4-26-2012

An Overview of Quasi-Monte Carlo Methods in Control Systems

Chaouki T. Abdallah

P.F. Hokayem

P. Dorato

S. Mastellone

An Overview of Quasi-Monte Carlo Methods in Control Systems

P.F. Hokayem*† C.T. Abdallah‡† P. Dorato‡ S. Mastellone‡

Introduction

Many control problems are so complex that analytic techniques fail to solve them [2]. Furthermore, even when analytic solutions are available, they may be computationally costly [2] and generally result in very high-order compensators [3]. Due to these reasons, we tend to accept approximate answers which provide us with certain performance guarantees for such problems. Sampling methods thus come into the picture to try and remedy the "cost of solution" problem by drawing samples from an appropriate space, and providing an approximate answer.

For many years, random sampling has dominated the afore mentioned arena [8, 11, 4]. Random sample generation, with a uniform underlying distribution, however tends to cluster the samples on the boundary of the sample space in higher dimensions. It is for this reason that we are interested in presenting a method that distributes the points *regularly* in the sample space while providing deterministic guarantees on the error involved. Recently, deterministic or quasi-Monte Carlo (QMC) methods have proven superior to random methods in several applications such as the calculation of certain integrals [6], financial derivatives [7] and motion planning in robotics [10]. They have also been used for stability analysis of high speed networks [9]. In this work, we provide an overview of such deterministic quasi-Monte Carlo method of sampling, and their applications to control systems analysis and design.

We present the basic concepts pertaining to quasi-Monte Carlo deterministic sampling. Such concepts include the following: Indicator functions, performance objective, generation of point sets, total variation, and error bounds.

Indicator Function

Our main justification for using various random and quasi-random sampling in control is that we are able to check whether a certain performance criterion is "approximately" met. To make this claim however, hinges on our ability to evaluate our criterion at single points in the parameter space. Hence the need for an indicator function that could provide us with such an answer. An example indicator function (Ψ) for the stability of a certain characteristic polynomial ($T(s,p,q)$) may be as follows

$$\Psi(P_i, Q_j) = \begin{cases} 1, & T(s,p,q) \text{ is stable} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where P_i and Q_j are sampled vectors from the plant parameter space and admissible controller parameter space, respectively. In a recent article by the authors [14], the above indicator function was relaxed into a differentiable one, hence expanding our ability to quantify the error involved in the approximation, and greatly enlarging the class of control problems solved.

Having the ability to evaluate our performance at single points, and taking a large enough number of samples (N), enable us to approximately evaluate our performance objective on the uncertain parameter space as follows,

$$f_{Q^*}(P_i) = f(P_i, Q^*) = \frac{1}{N} \sum_{i=1}^{N} \Psi(P_i, Q^*) \tag{2}$$

where f is called the counting function. The function $f_Q(P)$ can be interpreted as the average performance of the uncertain system with respect to a certain controller Q_i, in other words it is an approximation of the integration of the performance function over the plant parameter space. Hence, our problem is cast into an approximate integration setting. This integration perspective of

*Coordinated Science Laboratory, University of Illinois, 1308 W. Main St., Urbana, IL 61801, USA: hal@control.csl.uiuc.edu

†C.T. Abdallah and P.F. Hokayem research is partially supported by NSF-0233205.

‡Electrical & Computer Engineering Department, MSC01 1100, 1 University of New Mexico, Albuquerque, NM 87131-0001: {chaouki, peter, silvia}@eece.unm.edu

the problem permits us to exploit various sampling methods ranging from simple gridding to random and deterministic point set generation. The latter being our focus in this presentation, we review next the quasi-Monte Carlo techniques for generating deterministic samples.

Point Sets Generation

In this section we describe one method of generate "low discrepancy" quasi-Monte Carlo points in an d-dimensional sample space. The *discrepancy* is a measure of the regularity in distribution of a set of points in the sample space. Since the points result from a deterministic method of generation, they possess a certain regularity property of distribution in the sample space described by their discrepancy. This gives the method leverage over Monte Carlo methods in the sense that the guarantees over the error magnitude are deterministic.

Van Der Corput [1]

The van der Corput sequence in base b, where $b \geq 2 \in \mathbb{N}$, is a one dimensional sequence of points that possesses the property of having a certain regularity (low discrepancy) in the unit interval $\mathbb{I} = [0, 1] \subset \mathbb{R}$. *The main idea is to express every integer $n \in \mathbb{N}$ in base b and then reflect the expansion into the unit interval $\mathbb{I}$.* This is done as follows:

1. Let $R_b = \{0, 1, \ldots, b-1\}$ be the residue set modulo b
2. Any integer $n \geq 0$ can be expanded in base b as:

$$n = \sum_{k=0}^{\infty} a_k(n) b^k \tag{3}$$

   where $a_k(n) \in R_b, \forall k$.

3. Finally, we get the sequence $\{X_n\}$ as

$$X_n = \phi_b(n) = \sum_{k=0}^{\infty} a_k(n) b^{-j-1}. \tag{4}$$

As will be seen, the van der Corput sequence will be used to generate higher dimensional vector samples, with the variation of the expansion base b.

Halton Sequence [1]

The Halton sequence is a generalization of the van der Corput sequence to span an d-dimensional sample space. The main idea is to generate d 1-dimensional sequences and form the corresponding d-dimensional vector sample points. Let $b_1, b_2, \ldots, b_d$ be the corresponding expansion bases for each dimension, preferably relatively prime[1]. Let $\phi_{b_1}, \phi_{b_2}, \ldots, \phi_{b_d}$ be the corresponding reflected expansions according to the corresponding bases. Then the d-dimensional sequences $\{X_n^d\}$ are formed as follows:

$$X_n = (\phi_{b_1}, \phi_{b_2}, \ldots, \phi_{b_d}) \in \mathbb{I}^d \tag{5}$$

Discrepancy

As mentioned earlier, the *discrepancy* is a measure of the regularity in distribution of a set of points in the sample space. In order to define it mathematically, we need to introduce the following counting function:

$$A(B; P) = \sum_{i=1}^{N} I_B(X_i) \tag{6}$$

where $B \subset \mathbb{I}^d$ is an arbitrary set, $P = (X_1, \ldots, X_N)$ is a point set, N is the number of points, and I_B is an indicator function.

[1] Choosing the expansion bases relatively prime reduces the discrepancy, hence the error bound

Definition 1 *The general formula for the evaluation of the discrepancy is given by*

$$\mathcal{D}_N(\mathcal{B}, P) = \sup_{B \subset \mathcal{B}} \left| \frac{A(B,P)}{N} - \lambda_d(B) \right| \tag{7}$$

where $\lambda_d(B)$ is the d-dimensional Lebesgue measure of the arbitrary set B and $\mathcal{B}$ is the family of all lebesgue measurable subsets B of $\mathbb{I}^d$.

Definition 1 can be specialized into the following cases: The *star discrepancy* $D_N^\star(X_1, \ldots, X_N)$ is obtained by letting $\mathcal{B}$ in (7) be defined as follows

$$\mathcal{B}^\star = \{\forall B : B = \prod_{i=1}^{d} [0, u_i)\}$$

i.e. the set of all d-dimensional subsets of $\mathbb{I}^d$ that have a vertex at the origin, and u_i's being arbitrary points in the corresponding 1-dimensional space.

As an example for calculating star discrepancy of the Halton sequence with relatively prime expansion bases is given by (see [1])

$$D_N^\star(X_1, \ldots, X_N) < \frac{d}{N} + \frac{1}{N} \prod_{i=1}^{d} \left(\frac{b_i - 1}{2 \log b_i} \log N + \frac{b_i + 1}{2} \right). \tag{8}$$

Total Variation

The problem of bounding the error involved in evaluating the integral of a function using quasi-Monte Carlo methods depends on our ability to obtain the value of total variation of the function under consideration, as will be seen in the next section. Consequently, in this section we will concentrate on defining several notions of variation of a function defined on an interval $[0,1]^d$.

Definition 2 *[12] A finite function $f(x)$ defined on and interval [0,1] is said to have* 'bounded variation' *if there exists a number M, such that for any partition p of the interval [0,1]*

$$v_p = \sum_{i=1}^{n} |f(X_i) - f(X_{i-1})| < M.$$

Moreover, the 'total variation' *of $f(x)$ on $[0,1]$ is defined as $V(f) = \sup_{p \in \mathcal{P}} (v_p)$, where $\mathcal{P}$ is the set of all partitions on [0,1].*

Notice that Definition 2 pertains to functions of a single variable and does not require that the function be continuous. However, the function has to have a countable number of discontinuities on the interval under study. If it is further assumed that the function $f(x)$ is differentiable on $[0,1]$, then the total variation is defined as follows: $V(f) = \int_0^1 \left| \frac{df}{dx} \right| dx$.

Note 1 *The total variation of a function can be understood as the sum of all the heights of monotone segments. That is why we integrate over the absolute value of the gradient.*

The total variation of a function f defined on a one-dimensional unit interval $\mathbb{I} = [0,1]$ is fairly easy to calculate. However, if f is defined on $\mathbb{I}^d$ the problem of calculating $V^{(d)}(f)$ (the d-dimensional total variation) is more involved (see [15, 1]). In what follows we only present the definitions of the total variation for continuous and differentiable functions.

Definition 3 *The total variation of a function f defined on $\mathbb{I}^d$ in the sense of Vitali is defined as*

$$V^{(d)} = \int_0^1 \cdots \int_0^1 \left| \frac{\partial^{(d)} f}{\partial \eta_1 \partial \eta_2 \ldots \partial \eta_d} \right| d\eta_1 d\eta_2 \ldots d\eta_d. \tag{9}$$

whenever the indicated partial derivative is continuous on $\mathbb{I}^d$. If $V^{(d)} < +\infty$, then the function f is said to have a 'bounded total variation in the sense of Vitali'.

Note that the Definition 3 only measures the variation of f over all the variables at once. However, indicated partial derivative in (9) might be zero, but still the variation over the domain is not equal to zero (see [14]).

The problem encountered in Definition 3 can be remedied via the following enhanced definition of the total variation.

Definition 4 *[16, 1] Let f be a function defined on $\mathbb{I}^d$ with bounded variation in the sense of Vitali. Suppose that the restriction of f to each face F of $\mathbb{I}^d$ of dimension $k = 1, 2, \ldots, d-1$ is also of bounded variation on F in the sense of Vitali. Then the function f is said to be of* 'bounded variation in the sense of Hardy and Krause'.

Note 2 *The restriction of the function f to the face F in definition 4 is achieved through setting the $d-k$ variables equal to* 1.

Definition 4 overcomes the difficulties we encountered with Definition 3, as illustrated via a an example in [13].

Error in Quasi-Monte Carlo

The error in quasi-Monte Carlo methods of integration over the unit hypercube for N samples is defined as follows,

$$e = \int_{\mathbb{I}^d} f(\eta) d\eta - \frac{1}{N} \sum_{n=1}^{N} f(X_n) \tag{10}$$

The following two theorems provide bounds on the error (10), for the cases of 1-dimensional and d-dimensional integration, respectively.

Theorem 1 *Koksma's Inequality [1]*
Let $f(.)$ be a function defined on $\mathbb{I} = [0, 1]$ of bounded total variation
$\left| \int_{\mathbb{I}^d} f(\eta) d\eta - \frac{1}{N} \sum_{i=1}^{N} f(X_n) \right| \leq V(f) D_N^{\star}(X_1, \ldots, X_N)$.

Theorem 2 *Koksma-Hlawka Inequality [1]*
Let $f(.)$ be a function defined on $\mathbb{I}^d$ of bounded variation in the sense of Hardy and Krause
$\left| \int_{\mathbb{I}^d} f(\eta) d\eta - \frac{1}{N} \sum_{i=1}^{N} f(X_n) \right| \leq V^{(d)}(f) D_N^{\star}(X_1, \ldots, X_N)$.

Basically, Theorems 1 and 2 state that the magnitude of the error depends on the total variation of the function and the star discrepancy of the point set chosen. That is why we are always after low star discrepancy point sets in quasi-Monte Carlo methods. It is also worth mentioning that the error bounds are conservative, i.e. if the variation of the function is large, we get a large bound on the error, although the actual error might be small.

In our presentation we will provide several control analysis and design examples and a novel method that facilitates the use of total variation of the indicator function, and consequently calculating the error bound. Our approach is applied to both linear and nonlinear systems.

References

[1] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods,* SIAM, 1992.

[2] V. Blondel and J. Tsitsiklis, "A survey of computational complexity results in systems and control," Preprint available at http://www.ulg.ac.be/mathsys/blondel/publications.html, 1998.

[3] M. Vidyasagar, *Control Systems Synthesis: A Factorization Approach,* MIT Press, Cambridge, MA, 1985.

[4] R. Tempo and F. Dabbene, "Randomized Algorithms for Analysis and Control of Uncertain Systems: An Overview", *Perspectives in Robust Control - Lecture Notes in Control and Information Science*, (ed. S.O. Moheimani), pp. 347-362, Springer-Verlag, London, 2001.

[5] J.G. Truxal, "Control Systems-Some Unusual Design Problems", in *Adaptive Control Systems,* (Eds. E. Mishkin and L. Braun), McGraw-Hill, NY, 1961.

[6] A. Papageorgiou and J.G. Traub, "Faster Evaluation of Multidimesional Integrals", *Computers in Physics*, pp. 574-578, Nov., 1997.

[7] S. Paskov and J.G. Traub, "Faster Valuation of Financial Derivatives", *Journal of Portfolio Management*, Vol. 22:1, pp. 113-120, Fall, 1995.

[8] V. Koltchinskii, C.T. Abdallah, M. Ariola, P. Dorato and D. Panchenko, "Improved Sample Complexity Estimates for Statistical Learning Control of Uncertain Systems", *IEEE Trans. Automatic Control*, vol.45, no.12, pp.2383-2388, Dec. 2000.

[9] T. Alpcan, T. Basar, and R. Tempo, "Randomized Algorithms for Stability and Robustness Analysis of High Speed Communication Networks", *IEEE Trans. on Control Systems Technology*, submitted, July 2002.

[10] M.S. Branicky, S.M. LaValle, K. Olson, L. Yang, "Deterministic vs. Probabilistic Roadmaps", *IEEE Trans. on Robotics and Automation*, submitted, Jan. 2002.

[11] V. Koltchinskii, C.T. Abdallah, M. Ariola, P. Dorato and D. Panchenko, "Statistical learning control of uncertain systems: It is better than it seems", Tech. Rep.EECE 99-001, EECE Dept., The University of New Mexico, Feb. 1999.

[12] I.N. Bronshtein and K.A. Semendyayev, *Handbook of Mathematics*, Springer, Berlin, 1998.

[13] P.F. Hokayem, C.T. Abdallah, P. Dorato, "Quasi-Monte Carlo Methods in Robust Control", 11^{th} Mediterranean Conference on Control and Automation, Rhodes, 2003.

[14] P.F. Hokayem, C.T. Abdallah, P. Dorato, "Quasi-Monte Carlo Methods in Robust Control", Accepted to 42^{nd} IEEE Conference on Decision and Control, Maui, Hawaii, USA, 2003.

[15] L.K. Hua and Y. Wang, *Applications of Number Theory to Numerical Analysis*, Springer-Verlag, Berlin, 1981.

[16] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley, NY, 1974.