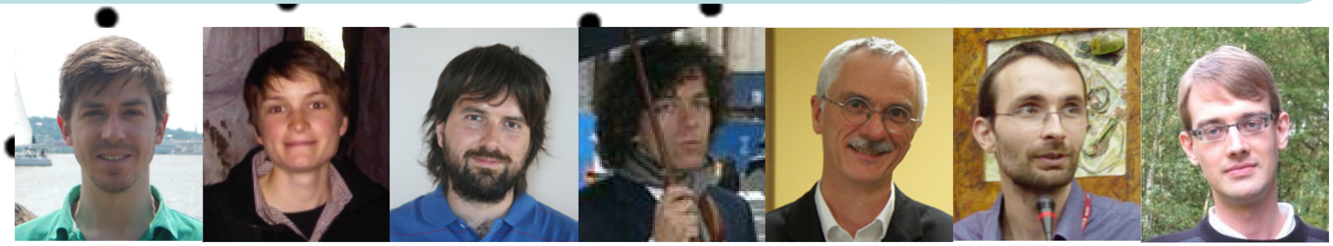
Constructing low star discrepancy point sets with genetic algorithms

François-Michel De Rainville, Carola Doerr,
Christian Gagné, Michael Gnewuch,
Denis Laurendeau, Olivier Teytaud, Magnus Wahlström



Numerical Integration

- One of the most challenging questions in numerical analysis: compute $\int f(x)\mathrm{d}x$ for a (possibly complicated) function $f\colon \mathbb{R}^d \to \mathbb{R}$
- FAR from being a purely academic problem: applications in financial derivate pricing, scenario reduction, computer graphics, pseudo-random number generators, stochastic programming...
- One of the oldest problems in mathematics

Monte Carlo Integration

- Instead of computing $\int f(x)\mathrm{d}x$, evaluate f in random samples
- Approximate the integral by the mean value $\frac{1}{n}\sum_{i=1}^{n} f(x^i)$
- How good is this approximation?

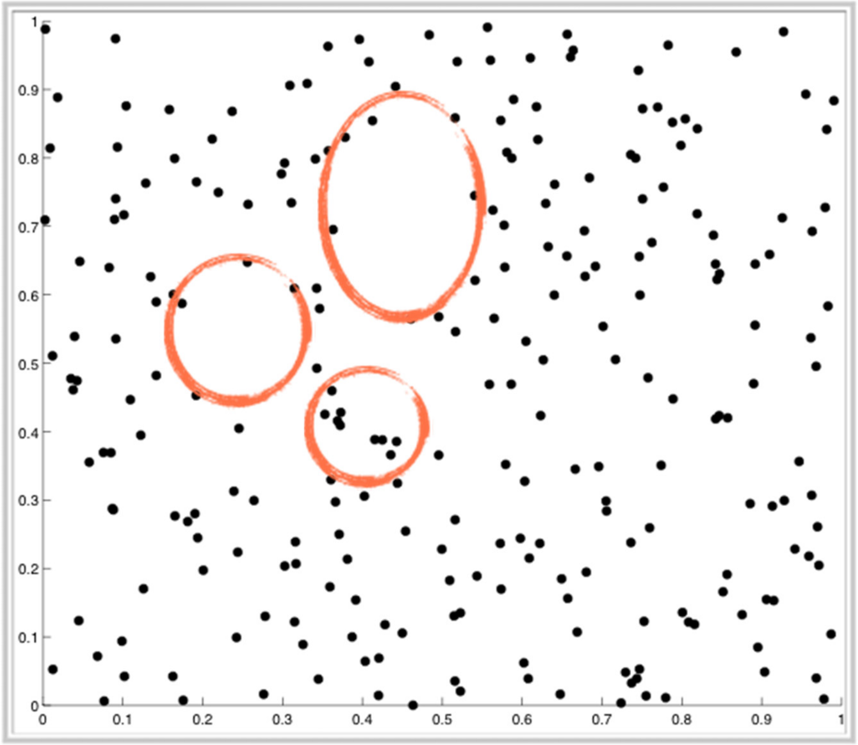
Approximation error can be measured by $V(f)d_\infty^*(x^1, \dots, x^n)$, where

  - $V(f)$ depends only on f
  - $d_\infty^*(x^1, \dots, x^n)$ depends only on $x^1, \dots, x^n$

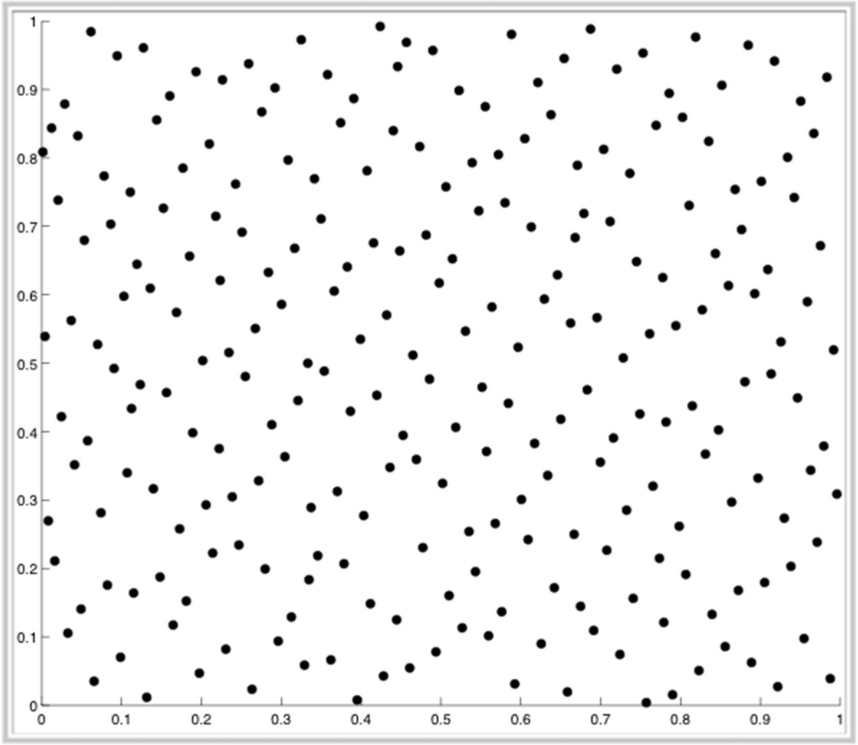
We cannot influence $V(f)$, but we can very well choose $x^1, \dots, x^n$

Low Star Discrepancy Point Sets

- Idea of Quasi-Monte Carlo integration: evaluate f in low discrepancy point sets



(Pseudo) Random



Quasi Random

Low Star Discrepancy Point Sets

- Idea of Quasi-Monte Carlo integration: evaluate f in low discrepancy point sets
- 2 Main Problems:
  - Where to place the points?
    (high-dimensional problem!)
  - Computation of star discrepancies is provably hard
    (NP-hard and W[1]-hard in the dimension, cf. [GSW09,GKWW12])

Our algorithm(s) are the gold-standard to address both problems

Human-Competitiveness 1/5

Criterion (B): *The results are **equal to or better than a result that was accepted as a new scientific result** at the time when it was published in a peer-reviewed scientific journal*

- Our algorithms clearly outperform all previous works
  - Exponential performance increase for our evaluation algorithm (previous work includes [WF97, Th01a, Th01b, Sh12])
  - Computed point sets are better by 36% on average when compared to results in [Th01a, Th01b, DGW10]

Human-Competitiveness 2/5

Criterion (D): *The results are **publishable in its own right** as new scientific results independent of the fact they were mechanically created*

- We have published our papers in the most prestigious journals of the field: *ACM Transactions on Modeling and Computer Simulation* & *SIAM Journal on Numerical Analysis*
- We have as well presented them in the relevant conferences of the different communities: *GECCO 2009*, *MCQMC 2008*, *MCM 2011*, *UDT2012*, *MCQMC 2012*, *GECCO 2013*, and at various relevant workshops

Human-Competitiveness 3/5 & 4/5

Criterion (E): *The results are equal to or better than the most recent human-created solution to* ***a long-standing problem*** *for which there has been a* ***succession of increasingly better human-created solutions***
Criterion (F): *The results are equal to or better than a result that was considered an* ***achievement in its field*** *at the time it was first discovered*

- There has been a long sequence of previous works on both problems (computing the discrepancy of a given point set and creating low discrepancy point configurations, respectively) [e.g., Nie72, De86, BZ93, DEM96, WF97, Th00, Th01a, Th01b, DGW10, and many more]
- Our algorithm is suited also for computing inverse star discrepancies
(i.e., for given dimension d and constant δ, what is the smallest n such that there exists $x^1, \ldots, x^n$ in $[0,1)^d$ with $d_\infty^*(x^1, \ldots, x^n) \leq \delta$?)

Human-Competitiveness 3/5 & 4/5, cont.

Criterion (E): *The results are equal to or better than the most recent human-created solution to* ***a long-standing problem*** *for which there has been a* ***succession of increasingly better human-created solutions***
Criterion (F): *The results are equal to or better than a result that was considered an* ***achievement in its field*** *at the time it was first discovered*

- Our algorithm is also much faster than previous approaches:

	Our algorithm		Thiémard Th01b	
Instance	Time	Result	Time to get same result	Result at same time
Faure-12-169	25s	0.2718	1s	0.2718
Sobol'-12-128	20s	0.1885	7.6m	0.1463
Sobol-12-256	35s	0.1110	1.6d	0.0873
Faure-20-1500	4.7m	0.0740	>4d	None
GLP-20-1619	5.2m	0.0844	>5d	None
Sobol-50-4000	42m	0.0665	9h	None
GLP-50-4000	42m	0.1201	>5d	None

Human-Competitiveness 5/5

Criterion (G): The result solves a problem of **indisputable difficulty** in its field

- The addressed problems are provably (!) difficult and subject to the curse of dimensionality
- Great interest by scientific and industrial researchers and engineers: we have started several new projects that build on our algorithms
- We could solve some open problems posed in the literature (e.g., open problem 42 in [NW10])

Achievements

✓ New genetic algorithms for

- computing low discrepancy point sets
- evaluating star discrepancy values
- computing inverse star discrepancies

✓ Our results clearly outperform previous results by a large margin, both in terms of quality and speed

✓ All computed point sets are available online:
http://qrand.gel.ulaval.ca/
(idea: maintain a database with low star discrepancy point sets)

✓ Great interest from different communities:
several new projects with further applications have been launched (both with mathematicians and engineers)

Full References of Our Papers

- Carola Doerr, François-Michel De Rainville
  Constructing low star discrepancy point sets with genetic algorithms
  In: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO ’13
  Association for Computing Machinery (ACM), 2013. To appear

- François-Michel De Rainville, Christian Gagné, Olivier Teytaud, Denis Laurendeau
  Evolutionary optimization of low-discrepancy sequences
  In: *ACM Transactions on Modeling and Computer Simulation*, 22(2):9:1--9:25
  Association for Computing Machinery (ACM), 2012

- Michael Gnewuch, Magnus Wahlström, Carola Winzen
  A new randomized algorithm to approximate the star discrepancy based on threshold accepting
  In: *SIAM Journal on Numerical Analysis*, 50:781--807
  Society for Industrial and Applied Mathematics (SIAM), 2012

Other References 1/4

- P. Bundschuh, Y. Zhu
  A method for exact calculation of the discrepancy of low-dimensional point sets (I)
  In: Abhandlungen aus dem Mathematischen Seminar der Univ. Hamburg, 63(1):115--133
  Springer, 1993

- L. De Clerck
  A method for exact calculation of the star-discrepancy of plane sets applied to the sequence of Hammersley
  In: Monatshefte für Mathematik, 101(4):261--278
  Springer, 1986

- David P. Dobkin, David Eppstein, Don P. Mitchell
  Computing the discrepancy with applications to supersampling patterns
  In: ACM Transactions on Graphics, 15(4):354--376
  Association for Computing Machinery (ACM), 1996

- Benjamin Doerr, Michael Gnewuch, Magnus Wahlström
  Algorithmic construction of low discrepancy point sets via dependent randomized rounding
  In: Journal of Complexity, 26:490--507
  Elsevier, 2010

Other References 2/4

- Panos Giannopoulos, Christian Knauer, Magnus Wahlström, Daniel Werner
  Hardness of discrepancy computation and epsilon-net verification in high dimension
  In: Journal of Complexity, 28(2):162--176
  Elsevier, 2012

- Michael Gnewuch, Anand Srivastav, Carola Winzen
  Finding optimal volume subintervals with k points and calculating the star discrepancy are NP-hard problems
  In: Journal of Complexity, 25(2):115--127
  Elsevier, 2009

- Harald Niederreiter
  Methods for estimating discrepancy
  In: Applications of Number Theory to Numerical Analysis, 203--236
  Academic Press, 1972

- Erich Novak and Henryk Wozniakowski
  Standard Information for Functionals
  In: Tractability of Multivariate Problems, vol. 2
  EMS Tracts in Mathematics, European Mathematical Society, 2010

Other References 3/4

- Manan Shah
  A genetic algorithm approach to estimate lower bounds of the star discrepancy
  In: Monte Carlo Methods and Applications, 16(3-4):379--398
  De Gruyter, 2010

- Eric Thiémard
  Computing bounds for the star discrepancy
  In: Computing 65(2):169--186
  Springer, 2000

- Eric Thiémard
  An algorithm to compute bounds for the star discrepancy
  In: Journal of Complexity, 17(4):850--880
  Elsevier, 2001

- Eric Thiémard
  Optimal volume subintervals with k points and star discrepancy via integer programming
  In: Mathematical Methods of Operations Research, 54(1):21--45
  Springer, 2001

Other References 4/4

- Peter Winker and Kai-Tai Fang
  Applications of threshold-accepting to the evaluation of the discrepancy of a set of points
  In: *SIAM Journal on Numerical Analysis*, 34:2028--2042
  Society for Industrial and Applied Mathematics (SIAM), 1997