# A remark on the minimal dispersion

**Alexander Litvak**

University of Alberta

# Problem.

**Problem.** Given $\varepsilon \in (0,1)$ and $d \geq 1$ what is the smallest $n$ such that there exist $n$ points in the unit $d$-dimensional cube $[0,1]^d$ with the following property:

*any axis-parallel box of volume* $\varepsilon$ *contains at least one point?*

# Problem.

**Problem.** Given $\varepsilon \in (0,1)$ and $d \geq 1$ what is the smallest $n$ such that there exist $n$ points in the unit $d$-dimensional cube $[0,1]^d$ with the following property:

*any axis-parallel box of volume* $\varepsilon$ *contains at least one point?*

Such an integer $n$ is denoted below by

$$N(\varepsilon, d).$$

# Problem.

**Problem.** Given $\varepsilon \in (0,1)$ and $d \geq 1$ what is the smallest $n$ such that there exist $n$ points in the unit $d$-dimensional cube $[0,1]^d$ with the following property:

*any axis-parallel box of volume* $\varepsilon$ *contains at least one point?*

Such an integer $n$ is denoted below by

$$N(\varepsilon, d).$$

**Equivalently:** Given integers $n, d \geq 1$ what is the largest $\varepsilon > 0$ such that for any $n$ points in the unit $d$-dimensional cube $[0,1]^d$ there exists an axis-parallel box of volume $\varepsilon$ containing none of these points?

# Problem.

**Problem.** Given $\varepsilon \in (0,1)$ and $d \geq 1$ what is the smallest $n$ such that there exist $n$ points in the unit $d$-dimensional cube $[0,1]^d$ with the following property:

*any axis-parallel box of volume $\varepsilon$ contains at least one point?*

Such an integer $n$ is denoted below by

$$N(\varepsilon, d).$$

**Equivalently:** Given integers $n, d \geq 1$ what is the largest $\varepsilon > 0$ such that for any $n$ points in the unit $d$-dimensional cube $[0,1]^d$ there exists an axis-parallel box of volume $\varepsilon$ containing none of these points?

Such an $\varepsilon$ is called dispersion (of the cube) or *minimal dispersion* and denoted by

$$\mathrm{disp}^*(n, d).$$

# Notations.

Consider the set of all axis parallel boxes contained in the cube $[0, 1]^d$,

$$\mathcal{R}_d := \left\{ \prod_{i=1}^{d} I_i \;\middle|\; I_i = [a_i, b_i) \subset [0, 1] \right\}.$$

# Notations.

Consider the set of all axis parallel boxes contained in the cube $[0,1]^d$,

$$\mathcal{R}_d := \left\{ \prod_{i=1}^{d} I_i \mid I_i = [a_i, b_i) \subset [0,1] \right\}.$$

The *dispersion* of a finite set of points $P \subset [0,1]^d$ is defined as

$$\operatorname{disp}(P) = \sup\{|B| \mid B \in \mathcal{R}_d,\ B \cap P = \emptyset\}.$$

# Notations.

Consider the set of all axis parallel boxes contained in the cube $[0,1]^d$,

$$\mathcal{R}_d := \left\{ \prod_{i=1}^{d} I_i \mid I_i = [a_i, b_i) \subset [0,1] \right\}.$$

The *dispersion* of a finite set of points $P \subset [0,1]^d$ is defined as

$$\mathrm{disp}(P) = \sup\{|B| \mid B \in \mathcal{R}_d,\ B \cap P = \emptyset\}.$$

Then the *minimal dispersion* is defined as the function of two variables, namely

$$\mathrm{disp}^*(n,d) = \inf_{|P|=n} \mathrm{disp}(P).$$

# Notations.

Consider the set of all axis parallel boxes contained in the cube $[0, 1]^d$,

$$\mathcal{R}_d := \left\{ \prod_{i=1}^{d} I_i \mid I_i = [a_i, b_i) \subset [0, 1] \right\}.$$

The *dispersion* of a finite set of points $P \subset [0, 1]^d$ is defined as

$$\mathrm{disp}(P) = \sup\{|B| \mid B \in \mathcal{R}_d,\ B \cap P = \emptyset\}.$$

Then the *minimal dispersion* is defined as the function of two variables, namely

$$\mathrm{disp}^*(n, d) = \inf_{|P|=n} \mathrm{disp}(P).$$

In this talk it will be more convenient to work with its inverse, the function

$$N(\varepsilon, d) = \min\{n \in \mathbb{N} \mid \mathrm{disp}^*(n, d) \leq \varepsilon\}.$$

# Known results.

Rote–Tichy (**96**), Larcher (**17**): $$N(\varepsilon, d) \le \frac{2^{7d+1}}{\varepsilon}$$

# Known results.

Rote–Tichy (**96**), Larcher (**17**): $$N(\varepsilon, d) \le \frac{2^{7d+1}}{\varepsilon}$$

Bukh–Chao (**20+**): $$N(\varepsilon, d) \le \frac{Cd^2 \log d}{\varepsilon}$$

# Known results.

Rote–Tichy (**96**), Larcher (**17**): $$N(\varepsilon, d) \leq \frac{2^{7d+1}}{\varepsilon}$$

Bukh–Chao (**20+**): $$N(\varepsilon, d) \leq \frac{Cd^2 \log d}{\varepsilon}$$

For a lower bound one trivially has

$$\text{disp}^*(n, d) \geq \frac{1}{n+1} \quad \Longleftrightarrow \quad N(\varepsilon, d) \geq \frac{1}{\varepsilon} - 1.$$

# Known results.

Rote–Tichy (**96**), Larcher (**17**): $$N(\varepsilon, d) \leq \frac{2^{7d+1}}{\varepsilon}$$

Bukh–Chao (**20+**): $$N(\varepsilon, d) \leq \frac{Cd^2 \log d}{\varepsilon}$$

For a lower bound one trivially has

$$\mathrm{disp}^*(n, d) \geq \frac{1}{n+1} \quad \Longleftrightarrow \quad N(\varepsilon, d) \geq \frac{1}{\varepsilon} - 1.$$

Dumitrescu–Jiang (**13**): $$\mathrm{disp}^*(n, d) \geq \frac{5}{4(n+5)} \quad \Longleftrightarrow \quad N(\varepsilon, d) \geq \frac{5}{4\varepsilon} - 5.$$

# Known results.

Rote–Tichy (**96**), Larcher (**17**): $$N(\varepsilon, d) \le \frac{2^{7d+1}}{\varepsilon}$$

Bukh–Chao (**20+**): $$N(\varepsilon, d) \le \frac{Cd^2 \log d}{\varepsilon}$$

For a lower bound one trivially has

$$\mathrm{disp}^*(n,d) \ge \frac{1}{n+1} \quad \Longleftrightarrow \quad N(\varepsilon, d) \ge \frac{1}{\varepsilon} - 1.$$

Dumitrescu–Jiang (**13**): $$\mathrm{disp}^*(n,d) \ge \frac{5}{4(n+5)} \quad \Longleftrightarrow \quad N(\varepsilon, d) \ge \frac{5}{4\varepsilon} - 5.$$

The first bound showing that the minimal dispersion grows with the dimension was obtained by Aistleitner–Hinrichs–Rudolf (**17**):

$$N(\varepsilon, d) \ge (1 - 4\varepsilon)\, \frac{\log_2 d}{4\varepsilon}.$$

# Known results.

$$(1-4\varepsilon)\,\frac{\log_2 d}{4\varepsilon} \le N(\varepsilon,d) \le \frac{Cd^2\log d}{\varepsilon}. \qquad (1)$$

Thus, when the dimension $d$ is fixed and $\varepsilon \to 0$ the problem is essentially solved:

$$N(\varepsilon,d) \approx \frac{C_d}{\varepsilon}.$$

# Known results.

$$(1-4\varepsilon)\,\frac{\log_2 d}{4\varepsilon} \le N(\varepsilon,d) \le \frac{Cd^2\log d}{\varepsilon}. \qquad (1)$$

Thus, when the dimension $d$ is fixed and $\varepsilon \to 0$ the problem is essentially solved:

$$N(\varepsilon,d) \approx \frac{C_d}{\varepsilon}.$$

However, there is a huge gap between the bounds when $d \to \infty$.

# Known results.

$$(1-4\varepsilon)\,\frac{\log_2 d}{4\varepsilon} \le N(\varepsilon, d) \le \frac{Cd^2 \log d}{\varepsilon}. \tag{1}$$

Thus, when the dimension $d$ is fixed and $\varepsilon \to 0$ the problem is essentially solved:

$$N(\varepsilon, d) \approx \frac{C_d}{\varepsilon}.$$

However, there is a huge gap between the bounds when $d \to \infty$.

Blumer–Ehrenfeucht–Haussler–Warmuth (**89**) provided a general bound in terms of VC dimension of $\mathcal{R}_d$. Using that this dimension equals to $2d$,

$$N(\varepsilon, d) \le \frac{Cd}{\varepsilon} \log_2\left(\frac{C}{\varepsilon}\right). \tag{2}$$

# Known results.

$$(1 - 4\varepsilon)\,\frac{\log_2 d}{4\varepsilon} \le N(\varepsilon, d) \le \frac{Cd^2 \log d}{\varepsilon}. \tag{1}$$

Thus, when the dimension $d$ is fixed and $\varepsilon \to 0$ the problem is essentially solved:

$$N(\varepsilon, d) \approx \frac{C_d}{\varepsilon}.$$

However, there is a huge gap between the bounds when $d \to \infty$.

Blumer–Ehrenfeucht–Haussler–Warmuth (**89**) provided a general bound in terms of VC dimension of $\mathcal{R}_d$. Using that this dimension equals to $2d$,

$$N(\varepsilon, d) \le \frac{Cd}{\varepsilon} \log_2\left(\frac{C}{\varepsilon}\right). \tag{2}$$

Rudolf (**18**): this can be proved using random points uniformly distributed in $[0, 1]^d$.

# Known results.

$$(1-4\varepsilon)\,\frac{\log_2 d}{4\varepsilon} \le N(\varepsilon, d) \le \frac{Cd^2 \log d}{\varepsilon}. \qquad (1)$$

Thus, when the dimension $d$ is fixed and $\varepsilon \to 0$ the problem is essentially solved:

$$N(\varepsilon, d) \approx \frac{C_d}{\varepsilon}.$$

However, there is a huge gap between the bounds when $d \to \infty$.

Blumer–Ehrenfeucht–Haussler–Warmuth (**89**) provided a general bound in terms of VC dimension of $\mathcal{R}_d$. Using that this dimension equals to $2d$,

$$N(\varepsilon, d) \le \frac{Cd}{\varepsilon} \log_2 \left(\frac{C}{\varepsilon}\right). \qquad (2)$$

Rudolf (**18**): this can be proved using random points uniformly distributed in $[0, 1]^d$.

Estimate (2) is better than the upper bound in (1) in the regime $\varepsilon \ge d^{-Cd}$.

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

Sosnovec (**18**): $N(\varepsilon, d) \leq C_\varepsilon \log_2 d$ for $\varepsilon < 1/4$.

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

Sosnovec (**18**): $N(\varepsilon, d) \leq C_\varepsilon \log_2 d$ for $\varepsilon < 1/4$.

Thus, when $\varepsilon \in (0, 1/4)$ is fixed and $d \to \infty$, we have

$$N(\varepsilon, d) \approx C_\varepsilon \log_2 d.$$

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

Sosnovec (**18**): $N(\varepsilon, d) \leq C_\varepsilon \log_2 d$ for $\varepsilon < 1/4$.

Thus, when $\varepsilon \in (0, 1/4)$ is fixed and $d \to \infty$, we have

$$N(\varepsilon, d) \approx C_\varepsilon \log_2 d.$$

**Dependence on $\varepsilon$.**

Sosnovec (**18**): $C_\varepsilon \approx (1/\varepsilon)^{(C/\varepsilon)^2}$.

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

Sosnovec (**18**): $N(\varepsilon, d) \leq C_\varepsilon \log_2 d$ for $\varepsilon < 1/4$.

Thus, when $\varepsilon \in (0, 1/4)$ is fixed and $d \to \infty$, we have

$$N(\varepsilon, d) \approx C_\varepsilon \log_2 d.$$

**Dependence on $\varepsilon$.**

Sosnovec (**18**): $C_\varepsilon \approx (1/\varepsilon)^{(C/\varepsilon)^2}$.

Ullrich and Vybíral (**18**):

$$C_\varepsilon = \frac{2^7}{\varepsilon^2} \left( \log_2 \left( \frac{1}{\varepsilon} \right) \right)^2 .$$

They also **conjectured** that $N(\varepsilon, d) \approx \frac{\log d}{\varepsilon}$.

# Known results.

**A natural conjecture:** $N(\varepsilon, d) \approx \frac{d}{\varepsilon}$. (Bukh–Chao: $N(\varepsilon, d) \geq \frac{d}{e\varepsilon}$ if $\varepsilon \leq (4d)^{-d}$.)

Sosnovec (**18**): $N(\varepsilon, d) \leq C_\varepsilon \log_2 d$ for $\varepsilon < 1/4$.

Thus, when $\varepsilon \in (0, 1/4)$ is fixed and $d \to \infty$, we have

$$N(\varepsilon, d) \approx C_\varepsilon \log_2 d.$$

**Dependence on $\varepsilon$.**

Sosnovec (**18**): $C_\varepsilon \approx (1/\varepsilon)^{(C/\varepsilon)^2}$.

Ullrich and Vybíral (**18**):

$$C_\varepsilon = \frac{2^7}{\varepsilon^2} \left( \log_2 \left( \frac{1}{\varepsilon} \right) \right)^2.$$

They also **conjectured** that $N(\varepsilon, d) \approx \frac{\log d}{\varepsilon}$.

Their upper bound is better in the regime

$$\varepsilon \geq \frac{C\,(\log_2 d)^2}{d}.$$

# Known results.

The Sosnovec–Ullrich–Vybíral proof is also based on a random choice of points, but instead of the uniform distribution on $[0, 1]^d$ they use uniform distribution on a certain lattice, gaining in the case of relatively large $\varepsilon$.

# Known results.

The Sosnovec–Ullrich–Vybíral proof is also based on a random choice of points, but instead of the uniform distribution on $[0,1]^d$ they use uniform distribution on a certain lattice, gaining in the case of relatively large $\varepsilon$.

**Remarks on very large $\varepsilon$.**

If $\varepsilon \in [1/2, 1]$ then $N(\varepsilon, d) = 1$, indeed one can take the point $(1/2, 1/2, ..., 1/2)$.

# Known results.

The Sosnovec–Ullrich–Vybíral proof is also based on a random choice of points, but instead of the uniform distribution on $[0,1]^d$ they use uniform distribution on a certain lattice, gaining in the case of relatively large $\varepsilon$.

**Remarks on very large $\varepsilon$.**

If $\varepsilon \in [1/2, 1]$ then $N(\varepsilon, d) = 1$, indeed one can take the point $(1/2, 1/2, ..., 1/2)$.

Sosnovec (**18**): $$N(\varepsilon, d) \leq 1 + \left\lfloor \frac{1}{\varepsilon - 1/4} \right\rfloor \qquad \text{for} \quad \varepsilon > 1/4.$$

Does not grow when $d \to \infty$. Recall, for $\varepsilon < 1/4$, we have $N(\varepsilon, d) \geq C_\varepsilon \log_2 d$.

# Known results.

The Sosnovec–Ullrich–Vybíral proof is also based on a random choice of points, but instead of the uniform distribution on $[0,1]^d$ they use uniform distribution on a certain lattice, gaining in the case of relatively large $\varepsilon$.

**Remarks on very large $\varepsilon$.**

If $\varepsilon \in [1/2, 1]$ then $N(\varepsilon, d) = 1$, indeed one can take the point $(1/2, 1/2, ..., 1/2)$.

Sosnovec (**18**): $$N(\varepsilon, d) \le 1 + \left\lfloor \frac{1}{\varepsilon - 1/4} \right\rfloor \qquad \text{for} \quad \varepsilon > 1/4.$$

Does not grow when $d \to \infty$. Recall, for $\varepsilon < 1/4$, we have $N(\varepsilon, d) \geq C_\varepsilon \log_2 d$.

Kurt MacKay (**20+**): $$N(\varepsilon, d) \le \frac{C}{\sqrt{\varepsilon - 1/4}} \qquad \text{for} \quad \varepsilon > 1/4.$$

# Known results.

The Sosnovec–Ullrich–Vybíral proof is also based on a random choice of points, but instead of the uniform distribution on $[0,1]^d$ they use uniform distribution on a certain lattice, gaining in the case of relatively large $\varepsilon$.

**Remarks on very large $\varepsilon$.**

If $\varepsilon \in [1/2, 1]$ then $N(\varepsilon, d) = 1$, indeed one can take the point $(1/2, 1/2, ..., 1/2)$.

Sosnovec (**18**): $$N(\varepsilon, d) \le 1 + \left\lfloor \frac{1}{\varepsilon - 1/4} \right\rfloor \qquad \text{for} \quad \varepsilon > 1/4.$$

Does not grow when $d \to \infty$. Recall, for $\varepsilon < 1/4$, we have $N(\varepsilon, d) \ge C_\varepsilon \log_2 d$.

Kurt MacKay (**20+**): $$N(\varepsilon, d) \le \frac{C}{\sqrt{\varepsilon - 1/4}} \qquad \text{for} \quad \varepsilon > 1/4.$$

**Problem.** What is $N(1/4, d)$? How does $N(\varepsilon, d)$ behave when $\varepsilon \to (1/4)^{\pm}$.

# Upper bounds: summary

Till very recently:

$$N(\varepsilon, d) \leq \begin{cases} \dfrac{C\ \ln d}{\varepsilon^2} \ln^2\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \varepsilon \geq \dfrac{\ln^2 d}{d}, \\ \dfrac{C\, d}{\varepsilon} \ln\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \dfrac{\ln^2 d}{d} \geq \varepsilon \geq \exp(-C^d), \\ \dfrac{C^d}{\varepsilon}, & \text{if} \quad \varepsilon \leq \exp(-C^d). \end{cases}$$

# Upper bounds: summary

Till very recently:

$$N(\varepsilon,d) \le \begin{cases} \dfrac{C\,\ln d}{\varepsilon^2}\,\ln^2\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \varepsilon \ge \dfrac{\ln^2 d}{d}, \\ \dfrac{C\,d}{\varepsilon}\,\ln\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \dfrac{\ln^2 d}{d} \ge \varepsilon \ge \exp(-C^d), \\ \dfrac{C^d}{\varepsilon}, & \text{if} \quad \varepsilon \le \exp(-C^d). \end{cases}$$

Using the Bukh–Chao result:

$$N(\varepsilon,d) \le \begin{cases} \dfrac{C\,\ln d}{\varepsilon^2}\,\ln^2\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \varepsilon \ge \dfrac{\ln^2 d}{d}, \\ \dfrac{C\,d}{\varepsilon}\,\ln\left(\dfrac{1}{\varepsilon}\right), & \text{if} \quad \dfrac{\ln^2 d}{d} \ge \varepsilon \ge \exp(-Cd\ln d), \\ \dfrac{C\,d^2\ln d}{\varepsilon}, & \text{if} \quad \varepsilon \le \exp(-Cd\ln d). \end{cases}$$

# New bounds.

## Theorem (small $\varepsilon$)

*Let* $d \geq 2$ *and* $\varepsilon \leq 1/2$. *Then*

(i) $$N(\varepsilon, d) \leq \frac{C \ln d}{\varepsilon} \ln\left(\frac{1}{\varepsilon}\right),$$ *provided that* $\varepsilon \leq \exp(-d)$,

(ii) $$N(\varepsilon, d) \leq \frac{C\, d}{\varepsilon} \ln\ln\left(\frac{2}{\varepsilon}\right),$$ *provided that* $\varepsilon \geq \exp(-d)$.

*Moreover, the random choice of points uniformly distributed on* $[0, 1]^d$ *works.*

# New bounds.

**Theorem (small $\varepsilon$)**

*Let* $d \geq 2$ *and* $\varepsilon \leq 1/2$. *Then*

(*i*) $$N(\varepsilon, d) \leq \frac{C \ln d}{\varepsilon} \ln\left(\frac{1}{\varepsilon}\right), \qquad \textit{provided that} \quad \varepsilon \leq \exp(-d),$$

(*ii*) $$N(\varepsilon, d) \leq \frac{C d}{\varepsilon} \ln\ln\left(\frac{2}{\varepsilon}\right), \qquad \textit{provided that} \quad \varepsilon \geq \exp(-d).$$

*Moreover, the random choice of points uniformly distributed on* $[0,1]^d$ *works.*

Thus, for $\varepsilon \leq \exp(-d)$ we have

$$\frac{\ln d}{6\varepsilon} \leq N(\varepsilon, d) \leq \frac{C \ln d}{\varepsilon} \ln\left(\frac{1}{\varepsilon}\right),$$

In $(ii)$ the improvement is only in substitution of $\ln(1/\varepsilon)$ with $\ln\ln(1/\varepsilon)$.

# New bounds.

Hinrichs–Krieg–Kunsch–Rudolf (**20**): for a random choice of points

$$\max\left\{\frac{c}{\varepsilon}\ln\left(\frac{1}{\varepsilon}\right),\ \frac{d}{2\varepsilon}\right\} \le N_{ran}(\varepsilon, d)$$

# New bounds.

Hinrichs–Krieg–Kunsch–Rudolf (**20**): for a random choice of points

$$\max\left\{\frac{c}{\varepsilon}\ln\left(\frac{1}{\varepsilon}\right), \frac{d}{2\varepsilon}\right\} \le N_{ran}(\varepsilon, d) \le \max\left\{\frac{C\ln d}{\varepsilon}\ \ln\left(\frac{1}{\varepsilon}\right), \frac{C\,d}{\varepsilon}\ \ln\ln\left(\frac{2}{\varepsilon}\right)\right\}$$

## Theorem (relatively large $\varepsilon$)

*Let* $d \ge 2$ *and* $\frac{\ln d}{d} \le \varepsilon \le 1/2$. *Then*

$$N(\varepsilon, d) \le \frac{C\ln d}{\varepsilon^2}\ \ln\left(\frac{1}{\varepsilon}\right).$$

# New bounds.

Hinrichs–Krieg–Kunsch–Rudolf (**20**): for a random choice of points

$$\max\left\{\frac{c}{\varepsilon}\ln\left(\frac{1}{\varepsilon}\right),\ \frac{d}{2\varepsilon}\right\} \le N_{ran}(\varepsilon, d) \le \max\left\{\frac{C\,\ln d}{\varepsilon}\ \ln\left(\frac{1}{\varepsilon}\right),\ \frac{C\,d}{\varepsilon}\ \ln\ln\left(\frac{2}{\varepsilon}\right)\right\}$$

## Theorem (relatively large $\varepsilon$)

*Let* $d \ge 2$ *and* $\frac{\ln d}{d} \le \varepsilon \le 1/2$. *Then*

$$N(\varepsilon, d) \le \frac{C\,\ln d}{\varepsilon^2}\ \ln\left(\frac{1}{\varepsilon}\right).$$

The bound here is better when $\varepsilon \ge \dfrac{\ln^2 d}{d\ln\ln(2d)}$.

# New bounds.

Hinrichs–Krieg–Kunsch–Rudolf (**20**): for a random choice of points

$$\max\left\{\frac{c}{\varepsilon}\ln\left(\frac{1}{\varepsilon}\right), \frac{d}{2\varepsilon}\right\} \le N_{ran}(\varepsilon, d) \le \max\left\{\frac{C\,\ln d}{\varepsilon}\ \ln\left(\frac{1}{\varepsilon}\right), \frac{C\,d}{\varepsilon}\ \ln\ln\left(\frac{2}{\varepsilon}\right)\right\}$$

**Theorem (relatively large $\varepsilon$)**

*Let* $d \ge 2$ *and* $\frac{\ln d}{d} \le \varepsilon \le 1/2$. *Then*

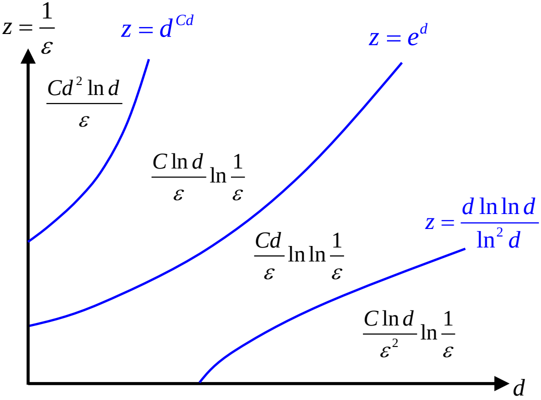
$$N(\varepsilon, d) \le \frac{C\,\ln d}{\varepsilon^2}\ \ln\left(\frac{1}{\varepsilon}\right).$$

The bound here is better when $\varepsilon \ge \dfrac{\ln^2 d}{d\ln\ln(2d)}$.

The prove also uses random points, but one needs to adjust the distribution.

# State of the art.

$$N(\varepsilon, d) \leq \begin{cases} \frac{C\,\ln d}{\varepsilon^2}\,\ln\left(\frac{1}{\varepsilon}\right), & \text{if } \varepsilon \geq \frac{\ln^2 d}{d \ln\ln(2d)}, \\ \frac{C\,d}{\varepsilon}\,\ln\ln\left(\frac{1}{\varepsilon}\right), & \text{if } \frac{\ln^2 d}{d\ln\ln(2d)} \geq \varepsilon \geq e^{-d}, \\ \frac{C\ln d}{\varepsilon}\,\ln\left(\frac{1}{\varepsilon}\right), & \text{if } e^{-d} \geq \varepsilon \geq \exp(-Cd\ln d), \\ \frac{C\,d^2\ln d}{\varepsilon}, & \text{if } \varepsilon \leq \exp(-Cd\ln d). \end{cases}$$



# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

$$\textit{each rectangle in } \mathcal{N} \textit{ contains a point from } P \quad \Longrightarrow$$

$$\textit{each rectangle in } \mathcal{R}_d \textit{ of volume at least } \varepsilon \textit{ contains a point from } P.$$

# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

$$\textit{each rectangle in } \mathcal{N} \textit{ contains a point from } P \qquad \Longrightarrow$$

$$\textit{each rectangle in } \mathcal{R}_d \textit{ of volume at least } \varepsilon \textit{ contains a point from } P.$$

Standardly, using union bound, one estimates the probability of the "bad" event

$$\textit{there exists a rectangle in } \mathcal{N} \textit{ containing no points from } P$$

as the sum of "bad" probabilities of individual events,

$$\textit{a given rectangle in } \mathcal{N} \textit{ containing no points from } P.$$

# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

*each rectangle in $\mathcal{N}$ contains a point from $P$* $\quad\Longrightarrow$

*each rectangle in $\mathcal{R}_d$ of volume at least $\varepsilon$ contains a point from $P$.*

Standardly, using union bound, one estimates the probability of the "bad" event

*there exists a rectangle in $\mathcal{N}$ containing no points from $P$*

as the sum of "bad" probabilities of individual events,

*a given rectangle in $\mathcal{N}$ containing no points from $P$.*

Individual bounds are simple – volume computations.

# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

$$\textit{each rectangle in } \mathcal{N} \textit{ contains a point from } P \qquad \Longrightarrow$$

$$\textit{each rectangle in } \mathcal{R}_d \textit{ of volume at least } \varepsilon \textit{ contains a point from } P.$$

Standardly, using union bound, one estimates the probability of the "bad" event

$$\textit{there exists a rectangle in } \mathcal{N} \textit{ containing no points from } P$$

as the sum of "bad" probabilities of individual events,

$$\textit{a given rectangle in } \mathcal{N} \textit{ containing no points from } P.$$

Individual bounds are simple – volume computations.

Thus the main difficulty is to construct the set $\mathcal{N}$ of not too large cardinality.

# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

$$\textit{each rectangle in } \mathcal{N} \textit{ contains a point from } P \quad \Longrightarrow$$

$$\textit{each rectangle in } \mathcal{R}_d \textit{ of volume at least } \varepsilon \textit{ contains a point from } P.$$

Standardly, using union bound, one estimates the probability of the "bad" event

$$\textit{there exists a rectangle in } \mathcal{N} \textit{ containing no points from } P$$

as the sum of "bad" probabilities of individual events,

$$\textit{a given rectangle in } \mathcal{N} \textit{ containing no points from } P.$$

Individual bounds are simple – volume computations.

Thus the main difficulty is to construct the set $\mathcal{N}$ of not too large cardinality.

Rudolf used the concept of $\delta$-cover to construct $\mathcal{N}$ (and bounds due to Gnewuch).

# Some ideas of the proof.

Consider a set $P$ of random points independently and uniformly distributed in the cube. To show that (with good probability) every box of volume $\varepsilon$ contains at least one point from $P$, we need to construct a set $\mathcal{N}$ of "test boxes." It should satisfy

$$\textit{each rectangle in } \mathcal{N} \textit{ contains a point from } P \quad \Longrightarrow$$

$$\textit{each rectangle in } \mathcal{R}_d \textit{ of volume at least } \varepsilon \textit{ contains a point from } P.$$

Standardly, using union bound, one estimates the probability of the "bad" event

*there exists a rectangle in $\mathcal{N}$ containing no points from $P$*

as the sum of "bad" probabilities of individual events,

*a given rectangle in $\mathcal{N}$ containing no points from $P$.*

Individual bounds are simple – volume computations.

Thus the main difficulty is to construct the set $\mathcal{N}$ of not too large cardinality.

Rudolf used the concept of $\delta$-cover to construct $\mathcal{N}$ (and bounds due to Gnewuch).

We use a different approach, which fits better this problem.

# Some ideas of the proof.

Denote

$$\mathcal{B}_{\varepsilon,d} := \Big\{ B \in \mathcal{R}_d \ \mid \ |B| \geq \varepsilon \Big\}.$$

**Definition.** *We say that* $\mathcal{N} \subset \mathcal{R}_d$ *is a* $\delta$*-net for* $\mathcal{B}_{\varepsilon,d}$ *if*

$$\forall\, B \in \mathcal{B}_{\varepsilon,d} \ \exists\, B_0 \in \mathcal{N} : \quad B_0 \subset B \quad \text{and} \quad |B_0| \geq (1-\delta)|B|.$$

# Some ideas of the proof.

Denote

$$\mathcal{B}_{\varepsilon,d} := \Big\{ B \in \mathcal{R}_d \mid |B| \geq \varepsilon \Big\}.$$

**Definition.** *We say that* $\mathcal{N} \subset \mathcal{R}_d$ *is a* $\delta$*-net for* $\mathcal{B}_{\varepsilon,d}$ *if*

$$\forall B \in \mathcal{B}_{\varepsilon,d} \;\exists B_0 \in \mathcal{N}: \quad B_0 \subset B \quad \textit{and} \quad |B_0| \geq (1-\delta)|B|.$$

A variant of the following lemma using random points and the union bound was proved by Rudolf.

## Lemma (size of a net)

*Let* $d \geq 1$ *and* $\varepsilon, \delta \in (0,1)$. *Let* $\mathcal{N}$ *be a* $\delta$*-net for* $\mathcal{B}_{\varepsilon,d}$ *with* $|\mathcal{N}| \geq 3$. *Then with probability at least* $1 - 1/|\mathcal{N}|$

$$N = \left\lfloor \frac{3 \ln |\mathcal{N}|}{(1-\delta)\varepsilon} \right\rfloor.$$

*a random choice of N points satisfies the desire property.*

# Proof of the lemma.

Consider $N$ independent random points $X_1, ..., X_N$ uniformly drawn from $[0,1]^d$. It is enough to show that

$$\forall B \in \mathcal{N} \quad \text{with } |B| \geq v = (1-\delta)\varepsilon \;\; \exists j \leq N : \quad X_j \in B.$$

# Proof of the lemma.

Consider $N$ independent random points $X_1, ..., X_N$ uniformly drawn from $[0, 1]^d$.
It is enough to show that

$$\forall B \in \mathcal{N} \quad \text{with } |B| \geq v = (1-\delta)\varepsilon \;\; \exists j \leq N : \quad X_j \in B.$$

Fix such a box $B$. Using independence of $X_j$'s,

$$\mathbb{P}\left(\{\forall j \leq N : \;\; X_j \notin B\}\right) = (1-v)^N < \exp(-vN).$$

# Proof of the lemma.

Consider $N$ independent random points $X_1, ..., X_N$ uniformly drawn from $[0, 1]^d$. It is enough to show that

$$\forall B \in \mathcal{N} \quad \text{with } |B| \geq v = (1-\delta)\varepsilon \ \ \exists j \leq N : \quad X_j \in B.$$

Fix such a box $B$. Using independence of $X_j$'s,

$$\mathbb{P}\left(\{\forall j \leq N : \ \ X_j \notin B\}\right) = (1-v)^N < \exp(-vN).$$

Therefore, by the union bound,

$$\mathbb{P}\left(\{\exists B \in \mathcal{N} : \ \ |B| \geq v \quad \text{and} \quad \forall j \leq N : \ \ X_j \notin B\}\right) < |\mathcal{N}| \exp(-vN).$$

# Proof of the lemma.

Consider $N$ independent random points $X_1, ..., X_N$ uniformly drawn from $[0, 1]^d$. It is enough to show that

$$\forall B \in \mathcal{N} \quad \text{with } |B| \geq v = (1-\delta)\varepsilon \;\; \exists j \leq N: \quad X_j \in B.$$

Fix such a box $B$. Using independence of $X_j$'s,

$$\mathbb{P}\left(\{\forall j \leq N: \;\; X_j \notin B\}\right) = (1-v)^N < \exp(-vN).$$

Therefore, by the union bound,

$$\mathbb{P}\left(\{\exists B \in \mathcal{N}: \;\; |B| \geq v \quad \text{and} \quad \forall j \leq N: \;\; X_j \notin B\}\right) < |\mathcal{N}| \exp(-vN).$$

Thus, as far as

$$|\mathcal{N}| \exp(-vN) \leq 1$$

there exists a realization of $X_j$'s with the desired property.

# Construction of a net.

# Dispersion on the torus.

We consider periodic axis parallel boxes, that is, boxes of the form

$$\prod_{i=1}^{d} I_i(a_i, b_i), \quad a_i, b_i \in [0, 1],$$

where

$$I_i(a, b) := \begin{cases} (a_i, b_i), & \text{whenever } 0 \le a_i < b_i \le 1, \\ [0, 1] \setminus [b_i, a_i], & \text{whenever } 0 \le b_i < a_i \le 1. \end{cases}$$

# Dispersion on the torus.

We consider periodic axis parallel boxes, that is, boxes of the form

$$\prod_{i=1}^{d} I_i(a_i, b_i), \quad a_i, b_i \in [0,1],$$

where

$$I_i(a,b) := \begin{cases} (a_i, b_i), & \text{whenever } 0 \le a_i < b_i \le 1, \\ [0,1] \setminus [b_i, a_i], & \text{whenever } 0 \le b_i < a_i \le 1. \end{cases}$$

Combining recent results of M. Ullrich (**18**) and Rudolf (**18**)

$$\frac{d}{\varepsilon} \le \widetilde{N}(\varepsilon, d) \le \frac{8d}{\varepsilon} \left( \ln d + \ln \left( \frac{8}{\varepsilon} \right) \right).$$

# Dispersion on the torus.

We consider periodic axis parallel boxes, that is, boxes of the form

$$\prod_{i=1}^{d} I_i(a_i, b_i), \quad a_i, b_i \in [0, 1],$$

where

$$I_i(a, b) := \begin{cases} (a_i, b_i), & \text{whenever } 0 \le a_i < b_i \le 1, \\ [0, 1] \setminus [b_i, a_i], & \text{whenever } 0 \le b_i < a_i \le 1. \end{cases}$$

Combining recent results of M. Ullrich (**18**) and Rudolf (**18**)

$$\frac{d}{\varepsilon} \le \widetilde{N}(\varepsilon, d) \le \frac{8d}{\varepsilon} \left( \ln d + \ln \left( \frac{8}{\varepsilon} \right) \right).$$

Note that the VC dimension of the set of periodic axis parallel boxes is of the order $d \ln d$, therefore the Blumer–Ehrenfeucht–Haussler–Warmuth result leads to

$$\widetilde{N}(\varepsilon, d) \le \frac{8d \ln d}{\varepsilon} \ln \left( \frac{8}{\varepsilon} \right)$$

— worse than the Rudolf bound.

# Dispersion on the torus.

$$\frac{d}{\varepsilon} \leq \widetilde{N}(\varepsilon, d) \leq \frac{8d}{\varepsilon} \left( \ln d + \ln \left( \frac{8}{\varepsilon} \right) \right).$$

We improve the Rudolf upper bound in the case $\varepsilon \leq 1/d^C$.

## Theorem (bounds in the periodic case)

*Let* $d \geq 2$ *and* $\varepsilon \in (0, 1/2]$. *Then*

(*i*) $\widetilde{N}(\varepsilon, d) \leq \frac{C \ln d}{\varepsilon} \ln \left( \frac{1}{\varepsilon} \right)$, *provided that* $\varepsilon \leq \exp(-d)$,

(*ii*) $\widetilde{N}(\varepsilon, d) \leq \frac{C\, d\, \ln d}{\varepsilon}$, *provided that* $\varepsilon \geq \exp(-d)$.

*Moreover, the random choice of points uniformly distributed on* $[0, 1]^d$ *works.*