Solvable Integration Problems and Optimal Sample Size Selection

Robert J. Kunsch
(joint work with Erich Novak and Daniel Rudolf)

UNIVERSITÄT OSNABRÜCK

Special Session: When to stop a simulation

MCQMC2018, mercredi, le 4 julliet 2018, Rennes

Standard Monte Carlo integration

Task: Compute $\mathbb{E}\, Y$ from random samples $Y_1, Y_2, \ldots \overset{\text{iid}}{\sim} Y$.

Standard Monte Carlo method:

$$\widehat{M}_n(Y) := \frac{1}{n} \sum_{i=1}^{n} Y_i$$

Application to integration w.r.t. prob. measure π on domain G, using $X, X_1, X_2, \ldots \overset{\text{iid}}{\sim} \pi$:

$$\mathrm{INT}\, f := \int_G f \,\mathrm{d}\pi \;\approx\; \widehat{M}_n(f(X)) \;=\; \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

Standard Monte Carlo error / root mean square error (RMSE):

$$e(\widehat{M}_n, Y) := \sqrt{\mathbb{E}\left(\widehat{M}_n(Y) - \mathbb{E}\, Y\right)^2} = \frac{\sigma(Y)}{\sqrt{n}}$$

Problem: What if variance $\sigma^2(Y)$ unknown?

Integration for a-priori unknown variance

Theorem (Hickernell, Jiang, Liu, Owen 2013)

Given $K > 1$, then for any error threshold $\varepsilon > 0$ and any uncertainty level $\delta > 0$, there exists an algorithm $A^K_{\varepsilon,\delta}$ such that

$$\mathbb{P}\left\{ |A^K_{\varepsilon,\delta}(Y) - \mathbb{E}\, Y| > \varepsilon \right\} \leq \delta$$

for all Y with bounded kurtosis, $\|Y - \mathbb{E}\, Y\|_4 \leq K \|Y - \mathbb{E}\, Y\|_2$.

Idea:

1. use fixed sample $Y_1, \ldots, Y_{n_1}$ for a variance estimate
2. choose size n_2 of sample $Y_{n_1+1}, \ldots, Y_{n_1+n_2}$ for mean estimate

There is no solution with fixed sample size!

Question: Optimality? Other assumptions? Mean squared error?

Integration for a-priori unknown variance

Theorem (Hickernell, Jiang, Liu, Owen 2013)

Given $K > 1$, then for any error threshold $\varepsilon > 0$ and any uncertainty level $\delta > 0$, there exists an algorithm $A^K_{\varepsilon,\delta}$ such that

$$\mathbb{P}\left\{|A^K_{\varepsilon,\delta}(Y) - \mathbb{E}\,Y| > \varepsilon\right\} \leq \delta$$

for all Y with bounded kurtosis, $\|Y - \mathbb{E}\,Y\|_4 \leq K\|Y - \mathbb{E}\,Y\|_2$.

Idea:

1. use fixed sample $Y_1, \ldots, Y_{n_1}$ for a variance estimate
2. choose size n_2 of sample $Y_{n_1+1}, \ldots, Y_{n_1+n_2}$ for mean estimate

There is no solution with fixed sample size!

Question: Optimality? Other assumptions? ~~Mean squared error?~~

Definition of solvability

Inputset $\mathcal{F} \subset L_1(\pi)$ — class of integrands f
with respect to prob. measure π on domain G

General randomized algorithm = mapping $A : \Omega \times \mathcal{F} \to \mathbb{R}$,
requiring only finitely many function evaluations, $n(\omega, f) < \infty$.

Special type: **i.i.d.-based algorithms** using $f(X_i)$ where $X_i \stackrel{\text{iid}}{\sim} \pi$

Definition (Solvable integration problems)

The integration problem is solvable for $\mathcal{F}$
iff for all $\varepsilon, \delta > 0$ exists an (ε, δ)-approximating algorithm A, i.e.

$$\mathbb{P}\{|A(f) - \mathrm{INT}\, f| > \varepsilon\} \leq \delta \qquad \text{for all } f \in \mathcal{F}.$$

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

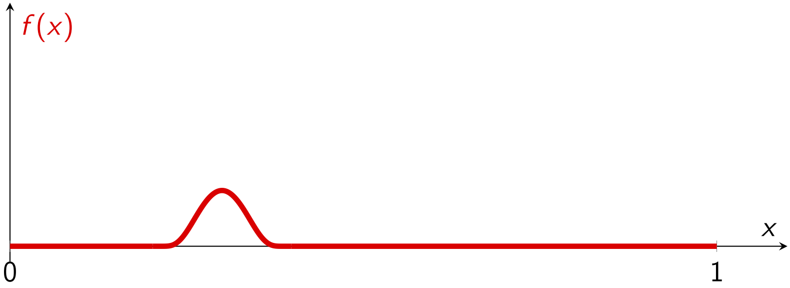
Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!



Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

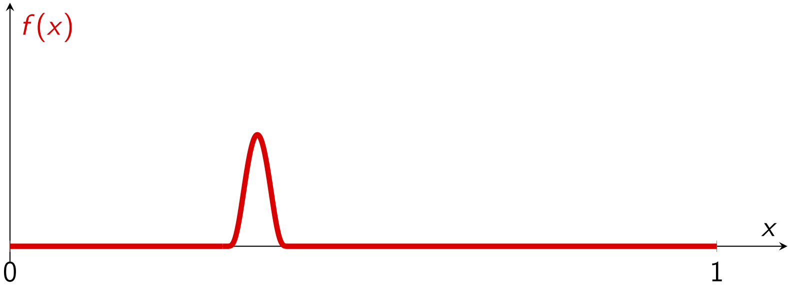
Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

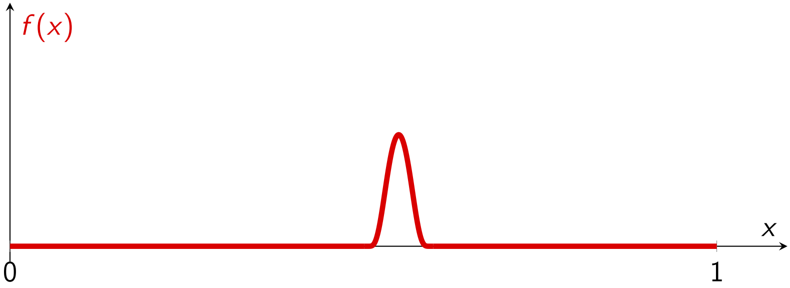


Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$
from functions f with huge integral but small support? Impossible!



Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

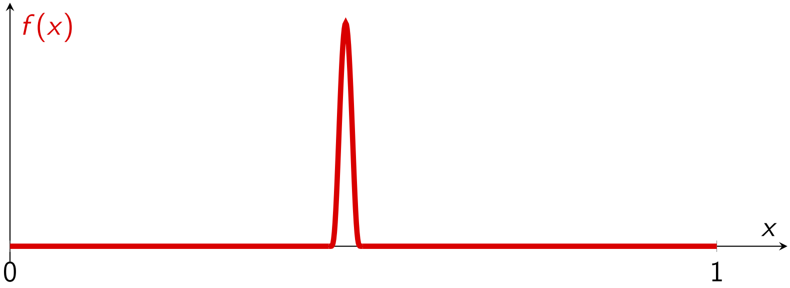
Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!



Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Example of an unsolvable problem

Theorem (K., Novak, Rudolf 2018)

The integration problem $\int_0^1 f \, \mathrm{d}x$
is unsolvable for the class $\mathcal{F} = C^\infty([0,1])$.

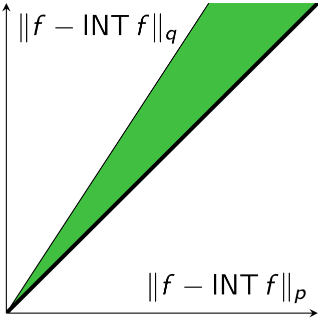
Idea of the proof: How would you distinguish $f = 0$ from functions f with huge integral but small support? Impossible!

Generalizing the kurtosis – a cone-shaped input set

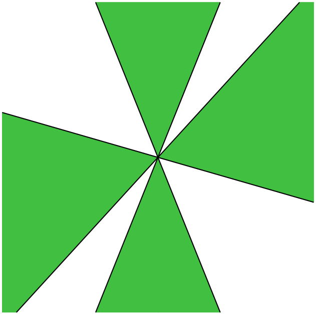
For $1 \leq p < q \leq \infty$ and $K > 1$, (Hickernell et al.: $p = 2$, $q = 4$)

$$\mathcal{F}_{p,q,K} := \{f \in L_q(\pi) \colon \|f - \mathrm{INT}\, f\|_q \leq K \|f - \mathrm{INT}\, f\|_p\}$$

Scaling property: $f \in \mathcal{F}_{p,q,K} \quad \Rightarrow \quad cf \in \mathcal{F}_{p,q,K}$ for all $c \in \mathbb{R}$.



cone condition



within 2-dim. subspace of L_q

Upper bound for the expected cost

Recall input set for $1 \leq p < q \leq \infty$ and $K > 1$,

$$\mathcal{F}_{p,q,K} := \{f \in L_q(\pi) \colon \|f - \mathrm{INT}\, f\|_q \leq K \|f - \mathrm{INT}\, f\|_p\}$$

Theorem (K, Novak, Rudolf 2018)

For any error threshold $\varepsilon > 0$ and uncertainty level $\delta > 0$, there exists an (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}^{p,q,K}$ on $\mathcal{F}_{p,q,K}$. Writing $\widetilde{q} := \min\{q, 2\}$, the expected sample size $\bar{n}$ is bounded by

$$\bar{n}(A_{\varepsilon,\delta}^{p,q,K}, f) := \mathbb{E}\, n(\omega, f)$$

$$\leq \left[C_q\, K^{pq/(q-p)} + C_{q,K} \left(\frac{\|f - \mathrm{INT}\, f\|_1}{\varepsilon} \right)^{1+1/(\widetilde{q}-1)} \right] \log \frac{1}{\delta} .$$

Structure of the algorithm

Parameters: odd $k, k' \in \mathbb{N}$; $2 \leq m \in \mathbb{N}$; $1 < \eta \in \mathbb{R}$.

Information: random samples $Y_i := f(X_i)$ with $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \pi$

1. estimate $\|f - \mathrm{INT}\, f\|_1$ from $n_1 := k \cdot m$ observations:

$$\widehat{R}_{k,m} := \mathrm{med}(\widehat{R}_m^{(1)}, \ldots, \widehat{R}_m^{(k)}), \quad \text{where}$$

$$\widehat{R}_m^{(\ell)} \overset{\text{iid}}{\sim} \frac{1}{m} \sum_{i=1}^{m} |Y_i - \widehat{M}_m|, \quad \text{with} \quad \widehat{M}_m := \frac{1}{m} \sum_{i=1}^{m} Y_i$$

2. estimate $\mathrm{INT}\, f$ via $n_2 := k' \cdot m'$ additional observations, $m' := \max\{\lceil \eta\, (\widehat{R}_{k,m})^{1+1/(\tilde{q}-1)} \rceil, 1\}$, with $\tilde{q} := \min\{q, 2\}$,

$$A_{k,m}^{k',\eta,\tilde{q}}(f) := \mathrm{med}(\widetilde{M}_{m'}^{(1)}, \ldots, \widetilde{M}_{m'}^{(k')}), \quad \text{where} \quad \widetilde{M}_{m'}^{(\ell)} \overset{\text{iid}}{\sim} \widehat{M}_{m'}$$

Median trick $\rightsquigarrow$ probability amplification, finite expected cost

Alternative for 1 on $\mathcal{F}_{p,q,K}$: if $q \leq 2$, estimate $\|f - \mathrm{INT}\, f\|_p^p$,
if $2 < q$, estimate $\|f - \mathrm{INT}\, f\|_2^2$.

Structure of the algorithm

Parameters: odd $k, k' \in \mathbb{N}$; $2 \leq m \in \mathbb{N}$; $1 < \eta \in \mathbb{R}$.

Information: random samples $Y_i := f(X_i)$ with $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \pi$

1. estimate $\|f - \mathrm{INT}\, f\|_1$ from $n_1 := k \cdot m$ observations:

$$\widehat{R}_{k,m} := \mathrm{med}(\widehat{R}_m^{(1)}, \ldots, \widehat{R}_m^{(k)}), \quad \text{where}$$

$$\widehat{R}_m^{(\ell)} \overset{\text{iid}}{\sim} \frac{1}{m} \sum_{i=1}^{m} |Y_i - \widehat{M}_m|, \quad \text{with} \quad \widehat{M}_m := \frac{1}{m} \sum_{i=1}^{m} Y_i$$

2. estimate $\mathrm{INT}\, f$ via $n_2 := k' \cdot m'$ additional observations, $m' := \max\{\lceil \eta\, (\widehat{R}_{k,m})^{1+1/(\widetilde{q}-1)} \rceil, 1\}$, with $\widetilde{q} := \min\{q, 2\}$,

$$A_{k,m}^{k',\eta,\widetilde{q}}(f) := \mathrm{med}(\widetilde{M}_{m'}^{(1)}, \ldots, \widetilde{M}_{m'}^{(k')}), \quad \text{where} \quad \widetilde{M}_{m'}^{(\ell)} \overset{\text{iid}}{\sim} \widehat{M}_{m'}$$

Median trick $\rightsquigarrow$ probability amplification, finite expected cost

Alternative for 1 **on** $\mathcal{F}_{p,q,K}$: if $q \leq 2$, estimate $\|f - \mathrm{INT}\, f\|_p^p$,
if $2 < q$, estimate $\|f - \mathrm{INT}\, f\|_2^2$.

An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε,δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon,\delta,\mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f\in\mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta},f) \geq \frac{1}{2\log 2}\, K^{pq/(q-p)} \log\frac{1}{2\delta}\,.$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f+h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h=c\} = 1-\pi\{h=0\} \in (0,\frac{1}{2})$?

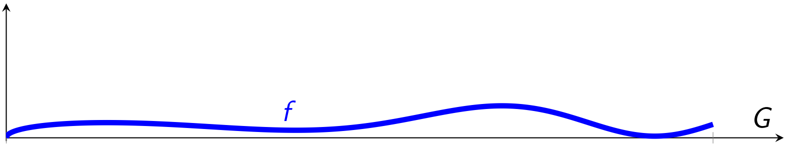
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε,δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon,\delta,\mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f\in\mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta},f)$$
$$\geq \frac{1}{2\log 2} K^{pq/(q-p)} \log\frac{1}{2\delta}\,.$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f+h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h=c\} = 1-\pi\{h=0\} \in (0,\frac{1}{2})$?



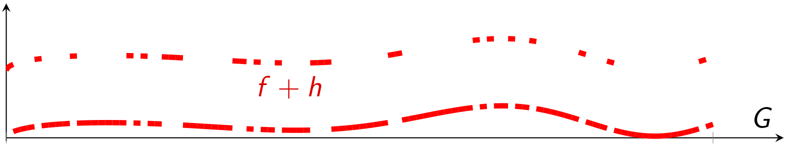
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon, \delta, \mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f \in \mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta}, f)$$
$$\geq \frac{1}{2 \log 2} K^{pq/(q-p)} \log \frac{1}{2\delta} .$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f + h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h = c\} = 1 - \pi\{h = 0\} \in (0, \frac{1}{2})$?



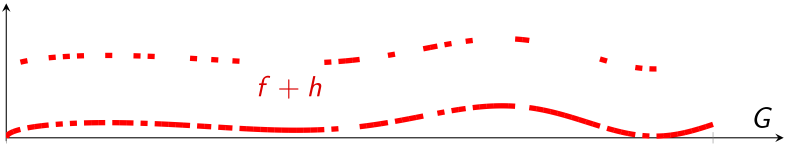
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon, \delta, \mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f \in \mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta}, f) \geq \frac{1}{2 \log 2} K^{pq/(q-p)} \log \frac{1}{2\delta} .$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f + h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h = c\} = 1 - \pi\{h = 0\} \in (0, \frac{1}{2})$?



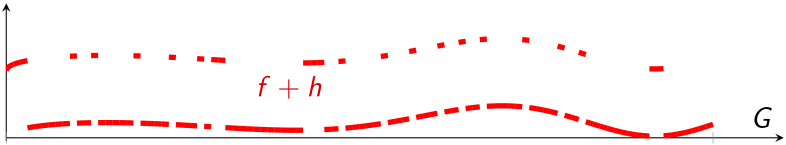
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\text{fix}}(\varepsilon, \delta, \mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f \in \mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta}, f) \geq \frac{1}{2 \log 2} K^{pq/(q-p)} \log \frac{1}{2\delta}.$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f + h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h = c\} = 1 - \pi\{h = 0\} \in (0, \frac{1}{2})$?



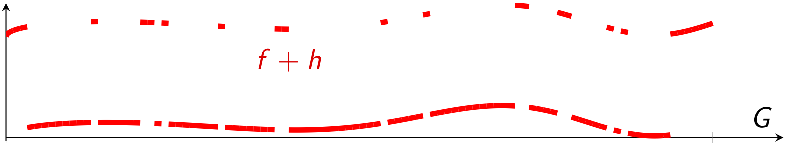
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\text{fix}}(\varepsilon, \delta, \mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f \in \mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta}, f) \geq \frac{1}{2 \log 2} K^{pq/(q-p)} \log \frac{1}{2\delta}\,.$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f + h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h = c\} = 1 - \pi\{h = 0\} \in (0, \frac{1}{2})$?



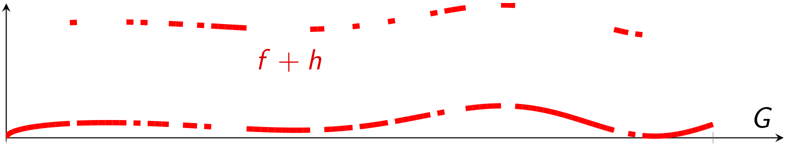
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε,δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon,\delta,\mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f\in\mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta},f) \geq \frac{1}{2\log 2} K^{pq/(q-p)} \log\frac{1}{2\delta}\,.$$

Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f+h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h=c\} = 1-\pi\{h=0\} \in (0,\frac{1}{2})$?



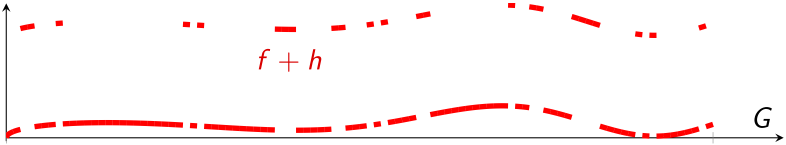
An absolute lower bound – the fixed cost

Theorem (K, Novak, Rudolf 2018)

Any (ε,δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$ requires a minimal amount of information, irrespective of the input,

$$\bar{n}^{\mathrm{fix}}(\varepsilon,\delta,\mathcal{F}_{p,q,K}) := \inf_{A_{\varepsilon,\delta}} \inf_{f\in\mathcal{F}_{p,q,K}} \bar{n}(A_{\varepsilon,\delta},f) \geq \frac{1}{2\log 2} K^{pq/(q-p)} \log\frac{1}{2\delta}\,.$$

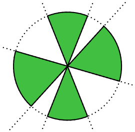
Proof: Distinguish $f \in \mathcal{F}_{p,q,K}$ from perturbed $f+h \in \mathcal{F}_{p,q,K}$, where $h \in \mathcal{F}_{p,q,K}$ with $\pi\{h=c\} = 1-\pi\{h=0\} \in (0,\frac{1}{2})$?



Worst case cost for bounded norm

Put $\widetilde{q} := \min\{q, 2\}$, we intersect $\mathcal{F}_{p,q,K}$ with

$$\tau\mathcal{B}_{\widetilde{q}} := \{f \in L_{\widetilde{q}}(\pi) : \|f - \mathrm{INT}\, f\|_{\widetilde{q}} \leq \tau\}$$



Theorem (K, Novak, Rudolf 2018)

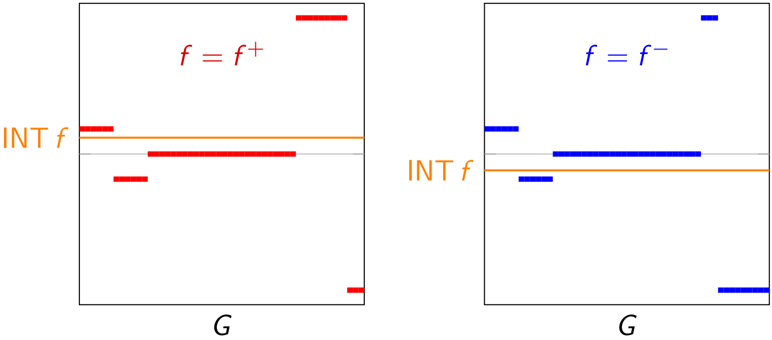
For any (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}$ on $\mathcal{F}_{p,q,K}$, the worst case cost for $\|f - \mathrm{INT}\, f\|_{\widetilde{q}} \leq \tau$ is bounded below by

$$\bar{n}^{\mathrm{wor}}(\varepsilon, \delta, \mathcal{F}_{p,q,K} \cap \tau\mathcal{B}_{\widetilde{q}}) := \inf_{A_{\varepsilon,\delta}} \sup_{f \in \mathcal{F}_{p,q,K} \cap \tau\mathcal{B}_{\widetilde{q}}} \bar{n}(A_{\varepsilon,\delta}, f)$$
$$\geq c_{q,K} \left(\frac{\tau}{\varepsilon}\right)^{1+1/(\widetilde{q}-1)} \log \frac{3}{4\delta}$$

for error thresholds $0 < \varepsilon < C_K \tau$ and uncertainty $0 < \delta \leq \frac{1}{4}$.

Worst case cost for bounded norm – proof

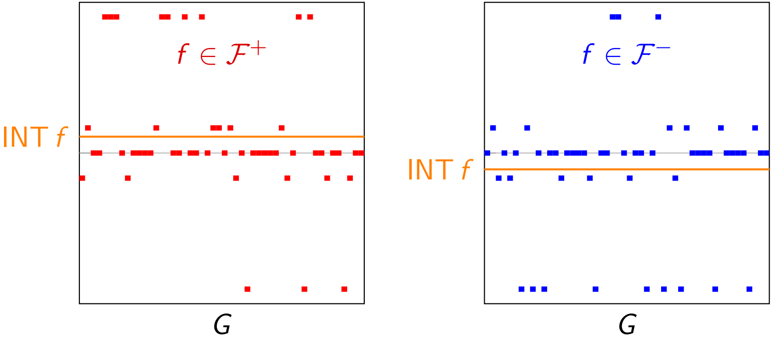
Worst case lower bounds on $\mathcal{F}_{p,q,K} \cap \tau\mathcal{B}_{\tilde{q}}$ using i.i.d. $Y_i := f(X_i)$
via cost bounds for statistical tests (Wald 1945) $\rightsquigarrow$ f^+ vs. f^-?



Same lower bounds for general algorithms using function values,
if π has no atoms $\rightsquigarrow$ distinguish shuffled classes $\mathcal{F}^+$ and $\mathcal{F}^-$

Worst case cost for bounded norm – proof

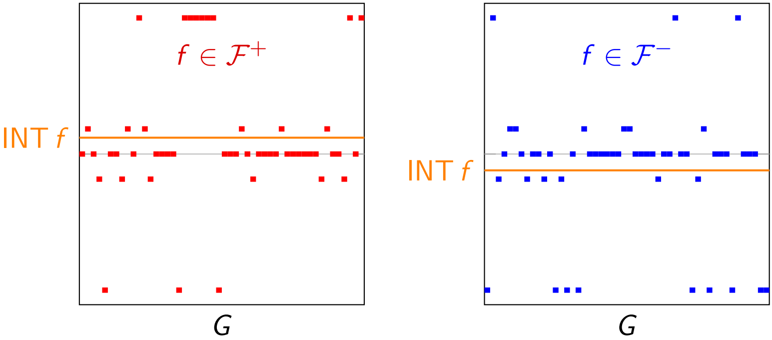
Worst case lower bounds on $\mathcal{F}_{p,q,K} \cap \tau\mathcal{B}_{\widetilde{q}}$ using i.i.d. $Y_i := f(X_i)$ via cost bounds for statistical tests (Wald 1945) $\rightsquigarrow$ f^+ vs. f^-?



Same lower bounds for general algorithms using function values, if π has no atoms $\rightsquigarrow$ distinguish shuffled classes $\mathcal{F}^+$ and $\mathcal{F}^-$

Worst case cost for bounded norm – proof

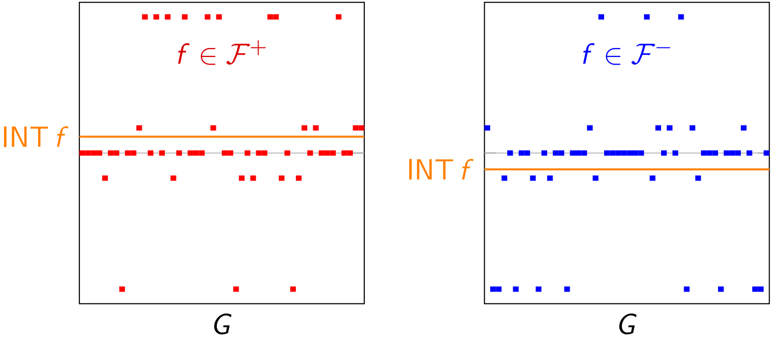
Worst case lower bounds on $\mathcal{F}_{p,q,K} \cap \tau \mathcal{B}_{\widetilde{q}}$ using i.i.d. $Y_i := f(X_i)$ via cost bounds for statistical tests (Wald 1945) $\rightsquigarrow$ f^+ vs. f^-?



Same lower bounds for general algorithms using function values, if π has no atoms $\rightsquigarrow$ distinguish shuffled classes $\mathcal{F}^+$ and $\mathcal{F}^-$

Worst case cost for bounded norm – proof

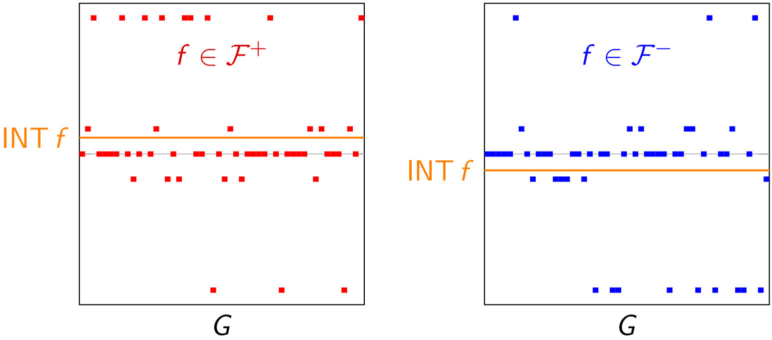
Worst case lower bounds on $\mathcal{F}_{p,q,K} \cap \tau \mathcal{B}_{\widetilde{q}}$ using i.i.d. $Y_i := f(X_i)$ via cost bounds for statistical tests (Wald 1945) $\rightsquigarrow$ f^+ vs. f^-?



Same lower bounds for general algorithms using function values, if π has no atoms $\rightsquigarrow$ distinguish shuffled classes $\mathcal{F}^+$ and $\mathcal{F}^-$

Worst case cost for bounded norm – proof

Worst case lower bounds on $\mathcal{F}_{p,q,K} \cap \tau\mathcal{B}_{\widetilde{q}}$ using i.i.d. $Y_i := f(X_i)$ via cost bounds for statistical tests (Wald 1945) $\rightsquigarrow$ f^+ vs. f^-?



Same lower bounds for general algorithms using function values, if π has no atoms $\rightsquigarrow$ distinguish shuffled classes $\mathcal{F}^+$ and $\mathcal{F}^-$

Summary

Input set for $1 \le p < q \le \infty$ and $K > 1$, $\tilde{q} := \min\{q, 2\}$

$$\mathcal{F}_{p,q,K} := \{f \in L_q(\pi) \colon \|f - \mathrm{INT}\, f\|_q \le K \|f - \mathrm{INT}\, f\|_p\}$$

Upper bound for (ε, δ)-approximating algorithm $A_{\varepsilon,\delta}^{p,q,K}$ on $\mathcal{F}_{p,q,K}$

$$\bar{n}(A_{\varepsilon,\delta}^{p,q,K}, f) \le C_{q,K} \left[1 + \left(\frac{\|f - \mathrm{INT}\, f\|_1}{\varepsilon}\right)^{1+1/(\tilde{q}-1)}\right] \log \frac{1}{\delta}$$

Fixed cost – an absolute lower bound

$$\bar{n}^{\mathrm{fix}}(\varepsilon, \delta, \mathcal{F}_{p,q,K}) \ge \frac{1}{2 \log 2} K^{pq/(q-p)} \log \frac{1}{2\delta}$$

Worst case cost for $\|f - \mathrm{INT}\, f\|_{\tilde{q}} \le \tau$,

$$\bar{n}^{\mathrm{wor}}(\varepsilon, \delta, \mathcal{F}_{p,q,K} \cap \tau \mathcal{B}_{\tilde{q}}) \ge c_{q,K} \left(\frac{\tau}{\varepsilon}\right)^{1+1/(\tilde{q}-1)} \log \frac{3}{4\delta}$$

References

F.J. Hickernell, L. Jiang, Y.W. Liu, A. Owen.
Guaranteed conservative fixed width confidence intervals via Monte Carlo sampling.
Proceedings of the MCQMC 2012, Springer, 2013.

R.J. Kunsch, E. Novak, D. Rudolf.
Solvable integration problems and optimal sample size selection.
Avaible on arXiv:1805.08637 [math.NA], 2018.

A. Wald.
Sequential tests of statistical hypotheses.
Annals of Mathematical Statistics, 16:117–186, 1945.