

$$P_1 = \frac{\text{Vol}(\text{vert})}{\text{Vol}(\text{boule})}$$

- ① On prend un cône ($\{x > 0\} \cap \{r \leq r \tan \theta\}$, $(x, \rho, \theta_1, \dots, \theta_d)$)
- ② On injecte un bruit au point $(0, \dots, \delta)$
On calcule $P_1 = P(V(0, \delta) \notin \text{cône})$ avec un Monte-Carlo
- ③ On calcule le single-noise certificate $P_2^{(SN)}$
- ④ On calcule le vrai certificat $P_2^{(opt)}$ par:
 $P_2^{(opt)} = P(V(\delta, \delta) \notin \text{cône})$ avec un monte-carlo
- ⑤ On fait le graphe de $P_2^{(opt)} - P_2^{(SN)}$ en fonction de δ et de la dimension $(1, 2, \dots, 10)$

⑤ Une fois qu'on a sample les points (par ex selon $U(0, r)$) on les convertit en coord polaires : $(x_1, \dots, x_d) \rightarrow (x_1, \sqrt{x_1^2 + \dots + x_d^2}, \underbrace{\theta_2, \dots, \theta_d}_{(= \rho)}$

n'apparaissent pas dans l'éq du cône
→ inutiles.

⑥ Idem avec un paraboloid de révolution ($\rho \leq a x_1^2$) avec paramètre a .

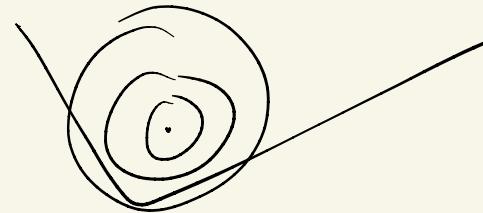
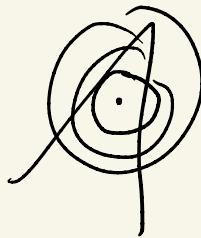
⑦ Idem avec bruit gaussien $N(0, \sigma^2)$
(nuance : le single-noise certificate se calcule différemment).

⑧ Tester nouveaux certificats en petite dimension

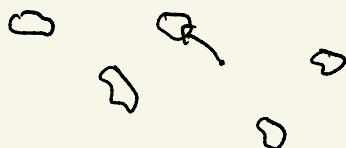
- * Bruits gaussiens concentriques
- * — uniformes —
- * Bruits gaussiens translates

(avec le cône / le paraboloid)

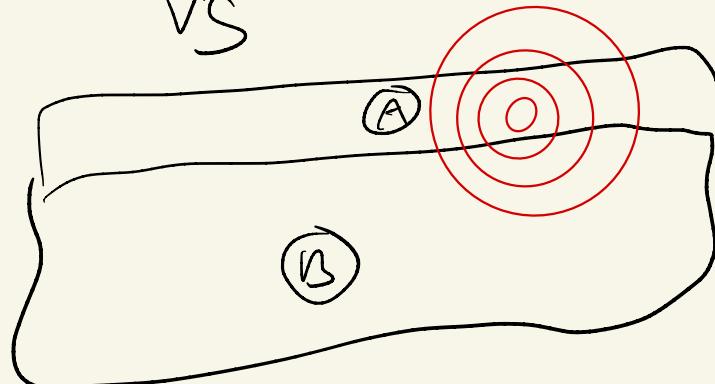
9



σ petit \rightarrow chang^t de p
 $\frac{\partial p}{\partial \sigma}$ est élevé



VS



p évalue
de façon
équilibrée

10 (maths) quel choix de bruits donne un certificat qui se calcule bien?

$$p_i = \int_{\text{Ney-Pearson set}} \underbrace{\prod_{j=1}^n f_j}_{\text{bruit utilisé}} d\mu \quad \begin{matrix} \text{mesure de comptage} \\ \text{par ex} \end{matrix}$$

\rightarrow se calculent
bien?

$$S = \left\{ z \in \mathbb{R}^d, f_0(z) \leq \sum h_i f_i(z) \right\}$$

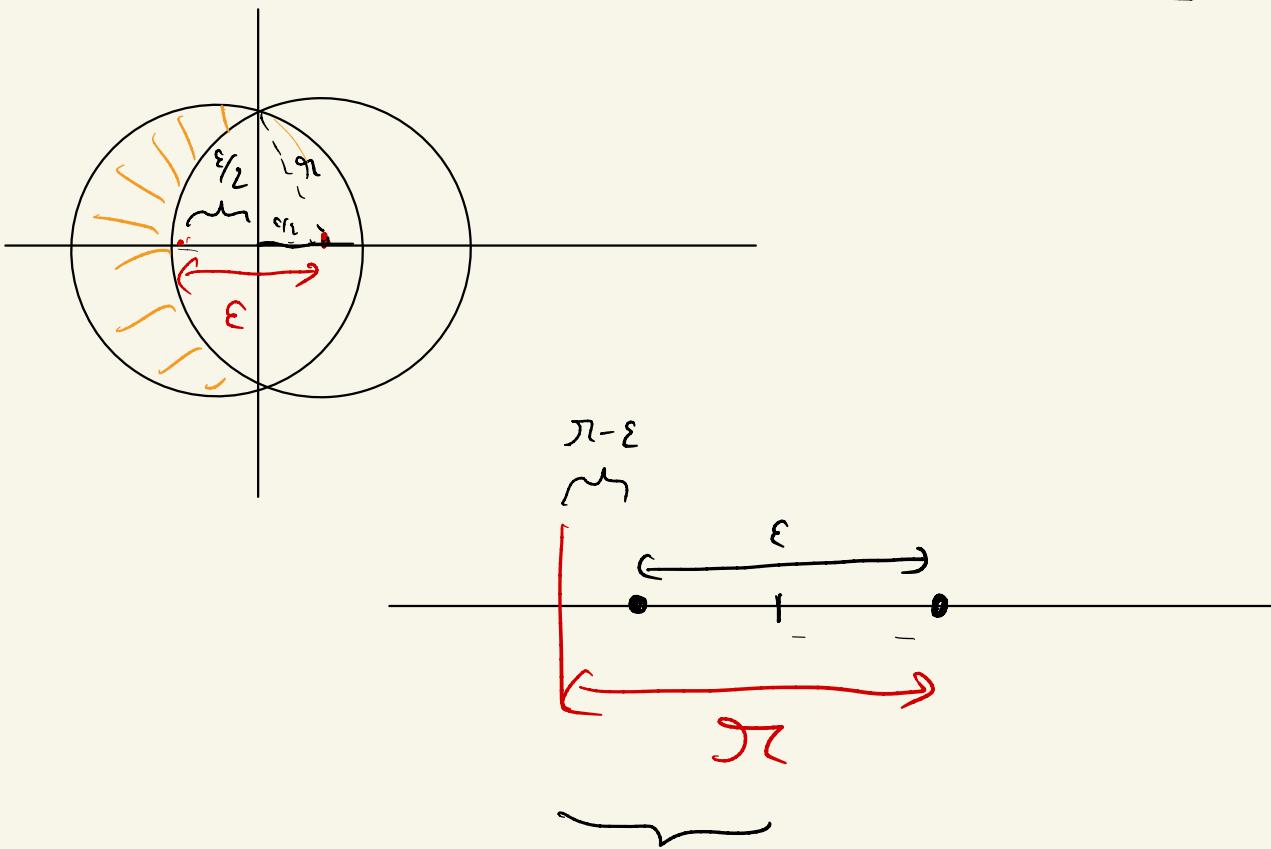
Pistes :

- Mettre des symétries dans les g_i (par ex bruits isotropes concentriques) puis par ex coord. polaires
 \rightarrow se ramène à un calcul en dim 2 ou 3

Ex : gaussiennes \rightarrow de $\mathcal{T} \neq$ $\{x_{SO}, e^{-\frac{\|x\|^2}{2\sigma_0^2}} \leq \sum h_i e^{-\frac{\|x\|^2}{2\sigma_i^2}}\}$

avec $x \sim \mathcal{N}(d)$ en dim 1.

Single-noise certificate en uniform



$$r - \varepsilon + \frac{\varepsilon}{2} = r - \frac{\varepsilon}{2}$$

Volume de la calotte sphérique

$$V_{cap} = \frac{1}{2} \underbrace{V_n(r)}_{\text{volume de la boule}} I_{1 - \left(\frac{\varepsilon}{2n}\right)^2} \left(\frac{d+1}{2}, \frac{1}{2} \right)$$

Incomplete regularized beta

$$I_8(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 t^{a-1} (1-t)^{b-1} dt$$

$$= \frac{\text{Incomplete beta}(a, b, 8)}{\text{beta}(a, b)}$$

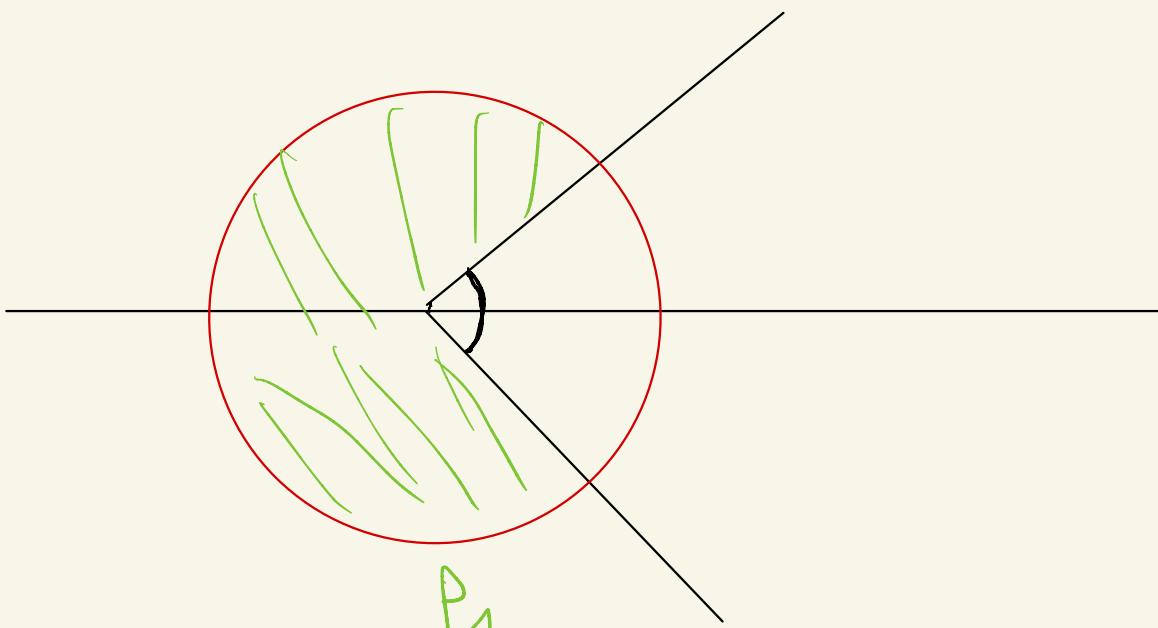
$$V_{orange} = V_n(r) - 2 V_{cap}$$

$$= V_n(r) \left(1 - I_{1 - \left(\frac{\varepsilon}{2n}\right)^2} \left(\frac{d+1}{2}, \frac{1}{2} \right) \right)$$

Single-noise certificate

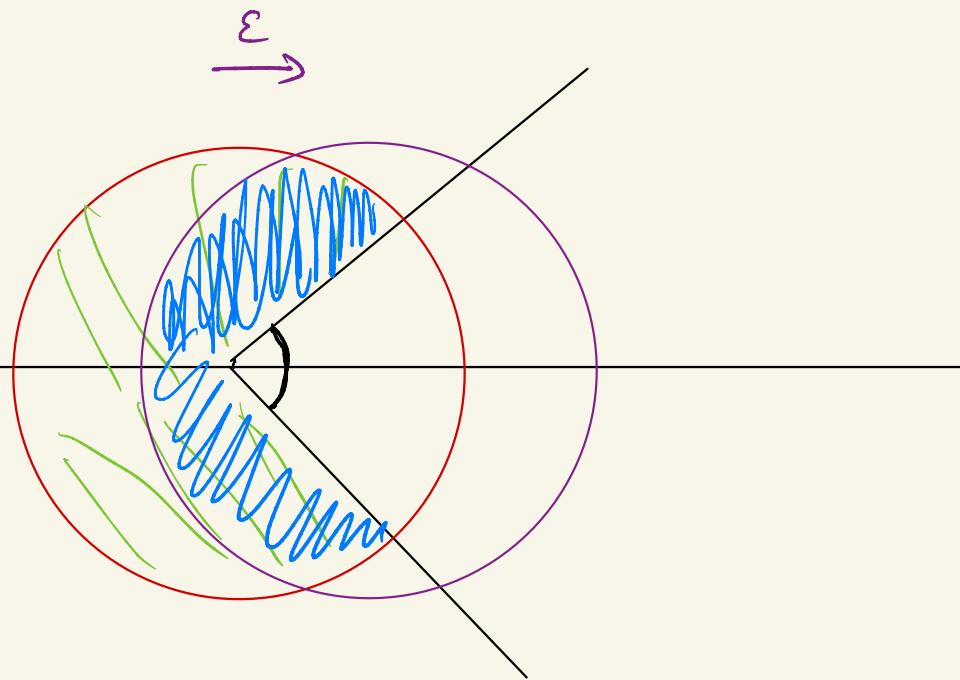
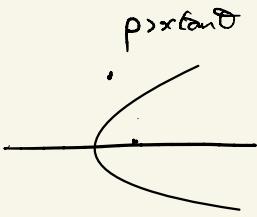
$$P_{SN} = \underbrace{P_1}_{\text{observed}} - \frac{V_{orange}}{V_n(n)}$$

$$P_{SN} = P_1 - \left(1 - I_{1 - \left(\frac{\epsilon}{2n}\right)^2} \left(\frac{d+1}{2}, \frac{1}{2} \right) \right)$$



$P_1 =$
relative volume
of this area

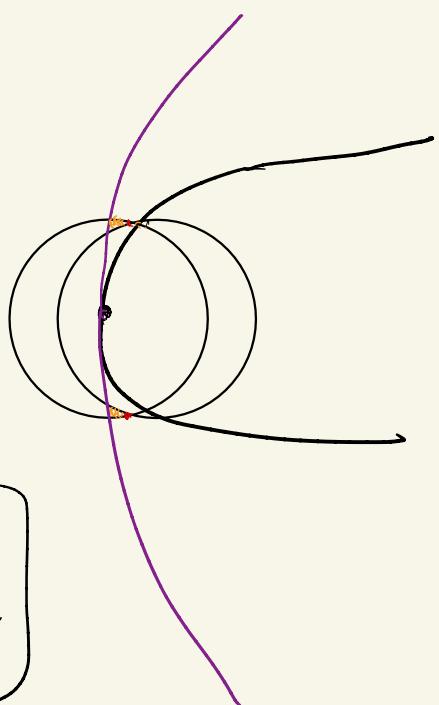
True certificate:



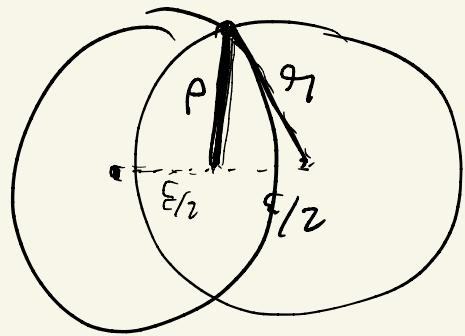
$$P_{\text{true}} = \frac{\text{vol (blue)}}{\text{vol (ball)}} \rightarrow \text{Monte-carlo sur la boule translatée.}$$

Single-noise certificate:

$$\phi \left(\phi^{-1}(p_1) - \frac{\varepsilon}{\sigma} \right)$$



$$p^2 < a\gamma$$

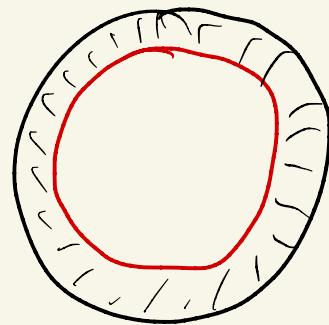
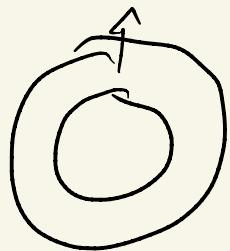
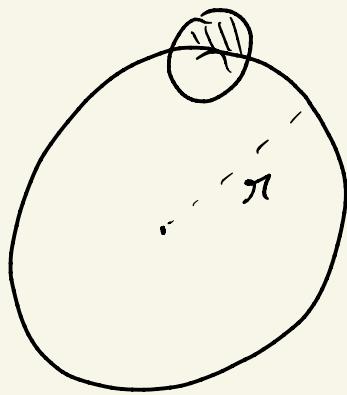


$$\begin{aligned} r &= 1 \\ \varepsilon &= 0,5 \end{aligned}$$

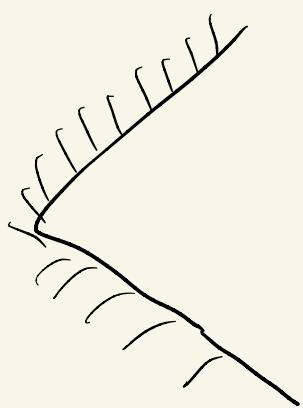
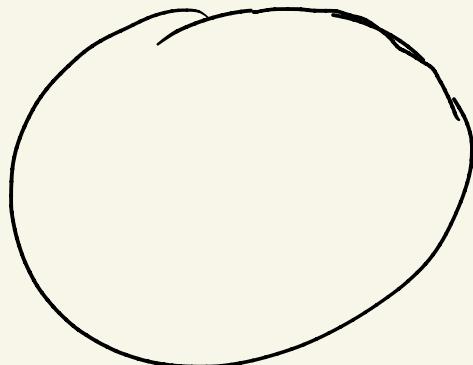
$$p^2 = r^2 - \left(\frac{\varepsilon}{2}\right)^2 = a \frac{\varepsilon}{2}$$

$$\pi^2 - \frac{\varepsilon^2}{4} = a \frac{\varepsilon}{2}$$

$$a = \frac{2 \left(\pi^2 - \frac{\varepsilon^2}{4} \right)}{\varepsilon}$$



nouvelle DB
après Randomized
smoothing.



Bruits uniformes concentriques

rayon r_1 utilisé pour le SNC.

Vs certificat rayon r_1 mais information par r_1, r_2, \dots, r_n .

Par ex dim 3 ou 5.

Generalized Neyman-Pearson set

$$S = \left\{ z \in \mathbb{R}^d, \frac{\mathbb{1}_{z \in B(0, r_1)}}{V(r_1)} \leq k_1 \frac{\mathbb{1}_{z \in B(0, r_1)}}{V(r_1)} \right.$$

$$\left. + k_2 \frac{\mathbb{1}_{z \in B(0, r_2)}}{V(r_2)} \right\}$$

paramètres

On calcule $\underline{P}_1 = P(\underline{B}(0, r_1) \notin \text{cone})$

$P_2 = P(\underline{B}(0, r_2) \notin \hat{\text{cone}})$

On trouve k_1 et k_2 tels que :

$$P(B(0, r_1) \in S) = p_1 \quad (\text{ou } \leq \text{de per})$$

$$P(B(0, r_2) \in S) = p_2 \quad (\text{ou } \leq \text{de per})$$

Et ensuite on aura (par Neyman - Pearson)

$$P(B(\varepsilon, r_1) \notin \text{cone}) \geq P(B(\varepsilon, r_1) \in S_{k_1, k_2})$$

↑
après attaque

Suabilité :

$$S = \left\{ z \in \mathbb{R}^d, \frac{\mathbf{1}_{z \in B(\varepsilon, r_1)}}{\mathbf{1}_{z \in B(0, r_1)}} \leq \dots \right.$$

$$\left. + \frac{k_2}{k_1} \frac{V(r_2)}{V(r_1)} \frac{\mathbf{1}_{z \in B(0, r_2)}}{\mathbf{1}_{z \in B(0, r_1)}} \right\}$$

paramètres

$$\text{et } \frac{V(\pi_1)}{V(\pi_2)} = \left(\frac{\pi_1}{\pi_2} \right)^d = \exp \left(d \ln \left(\frac{\pi_1}{\pi_2} \right) \right)$$

$$\rho_2 \frac{V(\pi_1)}{V(\pi_2)} = \exp \left(\ln(k_2) \right) \exp \left(d \ln \left(\frac{\pi_1}{\pi_2} \right) \right)$$

$$= \exp \left(\underbrace{\ln(k_2)}_{\hookrightarrow \text{param en grande dim.}} + d \ln \left(\frac{\pi_1}{\pi_2} \right) \right)$$

Mais pour le moment faire ça en petite dimension