Halftoning and quasi-Monte Carlo

Ken Hanson

CCS-2, Continuum Dynamics
Los Alamos National Laboratory

Los Alamos NATIONAL LABORATORY

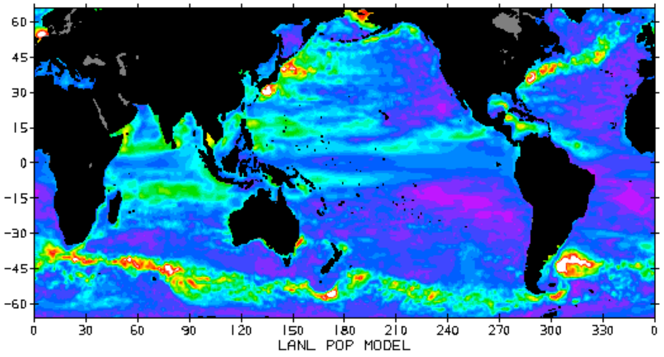
This presentation available at
http://www.lanl.gov/home/kmh/

Overview

- Digital halftoning – purpose and constraints
  - direct binary search (DBS) algorithm for halftoning
  - minimize cost function based on human visual system
- Quasi-Monte Carlo (QMC) – purpose, examples
- Minimum Visual Discrepancy (MVD) algorithm for points, analogous to DBS
  - examples; integration tests
- Voronoi diagrams – calculation via Monte Carlo
  - Voronoi weighted integration – lowers rms error in MC integr.
- Extensions
  - interacting particle model – good for higher dimensions

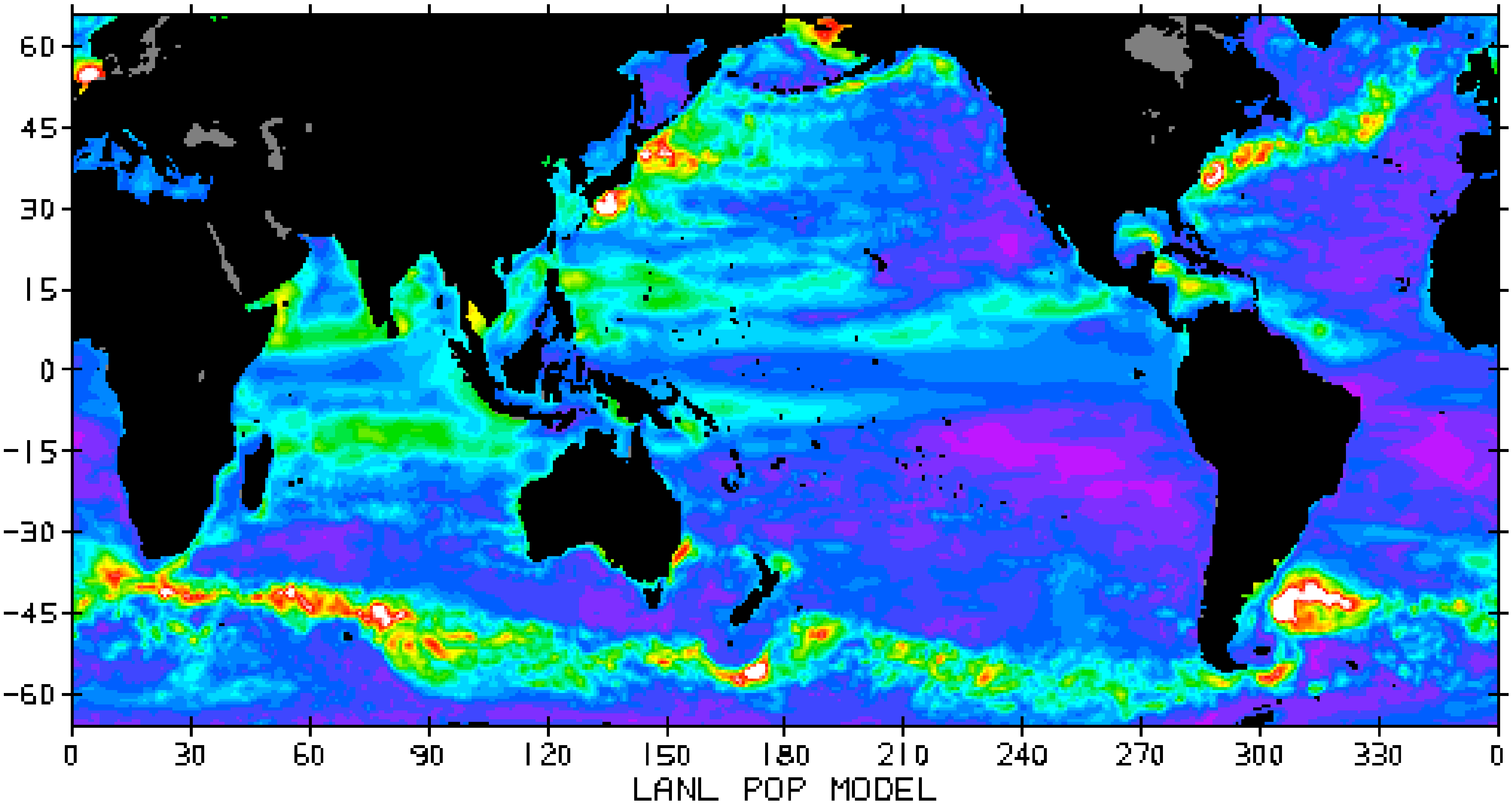
Validation of physics simulation codes

- Computer simulation codes
  - many input parameters, many output variables
  - very expensive to run; up to weeks on super computers
- It is important to validate codes - therefore need
  - to compare codes to experimental data; make inferences
  - advanced methods to estimate sensitivity of simulation outputs on inputs
    - Latin square (hypercube), stratified sampling, quasi-Monte Carlo
- Examples of complex simulations
  - ocean and atmosphere modeling
  - aircraft design, etc.
  - casting of metals



Example of ocean model simulation

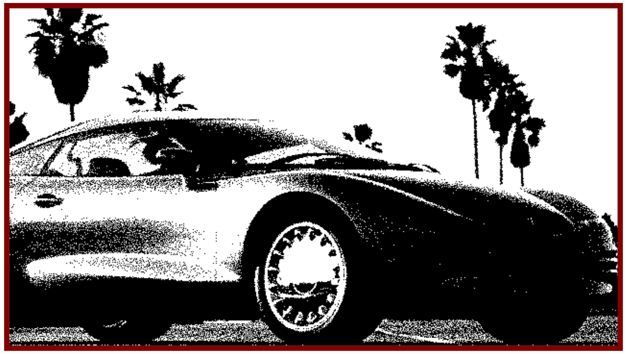
1/6 degree resolution – rms dev. in ocean height



calculation time ≈ one month!

Digital halftoning techniques

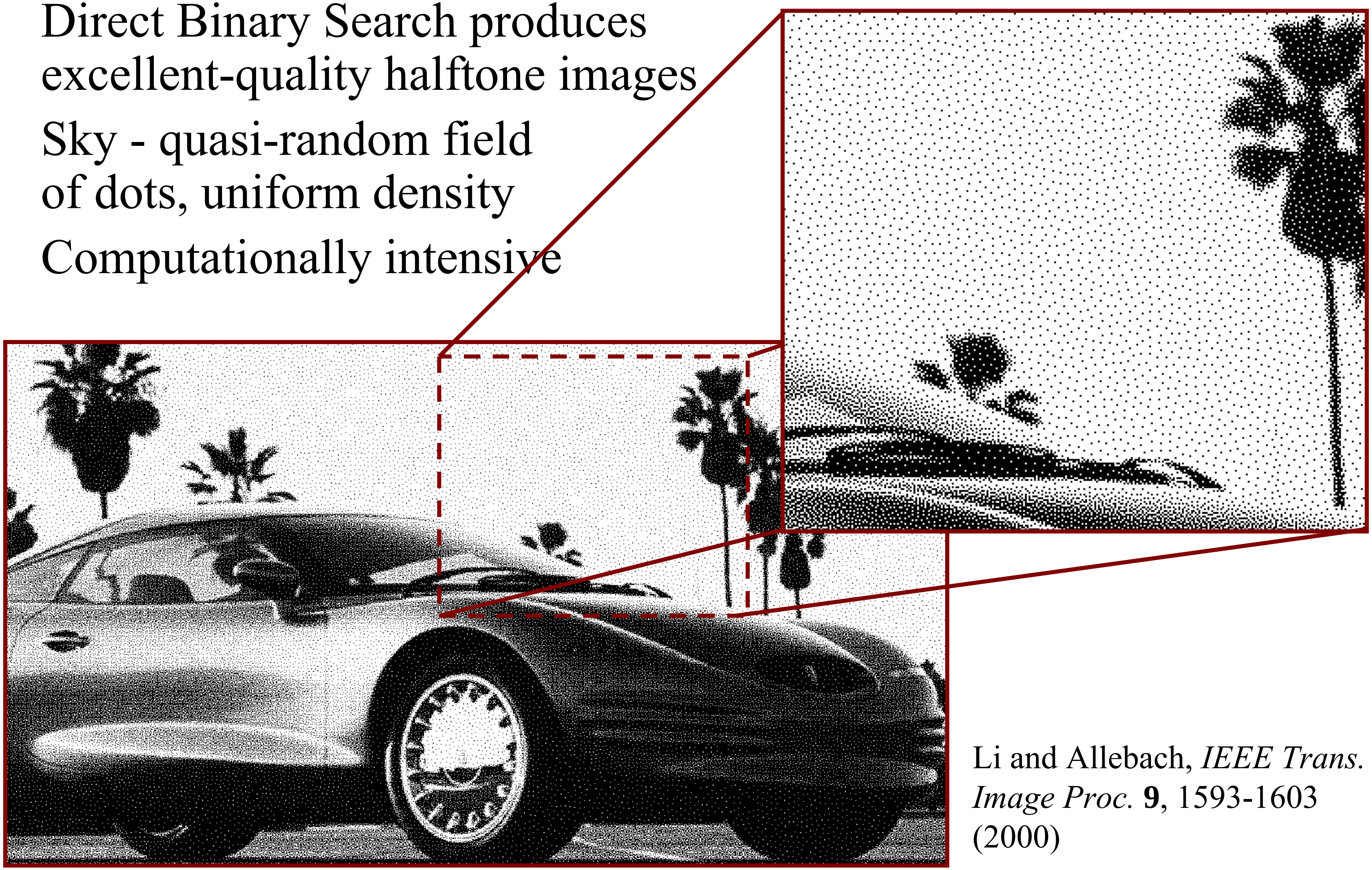
- Purpose
  - render a gray-scale image by placing black dots on white background
  - make halftone rendering **look like** original gray-scale image
- Constraints
  - resolution – size and spacing of dots, number of dots
  - speed of rendering



- Various algorithmic approaches
  - error diffusion, look-up tables, blue-noise, …
  - focus here on Direct Binary Search

DBS example

- Direct Binary Search produces excellent-quality halftone images
- Sky - quasi-random field of dots, uniform density
- Computationally intensive



Li and Allebach, *IEEE Trans. Image Proc.* **9**, 1593-1603 (2000)

Direct Binary Search (DBS) algorithm

- Digital halftone image is composed of black or white pixels
- Cost function is based on perception of two images

$$\varphi = \left| \mathbf{h} * (\mathbf{d} - \mathbf{g}) \right|^2$$

  - where **d** is the dot image, **g** is the gray-scale image to be rendered, * represents convolution, and **h** is the image of the blur function of the human eye, for example, $h(r) = (w^2 + r^2)^{-3/2}$
- To minimize φ
  - start with a collection of dots with average local density ~ **g**
  - iterate sequentially through all image pixels
  - for each pixel, swap value with neighborhood pixels, or toggle its value to reduce φ

Monte Carlo integration techniques

- Purpose
  - estimate integral of a function over a specified region R in m dimensions, based on evaluations at n sample points

$$\int_R f(\boldsymbol{x})\, d\boldsymbol{x} = \frac{V_R}{n} \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} f(\boldsymbol{x_i})$$

- Constraints
  - integrand not available in analytic form, but calculable
  - function evaluations may be expensive, so minimize them
- Algorithmic approaches
  - focus on accuracy in terms of # of function evaluations n
  - quadrature (Simpson) – good for few dimensions; rms err ~ n^{-1}
  - Monte Carlo – useful for many dimensions; rms err ~ $n^{-1/2}$
  - quasi-Monte Carlo – reduce # of evaluations; rms err ~ n^{-1}

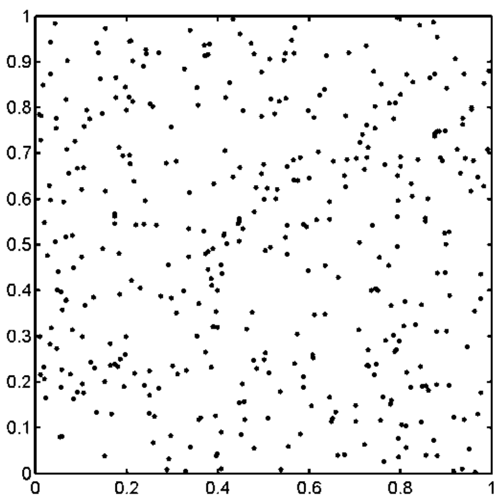
Quasi-Monte Carlo

- Purpose
  - estimate integral of a function over a specified domain in *m* dimensions
  - obtain better rate of convergence of integral estimation than seen in classic Monte Carlo
- Constraints
  - integrand function not available analytically, but calculable
  - function known (or assumed) to be well behaved
- Standard QMC approaches use low-discrepancy sequences in product space (Halton, Sobel, Faure,...)
- **Purpose here is to propose a new way of generating sets of sample points**

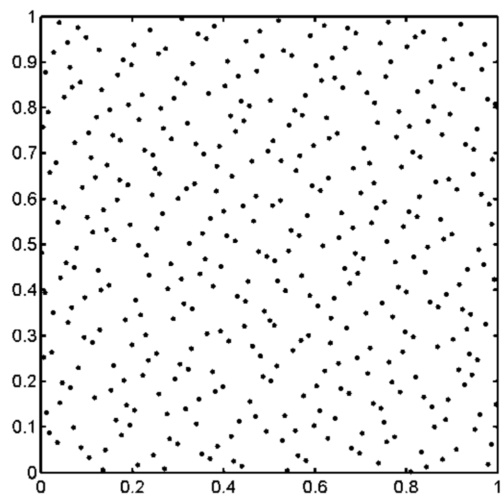
Point set examples

- Examples of different kinds of point sets
  - 400 points in each
- If quasi-MC sequences have better integration properties than random, is halftone pattern even better?

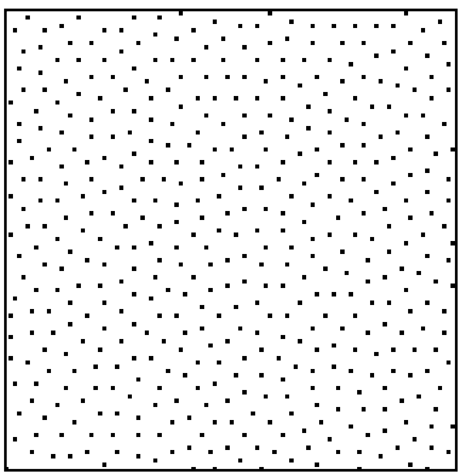
Random (independent)



Quasi-Random (Halton sequence)



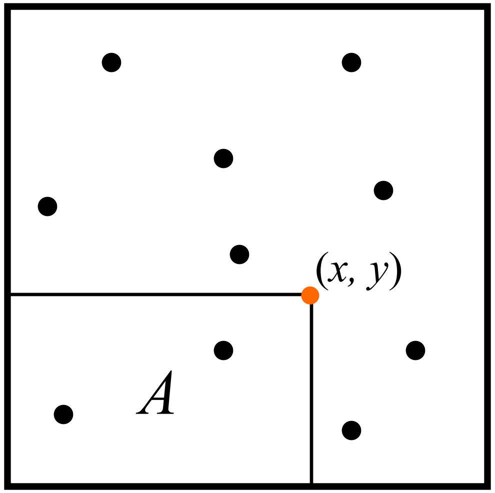
Halftone (DBS sky)



Discrepancy

- Much of QMC work is based on the discrepancy, defined for samples covering the unit square in 2D as

$$D_2 = \int_U \left[n(x,y) - A(x,y) \right]^2 \; dxdy$$



  - where integration is over unit square,
  - $n(x, y)$ is the number of points in the rectangle with opposing corners (0, 0) to (x, y), and
  - $A(x, y)$ is the area of the rectangle
- Can be related to upper bounds on integration error for some classes of functions
- Clearly a measure of uniformity of dot distribution; however, only for particular structure function

Minimum Visual Discrepancy (MVD) algorithm

Inspired by Direct Binary Search halftoning algorithm

- Start with an initial set of points
- Goal is to create uniformly distributed set of points
- Cost function is variance in blurred point image

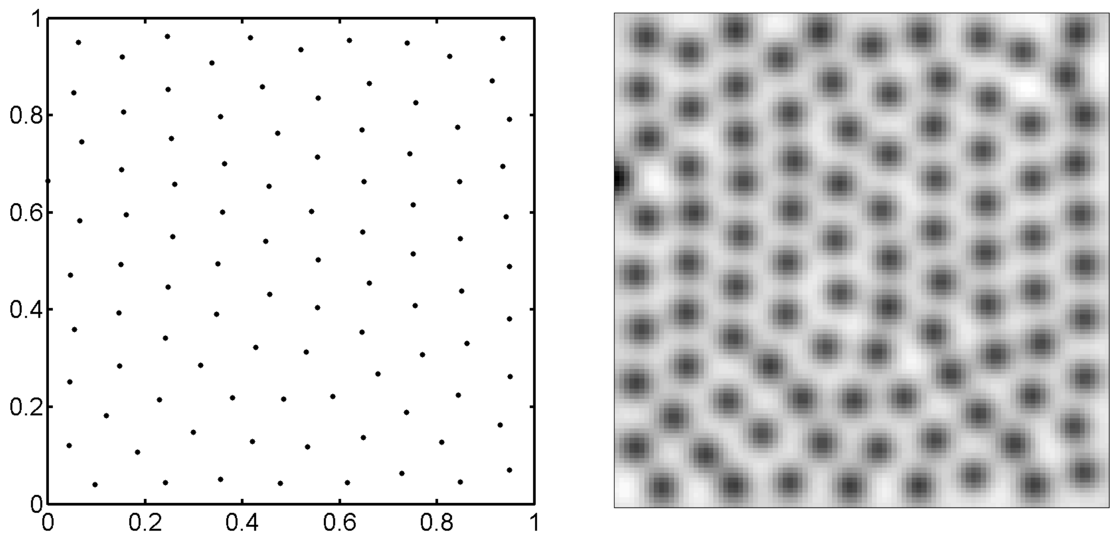
$$\psi = \mathrm{var}(\mathbf{h} * \mathbf{d})$$

  - where **d** is the point (dot) image, **h** is the blur function of the human eye, and * represents convolution
- To minimize ψ
  - start with some point set (random, stratified, Halton,…)
  - iterate through points in random order;
  - move each point in 8 directions, and accept move that has lowest ψ

Minimum Visual Discrepancy (MVD) algorithm

- MVD result; initialized with 100 points from Halton seq.
- MVD algorithm minimizes variance in blurred image
  - effect is to force points to be as far apart from each other as possible, constrained to unit square; thus, evenly distributed
  - expect global minimizer is a regular pattern; hexagonal in 2D

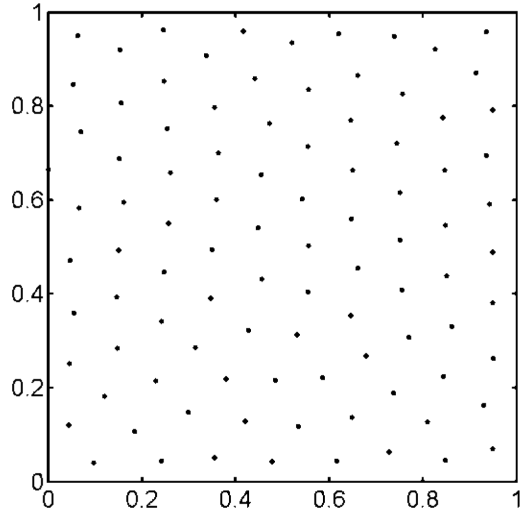
100, MVD   Blurred image



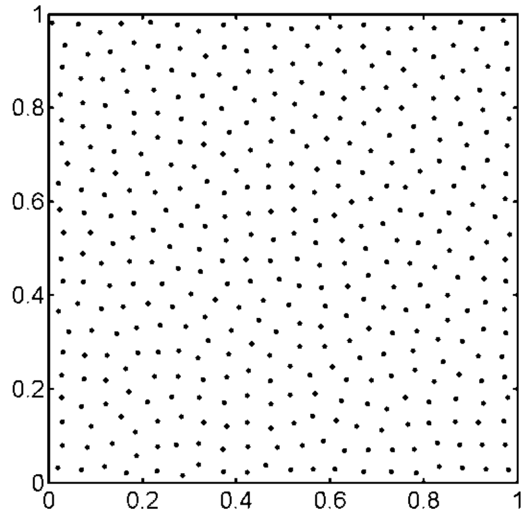
MVD point sets

- In each optimization, final pattern depends on initial point set
  - algorithm seeks local minimum, not global (similar to DBS)
- Patterns somewhat resemble regular hexagonal array
  - similar to lattice structure in crystals or glass
  - however, they lack long-range (coarse scale) order
  - best to start with point set with good long-range uniformity

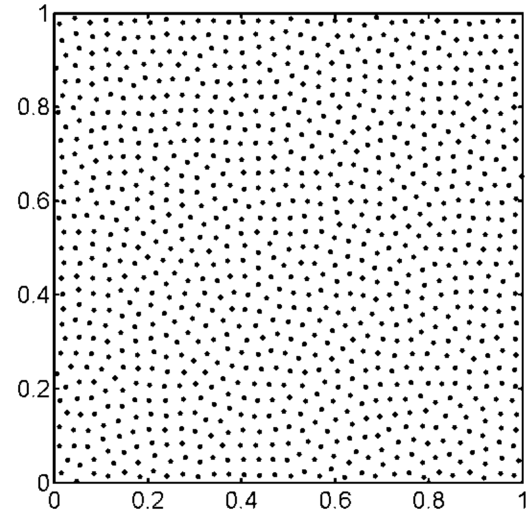
100, MVD



400, MVD



1000, MVD



Analogy to interacting particles

- Consider points as set of interacting particles
- Cost function is the total potential

$$\psi = \sum_{\boldsymbol{i},\boldsymbol{j} \geq \boldsymbol{i}+1} V(\boldsymbol{x}_{\boldsymbol{i}}, \boldsymbol{x}_{\boldsymbol{j}}) + \sum_{\boldsymbol{i}} U(\boldsymbol{x}_{\boldsymbol{i}})$$

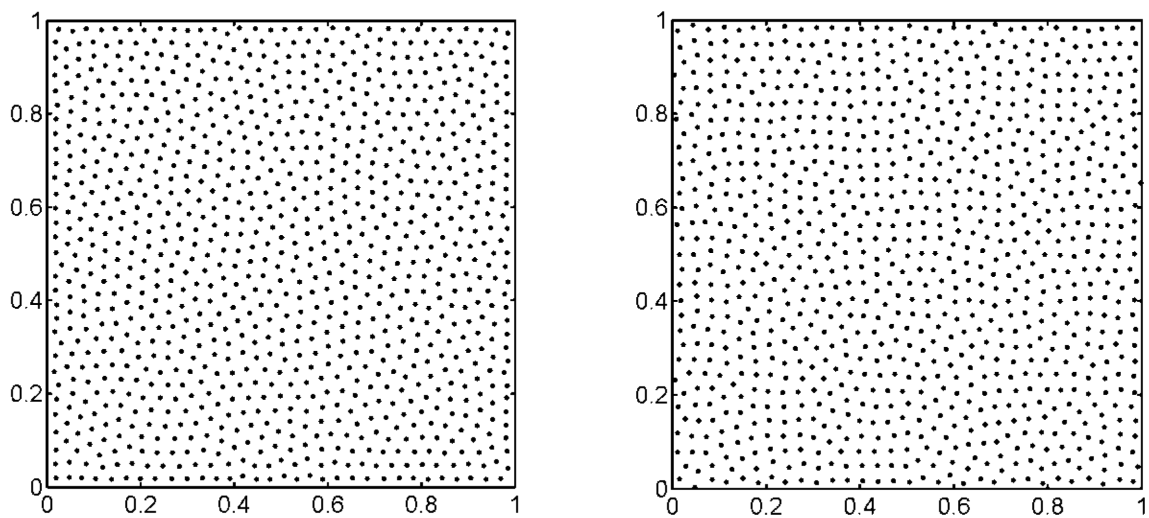
  - where $\boldsymbol{x}_i$ is location of *i*th particle
    V is particle-particle interaction potential
    and *U* is particle-boundary potential
  - particles are repelled by each other and boundary
- Minimize ψ by moving particles around
- This model is formally equivalent to Minimum Visual Discrepancy (*V* and *U* are directly related to blur fnc. **h**)
- Suitable for generating point sets in high dimensions

Interacting particle approach

- Example of interacting-particle calculation
  - resulting point pattern is visually indistinguishable from MVD pattern

Interacting Particle　　　　MVD

1000 points

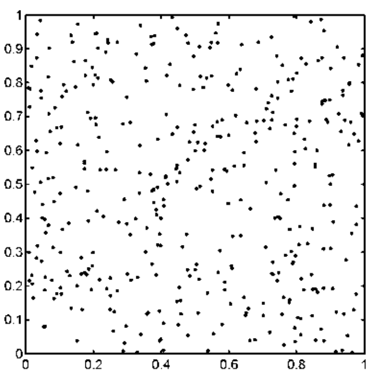


Comparison of various point sets

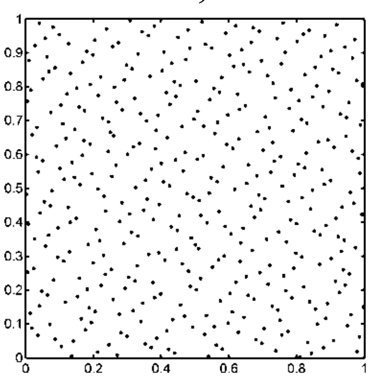
- Various kinds of point sets (400 points)
- Varying degrees of randomness and uniformity
- As the points become more uniformly distributed, the more accurate are the estimated integrals values

RMS relative accuracies of integral of $\text{func2} = \prod_i \exp\left(-2\left|\boldsymbol{x}_i - \boldsymbol{x}_i^0\right|\right); \quad 0 < \boldsymbol{x}_i^0 < 1$

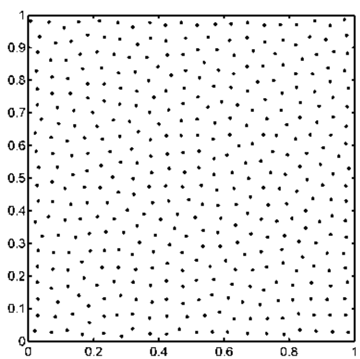
Random, 2.5%



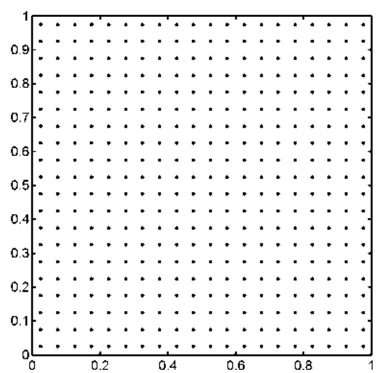
Halton, 0.5%



MVD, 0.14%

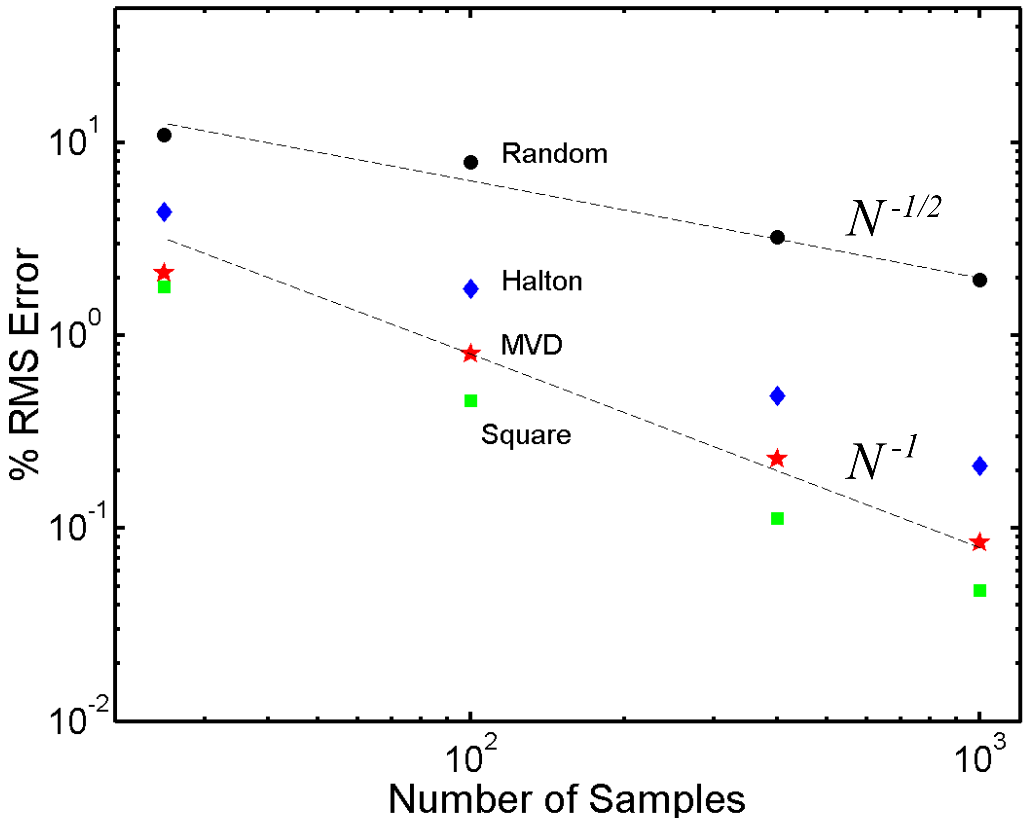


Grid, 0.09%



More Uniform, Higher Accuracy

Integration test results



- RMS error for integral of $\text{func2} = \prod_i \exp\left(-2\left|\boldsymbol{x}_i - \boldsymbol{x}_i^0\right|\right); \quad 0 < \boldsymbol{x}_i^0 < 1$
  - from worst to best: random, Halton, MVD, square grid
  - lines show $N^{-1/2}$ (expected for MC) and N^{-1} (expected for QMC)

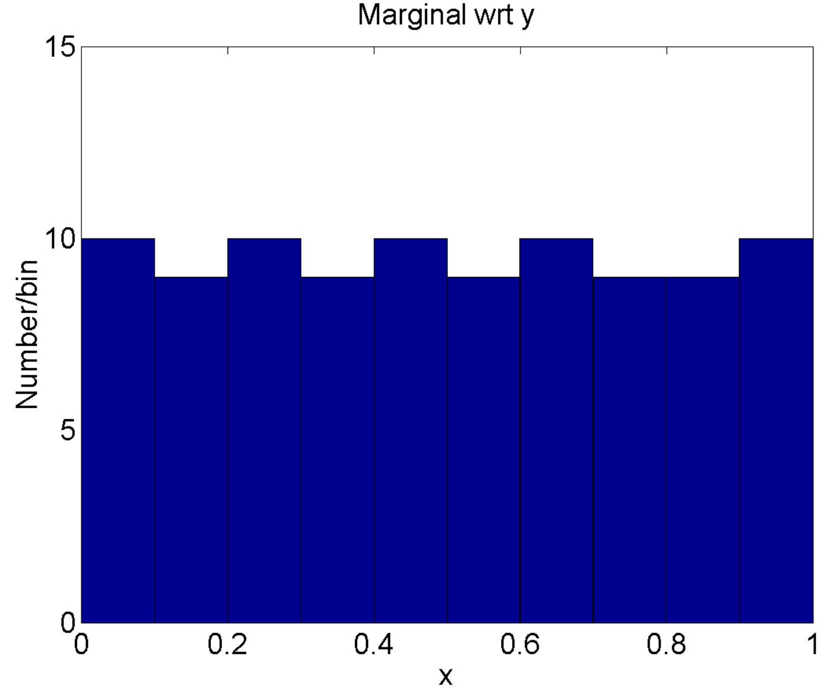
Regular versus random sampling

- If sampling on square grid gives lowest integration errors, why use random samples at all?
- Arguments for/against regular sampling:
  - pro - easy to do and good integr. accuracy (in low D)
  - con – only specific number of samples can be had (n^d), and difficult to add extra points;
    - many points required in high D
- Arguments for/against random sampling
  - pro – easy to add more points;
    - high D no problem
    - less likely to be fooled by periodic function;
  - con – lower accuracy and slow ($n^{-1/2}$) convergence
- QMC and MVD try to combine best of both

Marginals for MVD

- Desirable to have marginals of high dimensional point sets to uniformly sample in each parameter
- Latin hypercube sampling designed to achieve this property (for specified number of points)
- Plot shows histogram of 95 MVD samples along x-axis, i.e., marginalized over y direction
- MVD points have relatively uniform marginal distributions

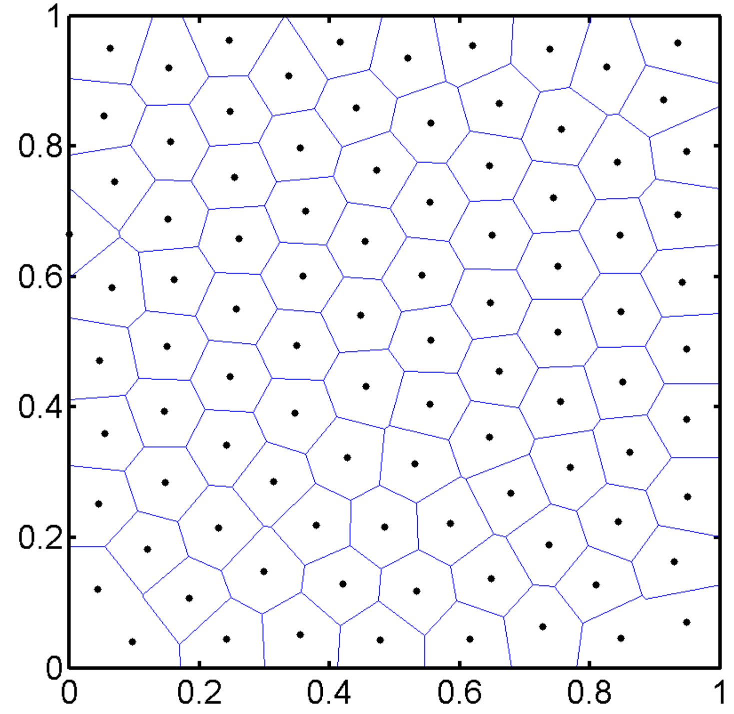
95 MVD points



Voronoi analysis via Monte Carlo

- Voronoi diagram
  - partitions region of interest into polygons
  - points within each polygon are closest to corresponding generating point, Z_i
- MC technique facilitates Voronoi analysis
  - randomly throw large number of points $\{X_i\}$ into region
  - compute distance of each X_i to all generating points $\{Z_i\}$
  - sort according to which Z_i they are closest to
  - can compute area A_i, radial moments,…
- Easily extended to high dimensions

100 MVD points



Voronoi analysis can improve classic MC

- Standard MC formula

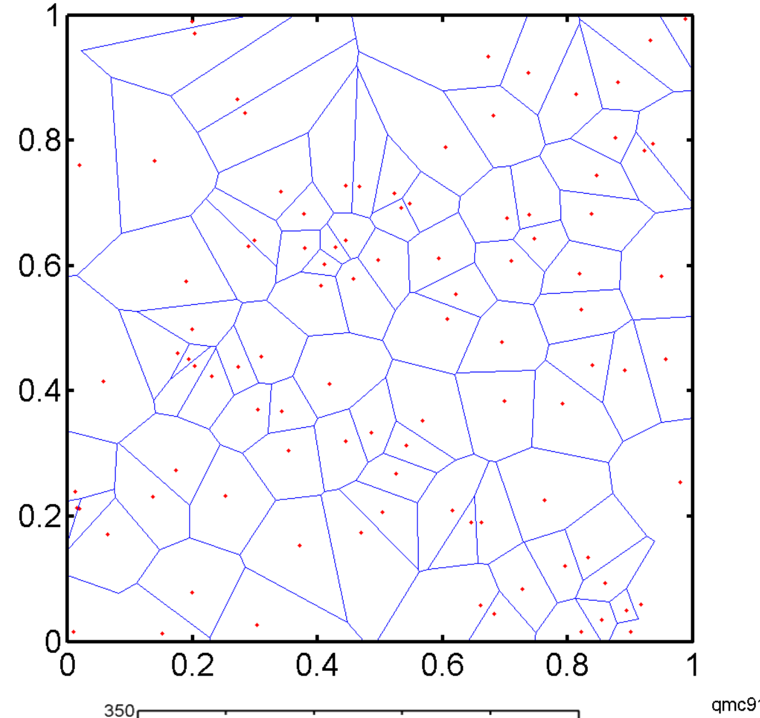
$$\int_R f(\mathbf{x})\,d\mathbf{x} = \frac{V_R}{n}\sum_{i=1}^{n} f(\mathbf{x}_i)$$

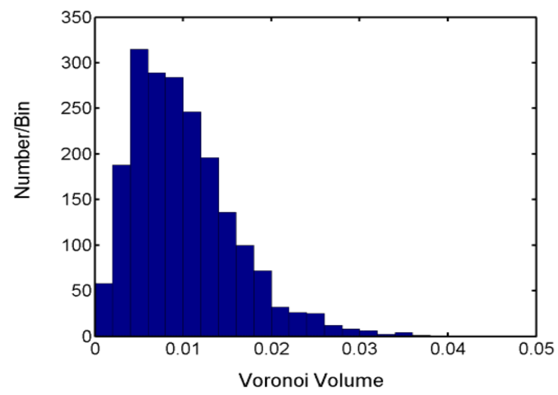
- Instead, use weighted average

$$\int_R f(\mathbf{x})\,d\mathbf{x} = \sum_{i=1}^{n} f(\mathbf{x}_i)V_i$$

  - where V_i is the volume of Voronoi region for ith point; Riemann integr.
- Accuracy of integral estimate dramatically improved in 2D:
  - factor of 6.3 for $N = 100$ (func2)
  - factor of > 20 for $N = 1000$ (func2)
- Suitable for adaptive sampling
- Less useful in high dimensions (?)

100 random samples

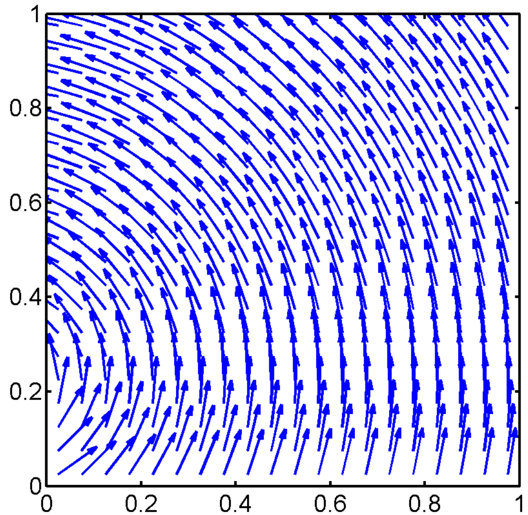




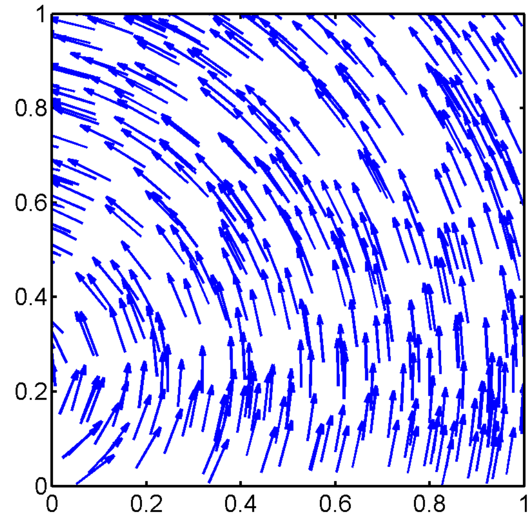
Visualization of fluid flow

- Fluid flow often visualized as field of vectors
- Location of vector bases may be chosen as
  - square grid (typical) - regular pattern produces visual artifacts
  - random points - fewer artifacts, but nonuniform placement
  - quasi-random - fewest artifacts and uniform placement

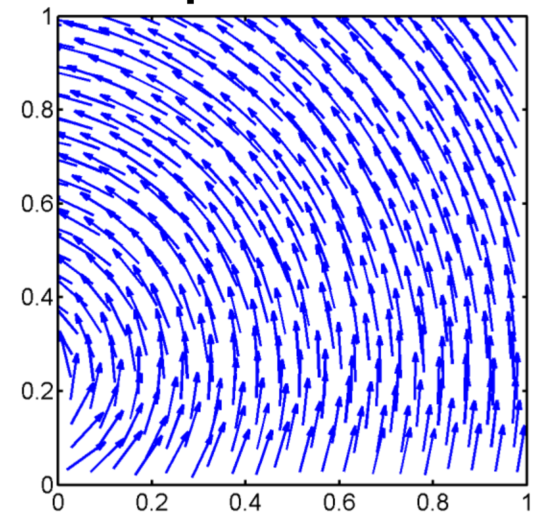
Square grid



Random points



Quasi-random (MVD) point set



Extensions

- Generation of optimal point sets in high dimensions
  - particle interaction model (equivalent to MVD)
- Sequential generation of point sets
  - add one point at a time to previous fixed point set
- Apply to arbitrary domains
- Draw MVD samples from specified pdf
- Use in visualization of flow fields, streamlines
- Adapt these ideas to MCMC for improved efficiency (??)

Conclusions

- Minimum Visual Discrepancy algorithm
  - ▸ produces point sets resembling uniform halftone images
  - ▸ yields better integral estimates than standard QMC sequences
  - ▸ equivalently, can use particle interaction model in high dimen.
- Voronoi analysis – can improve accuracy of classic MC
  - centroidal Voronoi tessellation (Gunzberger)

Bibliography

- K. M. Hanson, "Quasi-Monte Carlo: halftoning in high dimensions?," *Proc. SPIE* **5016**, 161-172 (2003)
- P. Li and J. P. Allebach, "Look-up-table based halftoning algorithm," *IEEE Trans. Image Proc.* **9**, 1593-1603 (2000)
- H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, (SIAM, 1992)
- Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi tesselations: applications and algorithms," *SIAM Review* **41**, 637-676 (1999)

This presentation and paper available at
http://www.lanl.gov/home/kmh/