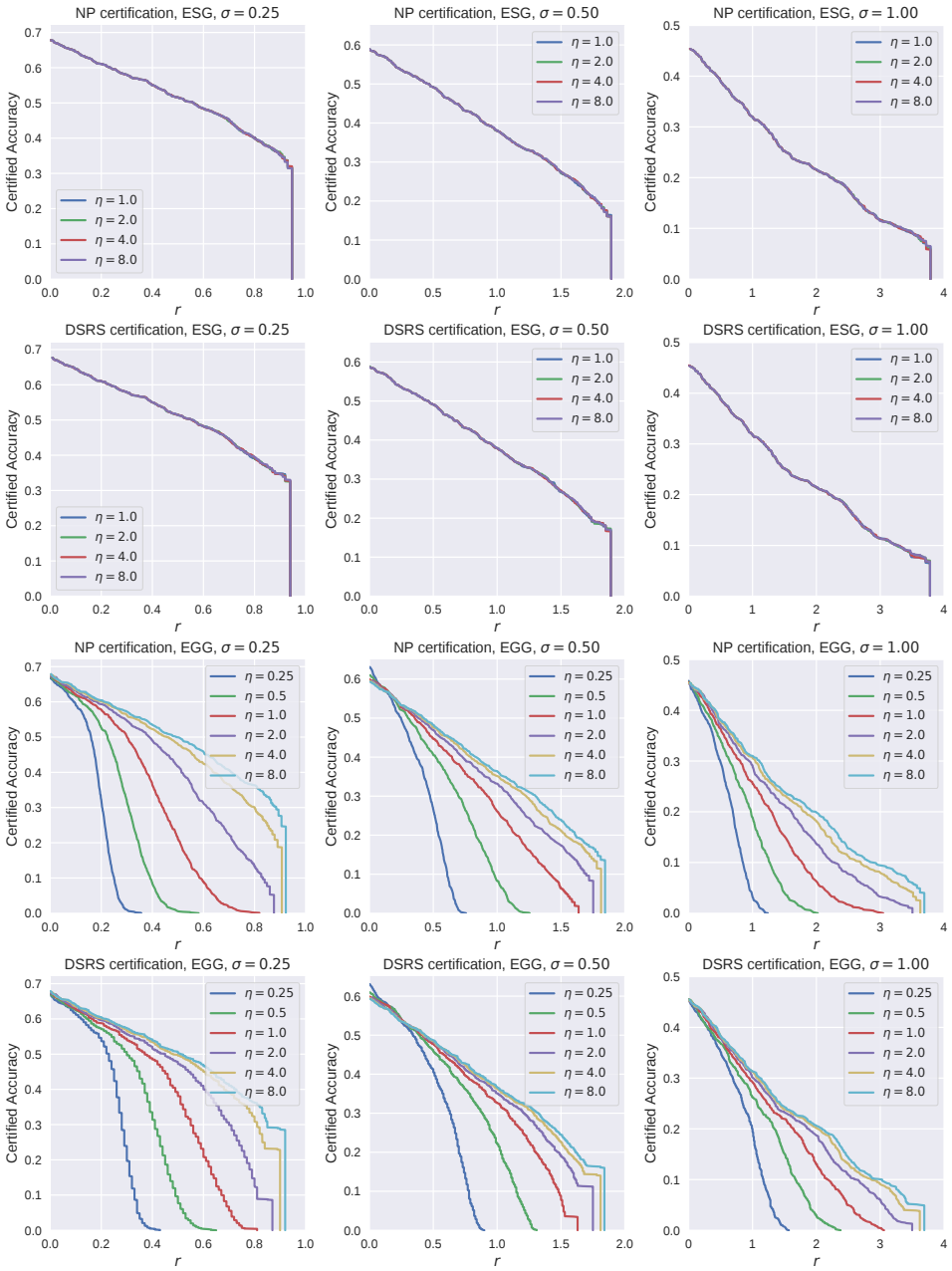


Figure 13. Certified accuracy of ℓ_2 for standardly augmented models, on ImageNet by General Gaussian, $k = 75260$.

H.4. Supplemental figures for Consistency-GGS and SmoothMix-GGS models

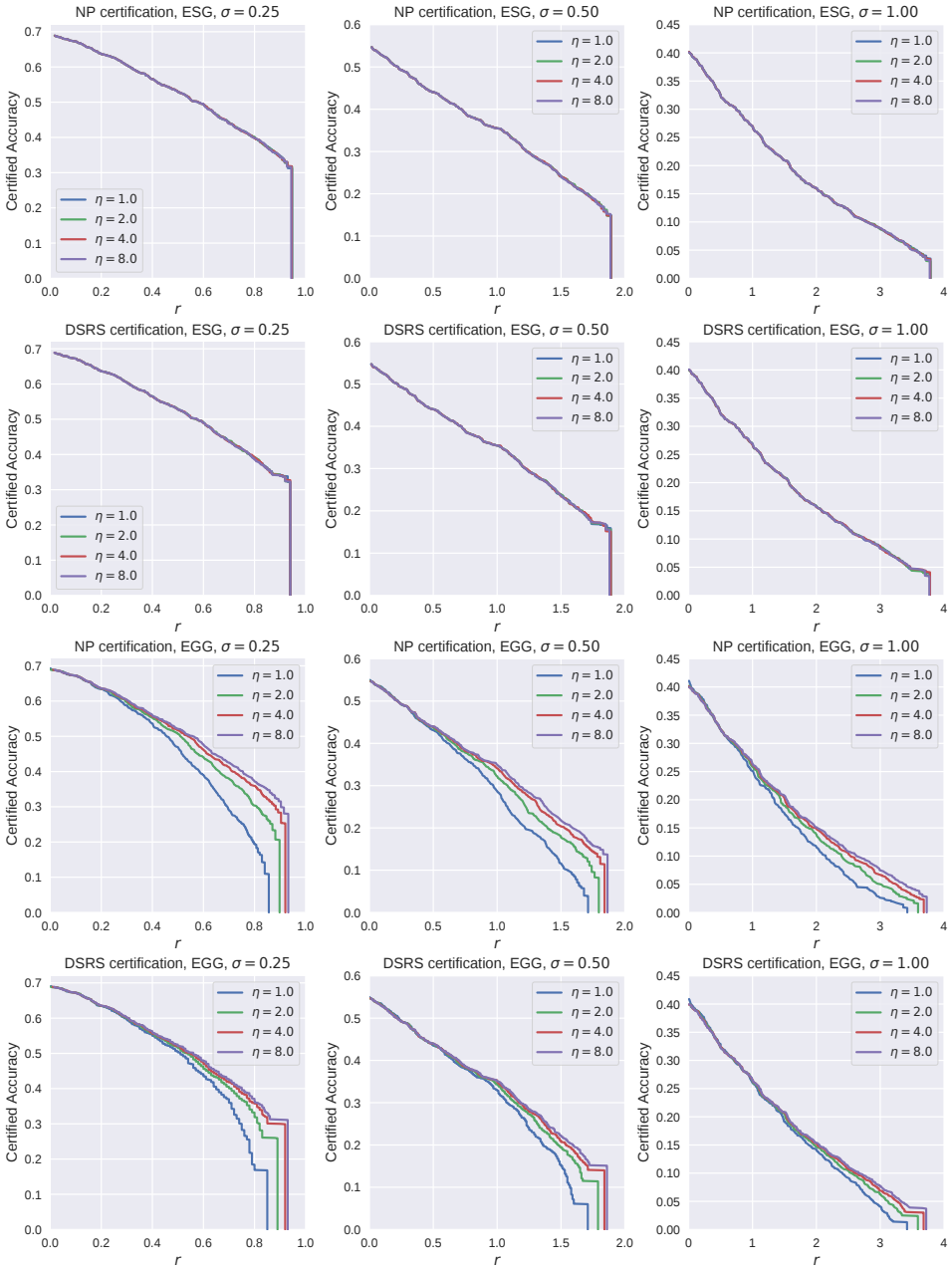


Figure 14. Certified accuracy of ℓ_2 for Consistency models, augmented on CIFAR-10 by General Gaussian, $k = 1530$.

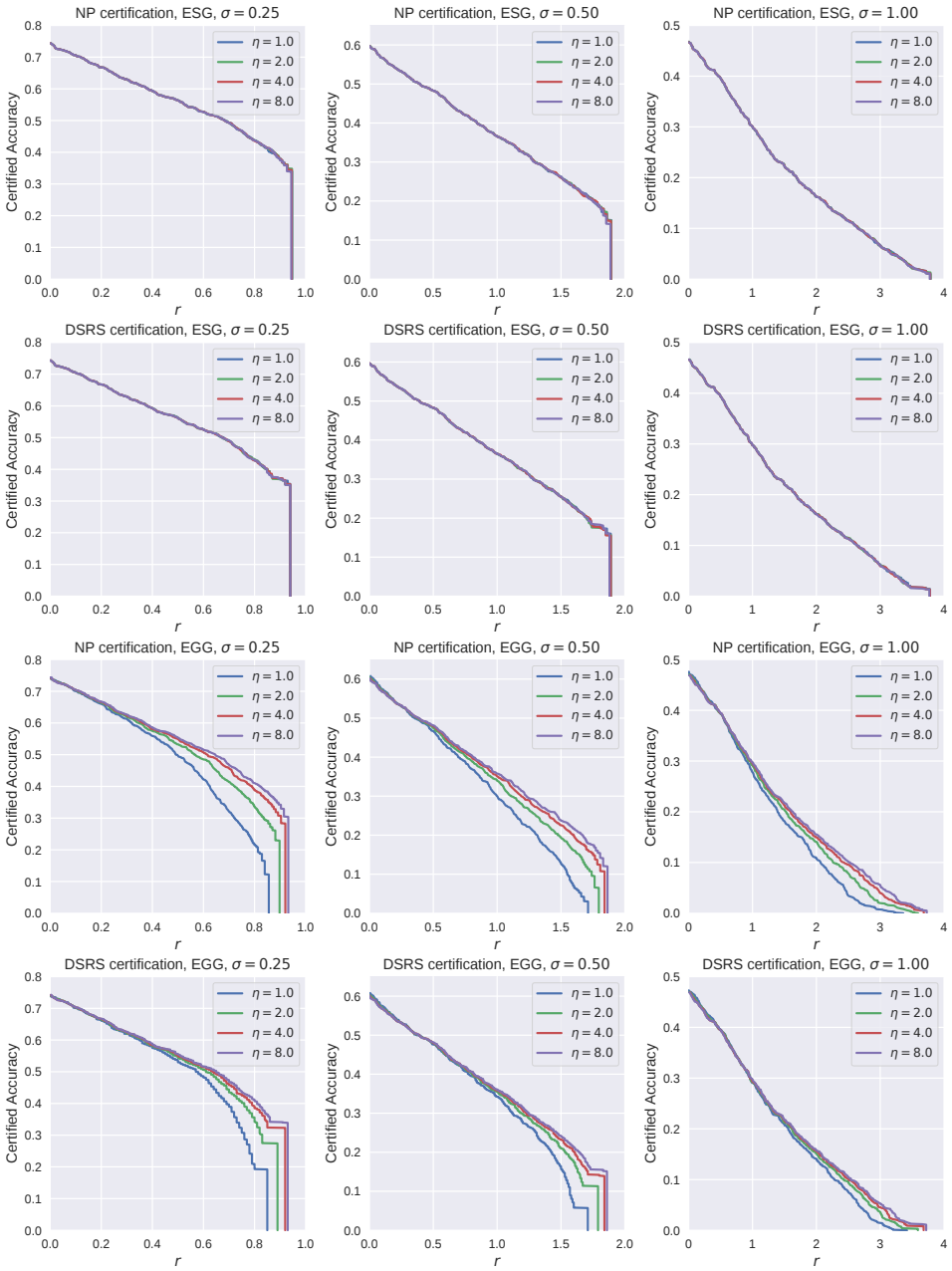


Figure 15. Certified accuracy of ℓ_2 for SmoothMix models, augmented on CIFAR-10 by General Gaussian, $k = 1530$.