

# Automatic Algorithms for Multidimensional Integration

Fred J. Hickernell

Department of Applied Mathematics

Center for Interdisciplinary Scientific Computation

Illinois Institute of Technology

[hickernell@iit.edu](mailto:hickernell@iit.edu)   [mypages.iit.edu/~hickernell](http://mypages.iit.edu/~hickernell)

Thanks to the GAIL team, esp. Lan Jiang, Lluís Antoni Jiménez Rugama, and Jagadees Rathinavel

NSF-DMS-1522687 and NSF-DMS-1638521 (SAMSI)

Matrix Workshop on the Frontiers of High Dimensional Computation, June 11, 2018





# When Do We Stop?

Compute an **integral**

$$\mu(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \varrho(\mathbf{x}) d\mathbf{x} \quad \text{Bayesian inference, financial risk, statistical physics, ...}$$

**Desired solution:** An adaptive algorithm,  $\hat{\mu}(\cdot, \cdot)$  of the form

$$\hat{\mu}(f, \varepsilon) = w_{0,n} + \sum_{i=1}^n w_{i,n} f(\mathbf{x}_i), \quad \text{e.g., } w_{0,n} = 0, w_{1,n} = \dots = w_{n,n} = \frac{1}{n}$$

where  $n$ ,  $\{\mathbf{x}_i\}_{i=1}^\infty$ ,  $w_{0,n}$ , and  $\mathbf{w} = (w_{i,n})_{i=1}^n$  are chosen to **guarantee**

$$|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon \text{ with high probability} \quad \forall \varepsilon > 0, \text{ reasonable } f$$



# When Do We Stop?

$$\mu(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \varrho(\mathbf{x}) d\mathbf{x}, \quad \hat{\mu}(f, \varepsilon) = w_{0,n} + \sum_{i=1}^n w_{i,n} f(\mathbf{x}_i),$$

Want  $|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon$  with high probability  $\forall \varepsilon > 0$ , reasonable  $f$

## Possible solutions

IID sampling:  $n = \left( \frac{2.58 \hat{\sigma}}{\varepsilon} \right)^2$

Low discrepancy sampling: choose  $n$  to make  $\text{disc}(\{\mathbf{x}_i\}_{i=1}^n) \text{Var}(f) \leq \varepsilon$

Randomized low discrepancy sampling: look at spread of sample means from random replications of low discrepancy points

## Drawbacks

Does CLT hold yet? How to determine  $\hat{\sigma}$ ?

Don't know  $\text{Var}(f)$ ; computing  $\text{disc}(\{\mathbf{x}_i\}_{i=1}^n)$  may be expensive

How many replications? What measure of spread?



# When Do We Stop?

$$\mu(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \varrho(\mathbf{x}) d\mathbf{x}, \quad \hat{\mu}(f, \varepsilon) = w_{0,n} + \sum_{i=1}^n w_{i,n} f(\mathbf{x}_i),$$

Want  $|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon$  with high probability  $\forall \varepsilon > 0$ , reasonable  $f$

## Possible solutions

IID sampling:  $n = \left( \frac{2.58 \hat{\sigma}}{\varepsilon} \right)^2$

Low discrepancy sampling: choose  $n$  to make  $\text{disc}(\{\mathbf{x}_i\}_{i=1}^n) \text{Var}(f) \leq \varepsilon$

Randomized low discrepancy sampling: look at spread of sample means from random replications of low discrepancy points

## Drawbacks

Does CLT hold yet? How to determine  $\hat{\sigma}$ ?

Don't know  $\text{Var}(f)$ ; computing  $\text{disc}(\{\mathbf{x}_i\}_{i=1}^n)$  may be expensive

How many replications? What measure of spread?

- Identify a data-driven measure of error
- Identify a cone of nice integrands for which that data-driven error can be proven to hold

Is integrand really in the cone?  
Well, sometimes there are necessary conditions that can be checked



# Automatic IID Sampling<sup>1</sup>

Replace the Central Limit Theorem by Berry-Esseen inequalities:

$$\mu := \int_{\mathcal{X}} f(\mathbf{x}) \varrho(\mathbf{x}) d\mathbf{x} =: \mathbb{E}[f(\mathbf{X})], \quad \mathcal{C} = \{f \in L^4(\mathcal{X}, \varrho) : \text{kurt}(f(\mathbf{X})) \leq \kappa_{\text{up}}\}$$

For  $n_\sigma$  determined by  $\kappa_{\text{up}}$ , and  $x_1, x_2, \dots$  IID let

$$\sigma_{\text{up}}^2 = \frac{(1.2)^2}{n_\sigma - 1} \sum_{i=1}^{n_\sigma} f(x_i), \quad \hat{\mu}(f, \varepsilon) = \frac{1}{n_\mu} \sum_{i=n_\sigma+1}^{n_\sigma+n_\mu} f(x_i), \quad \text{where}$$

$$n_\mu = \min \left\{ n \in \mathbb{N} : \Phi(-\sqrt{n}\varepsilon/\sigma_{\text{up}}) + \frac{1}{\sqrt{n}} \min \left( 0.3328(\kappa_{\text{up}}^{3/4} + 0.429), \frac{18.1139\kappa_{\text{up}}^{3/4}}{1 + |\sqrt{n}\delta/\sigma_{\text{up}}|^3} \right) \leq 0.25\% \right\}$$

Then  $\mathbb{P}\{|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon\} \geq 99\%$ .

---

<sup>1</sup>H., F. J. et al. Guaranteed Conservative Fixed Width Confidence Intervals Via Monte Carlo Sampling. in *Monte Carlo and Quasi-Monte Carlo Methods 2012* (eds Dick, J. et al.) 65 (Springer-Verlag, Berlin, 2013), 105–128.



# Automatic Quasi-Monte Carlo Sampling<sup>2</sup>

$$\mu = \int_{[0,1]^d} f(\mathbf{x}) d\mathbf{x} = \hat{f}(0), \quad f(\mathbf{x}) = \sum_{\mathbf{k}} \hat{f}(\mathbf{k}) \phi_{\mathbf{k}}(\mathbf{x}), \quad \hat{f}(\mathbf{k}) = \int_{[0,1]^d} f(\mathbf{x}) \bar{\phi}_{\mathbf{k}}(\mathbf{x}) d\mathbf{x}$$

$\phi_{\mathbf{k}}$  are complex exponentials (for lattices) or Walsh functions (for nets)  
 $\mathfrak{C}$  = integrands whose true Fourier coefficients do not decay erratically

The error bound is derived in terms of

discrete transform,  $\tilde{f}_n(\mathbf{k}) = \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i) \bar{\phi}_{\mathbf{k}}(\mathbf{x}_i)$   $\forall \mathbf{k}$ , can be computed in  $\mathcal{O}(n \log(n))$  operations

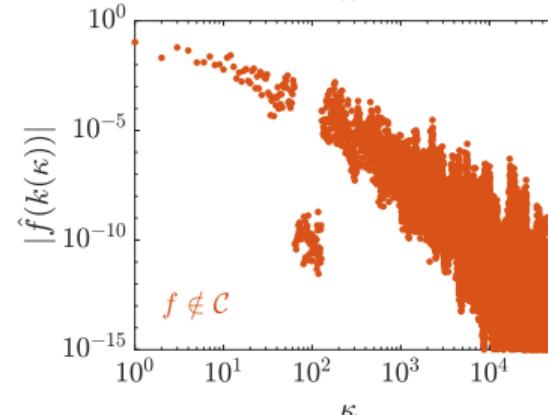
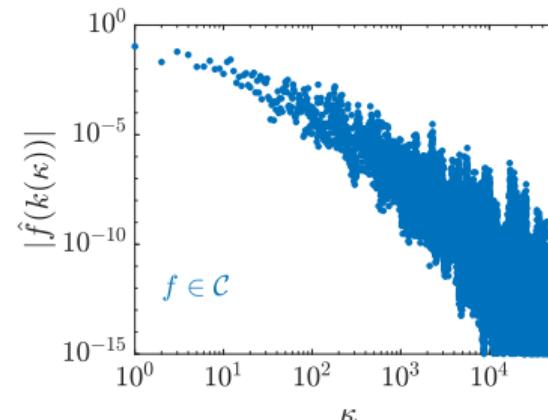
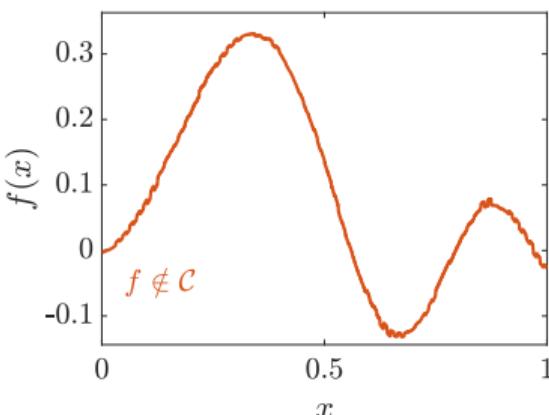
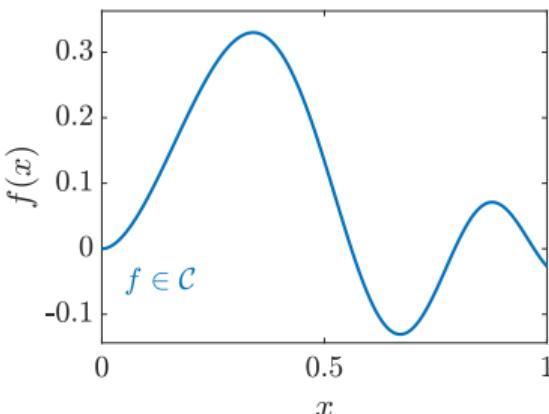
$$\mathfrak{C}(n) \sum_{\text{moderate } \mathbf{k}} |\tilde{f}_n(\mathbf{k})| \leq \varepsilon \implies |\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon$$

---

<sup>2</sup>H., F. J. & LI. A. Jiménez Rugama. *Reliable Adaptive Cubature Using Digital Sequences*. in *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014* (eds Cools, R. & Nuyens, D.) **163**. arXiv:1410.8615 [math.NA] (Springer-Verlag, Berlin, 2016), 367–383, LI. A. Jiménez Rugama & H., F. J. *Adaptive Multidimensional Integration Based on Rank-1 Lattices*. in *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014* (eds Cools, R. & Nuyens, D.) **163**. arXiv:1411.1966 (Springer-Verlag, Berlin, 2016), 407–422, H., F. J. et al. *Adaptive Quasi-Monte Carlo Methods for Cubature*. in *Contemporary Computational Mathematics — a celebration of the 80th birthday of Ian Sloan* (eds Dick, J. et al.) (Springer-Verlag, 2018), 597–619. doi:10.1007/978-3-319-72456-0.



# Integrands Inside and Outside the Cone





# Automatic QMC Sampling Assuming a Random $f$

Assume  $f \sim \mathcal{GP}(m, s^2 C_\theta)$ , a sample from a Gaussian process. Defining

$$c = \int_{[0,1]^d} \int_{[0,1]^d} C_\theta(t, x) dt dx, \quad \mathbf{c} = \left( \int_{[0,1]^d} C_\theta(t, x_i) dt \right)_{i=1}^n, \quad \mathbf{C} = (C_\theta(x_i, x_j))_{i,j=1}^n$$

and choosing the **weights** as

$$w_0 = m[1 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{1}], \quad \mathbf{w} = \mathbf{C}^{-1} \mathbf{c}, \quad \hat{\mu}(f, \varepsilon) = w_0 + \mathbf{w}^T \mathbf{f}, \quad \mathbf{f} = (f(x_i))_{i=1}^n.$$

yields an unbiased approximation:

$$\mu(f) - \hat{\mu}(f, \varepsilon) \mid \mathbf{f} = \mathbf{y} \sim \mathcal{N}\left(0, s^2(c - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c})\right)$$

Choosing  $n$  large enough to make

$$2.58s\sqrt{c - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c}} \leq \varepsilon,$$

assures us that

$$\mathbb{P}_f [|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon] \geq 99\%.$$

**Issues requiring attention:** choosing parameters of the covariance kernel, expensive matrix operations



# To Make This Approach Practical<sup>3</sup>

- Use MLE to estimate parameters
- Choose covariance kernels that match the low discrepancy design

$$c = \int_{[0,1]^d} \int_{[0,1]^d} C_\theta(t, x) dt dx = 1, \quad \mathbf{c} = \left( \int_{[0,1]^d} C_\theta(t, x_i) dt \right)_{i=1}^n = \mathbf{1}$$

$$\mathbf{C} = \left( C_\theta(x_i, x_j) \right)_{i,j=1}^n = \frac{1}{n} \mathbf{V} \Lambda \mathbf{V}^H, \quad \Lambda = \text{diag}(\lambda), \quad \mathbf{V}_1 = \mathbf{v}_1 = \mathbf{1}$$

$$\mathbf{V}^T \mathbf{z} \text{ is } \mathcal{O}(n \log n), \quad \lambda = \mathbf{V}^T \mathbf{C}_1, \quad \mathbf{C}^{-1} \mathbf{1} = \frac{\mathbf{1}}{\lambda_1}$$

Dealing with the fast transformed data,  $\hat{\mathbf{y}} = \mathbf{V}^T \mathbf{y}$ , where  $\mathbf{y} = (f(x_i))_{i=1}^n$ , it follows that

$$\hat{\mu}(f, \varepsilon) = \frac{\hat{y}_1}{n} = \frac{1}{n} \sum_{i=1}^n f(x_i), \quad \mathbb{P}_f [|\mu(f) - \hat{\mu}(f, \varepsilon)| \leq \varepsilon] \geq 99\%,$$

provided

$$2.58 \sqrt{\left(1 - \frac{n}{\lambda_1}\right) \frac{1}{n^2} \sum_{i=2}^n \frac{|\hat{y}_i|^2}{\lambda_i}} \leq \varepsilon$$

---

<sup>3</sup>Jagadeeswaran, R. & H., F. J. Automatic Bayesian Cubature. in preparation. 2018+.



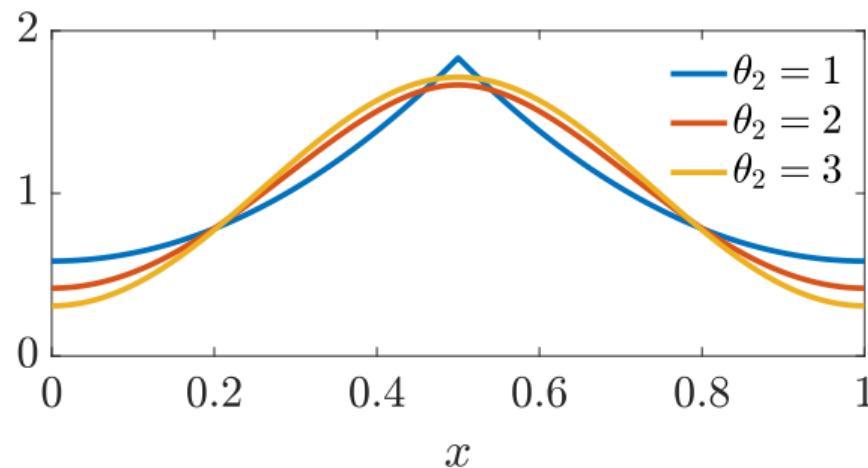
# Form of Matching Covariance Kernels

Typically the domain of  $f$  is  $[0, 1)^d$ , and

$$C(\mathbf{t}, \mathbf{x}) = \begin{cases} \tilde{C}(\mathbf{x} - \mathbf{t} \bmod 1) & \text{integration lattices} \\ \tilde{C}(\mathbf{x} \oplus \mathbf{t}) & \text{Sobol' sequences, } \oplus \text{ means digitwise addition modulo 2} \end{cases}$$

E.g., for integration lattices

$$C(\mathbf{x}, \mathbf{t}) = \prod_{k=1}^d [1 - \theta_1(-1)^{\theta_2} B_{2\theta_2}(x_k - t_k \bmod 1)], \quad \theta_1 > 0, \theta_2 \in \mathbb{N}$$

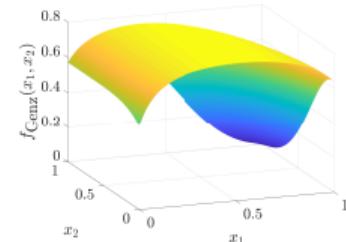




# Gaussian Probability

$$\mu = \int_{[\mathbf{a}, \mathbf{b}]} \frac{\exp\left(-\frac{1}{2}\mathbf{t}^T \Sigma^{-1} \mathbf{t}\right)}{\sqrt{(2\pi)^d \det(\Sigma)}} d\mathbf{t} = \int_{[0,1]^{d-1}} f(\mathbf{x}) d\mathbf{x}$$

For some typical choice of  $\mathbf{a}, \mathbf{b}, \Sigma, d = 3; \mu \approx 0.6763$



Rel. Error Tolerance	Median		Worst 10% $n$	Worst 10% Time (s)	
	Method	Error		Accuracy	
1E-2	IID	5E-4	100%	8.1E4	0.020
1E-2	Sh. Latt.	1E-5	100%	1.0E3	0.004
1E-2	Scr. Sobol'	4E-5	100%	1.0E3	0.005
1E-2	Bayes. Latt.	1E-5	100%	1.0E3	0.002

These algorithms are implemented in GAIL<sup>4</sup>

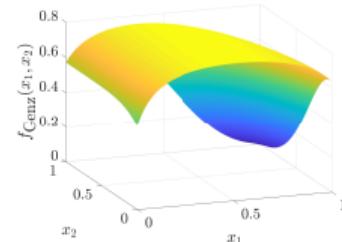
<sup>4</sup>Choi, S.-C. T. et al. GAIL: Guaranteed Automatic Integration Library (Versions 1.0–2.2). MATLAB software. 2013–2017. [http://gailgithub.github.io/GAIL\\_Dev/](http://gailgithub.github.io/GAIL_Dev/).



# Gaussian Probability

$$\mu = \int_{[\mathbf{a}, \mathbf{b}]} \frac{\exp\left(-\frac{1}{2}\mathbf{t}^T \Sigma^{-1} \mathbf{t}\right)}{\sqrt{(2\pi)^d \det(\Sigma)}} d\mathbf{t} = \int_{[0,1]^{d-1}} f(\mathbf{x}) d\mathbf{x}$$

For some typical choice of  $\mathbf{a}, \mathbf{b}, \Sigma, d = 3; \mu \approx 0.6763$



Rel. Error Tolerance	Method	Median		n	Worst 10% Time (s)
		Error	Accuracy		
1E-3	IID	9E-5	100%	2.0E6	0.400
1E-3	Sh. Latt.	8E-6	100%	2.0E3	0.007
1E-3	Scr. Sobol'	2E-5	100%	2.0E3	0.006
1E-3	Bayes. Latt.	1E-5	100%	1.0E3	0.002

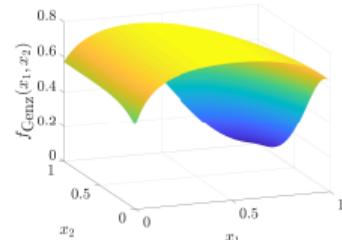
These algorithms are implemented in GAIL<sup>4</sup>

<sup>4</sup>Choi, S.-C. T. et al. GAIL: Guaranteed Automatic Integration Library (Versions 1.0–2.2). MATLAB software. 2013–2017. [http://gailgithub.github.io/GAIL\\_Dev/](http://gailgithub.github.io/GAIL_Dev/).



# Gaussian Probability

$$\mu = \int_{[\mathbf{a}, \mathbf{b}]} \frac{\exp(-\frac{1}{2}\mathbf{t}^T \Sigma^{-1} \mathbf{t})}{\sqrt{(2\pi)^d \det(\Sigma)}} d\mathbf{t} = \int_{[0,1]^{d-1}} f(\mathbf{x}) d\mathbf{x}$$



For some typical choice of  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\Sigma$ ,  $d = 3$ ;  $\mu \approx 0.6763$

Tolerance	Method	Median		n	Worst 10% Time (s)
		Error	Accuracy		
1E-4	Sh. Latt.	4E-7	100%	8.2E3	0.014
1E-4	Scr. Sobol'	4E-7	100%	1.6E4	0.018
1E-4	Bayes. Latt.	5E-7	100%	8.2E3	0.010

These algorithms are implemented in GAIL<sup>4</sup>

<sup>4</sup>Choi, S.-C. T. et al. GAIL: Guaranteed Automatic Integration Library (Versions 1.0–2.2). MATLAB software. 2013–2017. [http://gailgithub.github.io/GAIL\\_Dev/](http://gailgithub.github.io/GAIL_Dev/).



# Option Pricing

$$\mu = \text{fair price} = \int_{\mathbb{R}^d} e^{-rT} \max \left( \frac{1}{d} \sum_{j=1}^d S_j - K, 0 \right) \frac{e^{-z^T z / 2}}{(2\pi)^{d/2}} dz \approx \$13.12$$

$S_j = S_0 e^{(r-\sigma^2/2)jT/d + \sigma x_j}$  = stock price at time  $jT/d$ ,

$$x = Az, \quad AA^T = \Sigma = \left( \min(i, j) T/d \right)_{i,j=1}^d, \quad T = 1/4, \quad d = 13 \text{ here}$$



Abs. Error Tolerance	Method	Median Error	Median Accuracy	Worst 10% $n$	Worst 10% Time (s)
1E-2	IID	diff	2E-3	100%	6.1E7
1E-2	Scr. Sobol'	PCA	1E-3	100%	1.6E4
1E-2	Scr. Sob. cont. var.	PCA	2E-3	100%	4.1E3
1E-2	Bayes. Latt.	PCA	2E-3	99%	1.6E4



# Future Work

- Bayesian cubature with higher order nets and smoother kernels, control variates
- Stopping criteria for multilevel and multivariate decomposition methods
- Community QMC software that **combines the efforts of several research groups**
  - Skeleton with clearly defined properties for different kinds of objects
  - Plug-and-play functionality

```
%% Example
stopObj = CLTStopping %stopping criterion
distribObj = IIDDistribution; %IID sampling with uniform distribution
[sol, out] = integrate(KeisterFun, distribObj, stopObj)
stopObj.absTol = 1e-3; %decrease tolerance
[sol, out] = integrate(KeisterFun, distribObj, stopObj)

>> IntegrationExample
sol =
0.428567222603452
sol =
0.425203913946775
```

# Thank you

---

Slides available on SpeakerDeck at  
[speakerdeck.com/fjhickernell/matrix-week-two-2018-june](https://speakerdeck.com/fjhickernell/matrix-week-two-2018-june)



-  H., F. J., Jiang, L., Liu, Y. & Owen, A. B. *Guaranteed Conservative Fixed Width Confidence Intervals Via Monte Carlo Sampling.* in *Monte Carlo and Quasi-Monte Carlo Methods 2012* (eds Dick, J., Kuo, F. Y., Peters, G. W. & Sloan, I. H.) **65** (Springer-Verlag, Berlin, 2013), 105–128.
-  H., F. J. & Li. A. Jiménez Rugama. *Reliable Adaptive Cubature Using Digital Sequences.* in *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014* (eds Cools, R. & Nuyens, D.) **163**. arXiv:1410.8615 [math.NA] (Springer-Verlag, Berlin, 2016), 367–383.
-  Li. A. Jiménez Rugama & H., F. J. *Adaptive Multidimensional Integration Based on Rank-1 Lattices.* in *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014* (eds Cools, R. & Nuyens, D.) **163**. arXiv:1411.1966 (Springer-Verlag, Berlin, 2016), 407–422.
-  H., F. J., Jiménez Rugama, Li. A. & Li, D. *Adaptive Quasi-Monte Carlo Methods for Cubature.* in *Contemporary Computational Mathematics — a celebration of the 80th birthday of Ian Sloan* (eds Dick, J., Kuo, F. Y. & Woźniakowski, H.) (Springer-Verlag, 2018), 597–619.  
doi:10.1007/978-3-319-72456-0.
-  Jagadeeswaran, R. & H., F. J. *Automatic Bayesian Cubature.* in preparation. 2018+.



-  Choi, S.-C. T. *et al.* *GAIL: Guaranteed Automatic Integration Library (Versions 1.0–2.2)*. MATLAB software. 2013–2017. [http://gailgithub.github.io/GAIL\\_Dev/](http://gailgithub.github.io/GAIL_Dev/).
-  (eds Cools, R. & Nuyens, D.) *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014*. 163 (Springer-Verlag, Berlin, 2016).