Convergence of solutions of approximate random equations

W. Römisch

Humboldt-University Berlin
Institute of Mathematics

www.math.hu-berlin.de/~romisch

DFG Research Center MATHEON
Mathematics for key technologies

Workshop *PDEs with Random Coefficients*
WIAS Berlin, November 13-15, 2013

Introduction

(Partial) differential and integral equations with random coefficients and random right-hand sides are now within reach of <u>efficient</u> computational methods.

The latter require a combination of discretization and sampling techniques and a specific theoretical justification.

New sampling methods based on randomized lattice rules or (interlaced) polynomial lattice rules (specific Quasi-Monte Carlo methods) are available, which led to a breakthrough in high-dimensional numerical integration and to lifting the curse of dimension (recent work of Sloan-Kuo-Joe and Dick-Pillichshammer).

A general framework for studying approximations of random equations is already available since more than 20 years, but seemingly unknown. Aim of the talk is to start with a kind of review of those developments, to show its relevance for recent research and to add some updates.

Recent work on random partial differential equations

Simultaneous stochastic Galerkin and FE methods

Elliptic partial differential equation with random coefficients

$$-\mathrm{div}\,(a(\xi, z(\omega))\nabla x(\xi,\omega)) = f(\xi)\ (\xi \in D),\ x(\xi,\omega) = 0\ (\xi \in \partial D).$$

where $D \subset \mathbb{R}^m$, $z = (z_j)_{j\in\mathbb{N}}$ is uniformly distributed in $[0,1]^{\mathbb{N}}$,

$$a(\xi, z) = \bar{a}(\xi) + \sum_{j=1}^{\infty} z_j \psi_j(\xi)\ (\xi \in D)\ \ (\text{Karhunen-Loeve expansion}).$$

Variational formulation:

$$X = H_0^1(D) \subset H = L_2(D) \subset X^* = H^{-1}(D),\ \|v\|_X = \|\nabla v\|_H$$

$$\int_D a(\xi, z(\omega))\langle \nabla x(\xi), \nabla v(\xi)\rangle d\xi = \int_D f(\xi) v(\xi) d\xi \quad (\forall v \in X).$$

s-term N-sample QMC scheme: $a(\xi, z_N^s) = \bar{a}(\xi) + \sum_{j=1}^{s} z_j \psi_j(\xi)$
FE method: Replace X by a finite element subspace X_h.

(Kuo-Schwab-Sloan: SIAM J. Num. Anal. 2012)

Random operator equations: Earlier work

Let (X, d) be a separable complete metric space, Y and Z separable metric spaces and $0 \in Y$ fixed. All metric spaces are endowed with their Borel σ-fields. Consider the random operator equation

$$T(x, z(\omega)) = 0 \quad (\omega \in \Omega), \qquad (*)$$

where $T : X \times Z \to Y$ is a mapping and z is a Z-valued random variable given on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Question: (Existence of a random solution (Hanš 56))

Does there exist a measurable map $x : \Omega \to X$ such that

$$T(x(\omega), z(\omega)) = 0, \ \mathbb{P}\text{-almost surely}$$

if the equation $T(x, z(\omega)) = 0$ is solvable for all $\omega \in \Omega$? Is the solution concept appropriate for approximations ?

Note that measurability of $T(x, \cdot)$ $(\forall x \in X)$ together with $\mathbb{P}$ -a.s. solvability of $T(x, z(\omega)) = 0$ is not sufficient!

Theorem: (Engl 78, Nowak 78)

Assume that $S(\omega) = \{x \in X : T(x, z(\omega)) = 0\} \neq \emptyset$ $\mathbb{P}$ -a.s.

If $[T(\cdot, z(\cdot))]^{-1}(0) \in \mathcal{F} \times \mathcal{B}(X)$, there exists a random solution of $(*)$.

The condition $[T(\cdot, z(\cdot))]^{-1}(0) \in \mathcal{F} \times \mathcal{B}(X)$ is implied by (a) or (b):
(a) T is Borel measurable.
(b) T is a Carathéodory mapping, i.e., $T(\cdot, z)$ is continuous for all $z \in Z$ and $T(x, \cdot)$ is measurable for all $x \in X$.

Extensions to set-valued mappings and stochastic domains are available.

A (Borel) probability measure $\mu_X \in \mathcal{P}(X)$ is called weak solution of

$$T(x, z(\omega)) = 0 \quad (\omega \in \Omega)$$

iff there exists $\mu \in \mathcal{P}(X \times Z)$ such that

$$\mu T^{-1} = \delta_0 \,, \quad \mu_X = \mu p_X^{-1} \,, \quad \mathcal{L}(z) = \mu p_Z^{-1},$$

where $\mathcal{L}(z) = \mathbb{P} z^{-1} : \mathcal{B}(Z) \to [0, 1]$ is the probability distribution (or law) of z, p_X and p_Z are the projections from $X \times Z$ to X and Z, respectively, and $\delta_0 \in \mathcal{P}(Y)$ denotes the Dirac measure placing unit mass at $0 \in Y$.

Remark: If $x : \Omega \to X$ is a random solution, $\mathcal{L}(x)$ is a weak solution (by putting $\mu = \mathcal{L}(x, z)$). If $[T(\cdot, z)]^{-1}(0)$ is a singleton for all $z \in Z$, the weak solution is unique. A weak solution of a random operator equation is a random solution on some probability space.

Random operator equations: Approximations

Let X, Y and Z be complete separable metric (Polish) spaces, $T : X \times Z \to Y$ be Borel measurable, z be a Z-valued random variable (on $(\Omega, \mathcal{F}, \mathbb{P})$) and $0 \in Y$ be fixed. We consider

$$T(x, z(\omega)) = 0 \quad (\omega \in \Omega).$$

In addition, we consider the approximate random operator equations

$$T_n(x, z_n(\omega)) = 0 \quad (\omega \in \Omega_n; n \in \mathbb{N}),$$

where for each $n \in \mathbb{N}$, $X_n \subset X$, $Z_n \subset Z$, $T_n : X_n \times Z_n \to Y$ be Borel measurable and z_n be a Z_n-valued random variable (on $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$).

Let (x_n) be a sequence of random solutions to the approximate random operator equations.

Motivation: Approximation procedures for solving random equations require a '*discretization*' of T and an *approximation* ('sampling', 'estimation') of z.

A. T. Bharucha-Reid (Ed.): Approximate Solution of Random Equations, North-Holland, New York 1979.

Weak convergence in $\mathcal{P}(X)$: (μ_n) converges weakly to μ iff

$$\lim_{n\to\infty} \int_X f(x)\mu_n(dx) = \int_X f(x)\mu(dx) \qquad \forall f \in C_b(X,\mathbb{R}).$$

The topology of weak convergence is metrizable if X is separable.
Weak compactness is characterized by uniform tightness due to Prokhorov's theorem.

Problem: Find conditions on (T_n) and T that imply weak convergence of $(\mathcal{L}(x_n))$ if $(\mathcal{L}(z_n))$ converges weakly to $\mathcal{L}(z)$.

A sequence (T_n) converges discretely to T iff

(i) $d(x, X_n) = \inf_{y\in X_n} d_X(x,y) \to 0 \quad (\forall x \in X)$,
$d(z, Z_n) = \inf_{v\in Z_n} d_Z(z,v) \to 0 \quad (\forall z \in Z)$.

(ii) For all $(x,z) \in X \times Z$ and sequences $(x_n, z_n) \in X_n \times Z_n$ such that $x_n \to x$ in X and $z_n \to z$ in Z it holds

$$T_n(x_n, z_n) \to T(x,z) \quad (\text{in } Y).$$

(Stummel, Reinhardt, Vainikko)

Theorem: (Engl-Rö 87, Fiedler-Rö 92)
Let the following conditions be satisfied:

(a) $\bigcup_{n\in\mathbb{N}} T_n^{-1}(\{0\}) \cap (K \times B)$ is relatively compact in $X \times Z$ for each bounded $B \subset X$ and compact $K \subset Y$.

(b) (T_n) converges discretely to T.

(c) $(\mathcal{L}(z_n))$ converges weakly to $\mathcal{L}(z)$.

(d) The set of laws of random solutions $\{\mathcal{L}(x_n) : n \in \mathbb{N}\}$ is stochastically bounded, i.e., for each $\varepsilon > 0$ there exists a bounded Borel set $B_\varepsilon \subset X$ such that

$$\inf_{n\in\mathbb{N}} \mathcal{L}(x_n)(B_\varepsilon) \geq 1 - \varepsilon\,.$$

Then the set $\{\mathcal{L}(x_n) : n \in \mathbb{N}\}$ is relatively compact with respect to the weak topology and each weak limit of a subsequence is a weak solution of the random operator equation $T(x, z(\omega)) = 0 \;\; (\omega \in \Omega)$.
If the weak solution of $T(x, z(\omega)) = 0 \;\; (\omega \in \Omega)$ is unique, the sequence $(\mathcal{L}(x_n))$ converges weakly to this weak solution.

Corollary:

Let the following conditions be satisfied:

(a) Assume that for each compact $K \subset Z$ there exist $C = C(K) > 0$ and $n_0 = n_0(K) \in \mathbb{N}$ such that

$$d_X(x, \tilde{x}) \leq C\, d_Y(T_n(x, \tilde{z}), T_n(\tilde{x}, \tilde{z}))$$

holds for all $n \geq n_0$, $\tilde{z} \in K$, $x, \tilde{x} \in X_n$ (with d_X and d_Y denoting the distances in X and Y, respectively).

(b) (T_n) converges discretely to T.

(c) $(\mathcal{L}(z_n))$ converges weakly to $\mathcal{L}(z)$.

If a weak solution of $T(x, z(\omega)) = 0 \quad (\omega \in \Omega)$ exists, it is unique and the sequence $(\mathcal{L}(x_n))$ converges weakly to this weak solution.

There exist several other sets of more specific conditions implying the assumptions of the approximation theorem.

Nonlinear elliptic PDEs with random coefficients

We consider the random nonlinear elliptic PDE

$$-\sum_{i=1}^{m} \frac{\partial}{\partial \xi_i} a_i\Big(\xi, z(\xi,\omega), \frac{\partial x}{\partial \xi_1}, \dots, \frac{\partial x}{\partial \xi_m}\Big) = f(\xi) \ \ (\xi \in D), \ x(\xi) = 0 \ (\xi \in \partial D).$$

or

$$-\mathrm{div}\, a(\xi, z(\xi,\omega), \nabla x(\xi)) = f(\xi) \ \ (\xi \in D), \ x(\xi) = 0 \ (\xi \in \partial D).$$

where $D \subset \mathbb{R}^m$ is open such that $\bar{D}$ is a bounded polyhedron, $a \in C_m(D \times \mathbb{R} \times \mathbb{R}^m)$ such that $a(\xi, r, \cdot) \in C_m^1(\mathbb{R}^m)$, $a(\xi, r, 0, \dots, 0) = 0$ and

$$\Big|\frac{\partial a_i}{\partial r}(\xi, r, t_1, \dots, t_m)\Big| \le M\,, \ \ \Big|\frac{\partial a_i}{\partial t_j}(\xi, r, t_1, \dots, t_m)\Big| \le M \ \ (i, j = 1, \dots, m),$$

$$\sum_{i,j=1}^{m} \frac{\partial a_i}{\partial t_j}(\xi, r, t_1, \dots, t_m) y_i y_j \ge \gamma \|y\|^2 \ \ (y \in \mathbb{R}^m)$$

holds uniformly for all $\xi \in \bar{D}$, $t \in \mathbb{R}^m$, $r \in \mathbb{R}$ and some $\gamma > 0$ and $K > 0$; $f \in C(\bar{D})$, z a Z-valued random variable, where $Z = L_\infty(D)$.

Variational formulation:

$$X = H_0^1(D) \subset H = L_2(D) \subset X^* = H^{-1}(D), \ \|x\|_X = \|\nabla x\|$$

$$\langle A(x, z(\omega)), v\rangle = \int_D \langle a(\xi, z(\xi,\omega), \nabla x(\xi)), \nabla v(\xi)\rangle d\xi = \int_D f(\xi)v(\xi)d\xi \ \ (\forall v \in X)$$

or

$$\langle A(x, z(\omega)), v\rangle = \langle f, v\rangle \ \ (\forall v \in X)$$

or

$$T(x, z(\omega)) = A(x, z(\omega)) - f = 0, \qquad (*)$$

where $T : X \times Z \to Y = X^*$ and $\langle \cdot, \cdot \rangle$ denotes the dual pairing of X and X^*.

The assumptions imply that $T(\cdot, z)$ is strongly monotone (with γ) and Lipschitz continuous with modulus M (uniformly with respect to z).

Let X_n be a finite element subspace of X such that $d(x, X_n) \to 0 \ \forall x \in X$ and $I_n : X_n \to X$ denotes the embedding mapping.

We consider the Galerkin equations $T_n(x_n, z_n(\omega)) = 0$ $(x_n \in X_n)$ defined by

$$\langle I_n^* T(x_n, z_n(\omega)), v\rangle = \langle T(x_n, z_n), v\rangle = 0 \ \ (v \in X_n).$$

Corollary:

Assume that $(\mathcal{L}(z_n))$ converges weakly to $\mathcal{L}(z)$.
Then each Galerkin equation has a unique random solution and the sequence $(\mathcal{L}(x_n))$ converges weakly to the law $\mathcal{L}(x)$ of the unique random solution $x(\cdot)$.
Furthermore, it holds $\mathbb{P}$-a.s.

$$\|x_n(\omega) - x(\omega)\|_X \leq \frac{M}{\gamma}\big(d(x(\omega), X_n) + \|z_n(\omega) - z(\omega)\|_Z\big).$$

If we impose the regularity condition $d(x(\omega), X_n) \leq C\, r_n$ for each $n \in \mathbb{N}$, some uniform constant $C > 0$ and a sequence (r_n) converging to zero, one obtains

$$\rho_X(\mathcal{L}(x_n), \mathcal{L}(x)) \leq \frac{M}{\gamma}\big(C\, r_n + \rho_Z(\mathcal{L}(z_n), \mathcal{L}(z))\big),$$

where ρ_X and ρ_Z are the Prokhorov distances on the set of probability measures on X and Z, respectively, metrizing the topologies of weak convergence.

Hence, the rate of weak convergence of the approximate solutions to the original one is equal to the sum of the rates of convergence of the finite element method and of the approximate stochastic coefficients.

Assume that z admits a Karhunen-Loéve (KL) type expansion of the form

$$z(\xi, \omega) = \sum_{i=1}^{\infty} z_i(\omega)\psi_i(\xi) \;\; (\xi \in \bar{D}, \omega \in \Omega),$$

where the real random variables z_i, $i \in \mathbb{N}$ are independent and uniformly distributed in $[0, 1]$ (or $[-0, 5, 0, 5]$).

Hence, an approximation z_n may be obtained by considering s terms of the KL expansion and replacing the s-dimensional random vector $(z_1, \dots, z_s)$ by N uniform samples $z^j \in [0, 1]^s$, $j = 1, \dots, N$. This means that $z_n(\xi) = z_{s,N}(\xi)$ has the N realizations

$$z_{nj}(\xi) = \sum_{i=1}^{s} z_i^j \psi_i(\xi) \quad (\xi \in \bar{D}, j = 1, \dots, N).$$

Examples:

(a) Monte Carlo methods: Select N independent samples from the uniform distribution on $[0, 1]^s$. The best possible convergence rate for approximating $\mathcal{L}(z)$ is then $O(n^{-\frac{1}{2}})$.

(b) Quasi-Monte Carlo methods: Select N QMC points $z^j \in [0,1]^s$ such that

$$e_{N,s}(z^1,\ldots,z^N) = \sup_{g\in\mathbb{B}_s}\Big|\int_{[0,1]^s} g(t)dt - N^{-1}\sum_{j=1}^{N} g(z^j)\Big|,$$

where $\mathbb{B}_s$ is the unit ball of a linear normed space $\mathbb{F}_s$, has a good convergence rate with a constant not depending on s.

Suitable spaces are kernel reproducing Hilbert spaces (or related Banach spaces), e.g., the anchored weighted tensor product Sobolev space

$$\mathbb{F}_s = \mathcal{W}^{(1,\ldots,1)}_{2,\gamma,\mathrm{mix}}([0,1]^s) = \bigotimes_{j=1}^{s} W^1_{2,\gamma_j}([0,1])$$

with the weighted inner product ($c \in [0,1]^s$ denotes the "anchor")

$$\langle f,g\rangle_\gamma = \sum_{u\subseteq\{1,\ldots,d\}} \gamma_u^{-1}\int_{[0,1]^{|u|}} \frac{\partial^{|u|} f}{\partial x_u}(x_u,c_{-u})\frac{\partial^{|u|} g}{\partial x_u}(x_u,c_{-u})dx_u$$

or weighted Walsh spaces containing smooth functions.

Examples of QMC methods:

(i) randomly shifted lattice rules (Sloan-Kuo-Joe 02):

$$z^j = \Big\{\frac{j-1}{N} g + \triangle\Big\},\ j = 1, \ldots, N,$$

where $g \in \mathbb{Z}^s$ is the generator of the lattice, $\triangle$ is uniformly distributed in $[0,1]^s$ and $\{\cdot\}$ means taking the fractional part componentwise, allow the mean square convergence estimate

$$\hat{e}_{N,s}(z^1, \ldots, z^N) \le C(\delta) N^{-1+\delta} \quad (\delta \in (0, 0, 5])$$

if the weights (γ_j) converge sufficiently fast to 0 and γ_u is of product form.

(ii) interlaced polynomial lattice rules (Dick-Goda 13):

$$e_{N,s}(z^1, \ldots, z^N) \le C\, N^{-\frac{1}{p}} \quad (p \in (0,1)).$$

The good convergence rates require smooth integrands and smooth dependence of the solution on the random parameter.

Hence, we consider the case $a(\xi, r, t) = \sum_{j=1}^{m} A_j(\xi, r) t_j$ and KL expansion of $A(\xi, z(\omega))$.

We consider a lattice rule with uniform random shift $\triangle$ in $[0,1]^s$ and N points in s dimensions, and for each shifted lattice point solve the approximate elliptic problem

$$\int_D a(\xi, z(\omega))\langle \nabla x_h(\xi), \nabla v(\xi)\rangle d\xi = \int_D f(\xi)v(\xi)d\xi \quad (\forall v \in X_h).$$

in a finite element subspace X_h (on $D \subset \mathbb{R}^m$).

Theorem: (Kuo-Schwab-Sloan 12)
The convergence rate of the scheme is

$$\left(\mathbb{E}\left[\left|I(G(z)) - Q_{s,N}(\triangle; G(x_h))\right|^2\right]\right)^{\frac{1}{2}} \leq C\left(s^{-1} + N^{-1+\delta} + h^{\tau}\right),$$

where $0 \leq \tau = t + t' \leq 2$, N is prime, $\delta \in (0, \frac{1}{2}]$, $f \in H^{-1+t}(D)$, $G \in H^{-1+t'}(D)$, $p = \frac{2}{3}$ and $\sum_{j\in\mathbb{N}} \|\psi_j\|^p_{L_\infty(D)} < \infty$.

Conclusions

- Recently the numerical analysis of random differential equations became a very active field of research. In particular, combinations of generalized Wiener or Karhunen-Loéve expansions with multi-level Monte Carlo and Quasi-Monte Carlo methods became powerful tools.
- For linear elliptic PDE's with random coefficients new and promising error estimates are obtained for specific QMC finite element methods. They may be extended to certain regular analytic parametric operator equations (Schwab 13).
- A challenging question is: To which nonlinear random PDEs or even random variational problems may the QMC convergence rates be extended ? I am particularly interested in stochastic optimization models that may be rewritten as set-valued equations.

References

A. T. Bharucha-Reid and W. Römisch: Projective schemes for random operator equations. Weak compactness of approximate solution measures, *J. Integr. Equat.* 8 (1985), 95–111.

J. Dick, F. Y. Kuo, and I. H. Sloan: High dimensional integration – the Quasi-Monte Carlo way, *Acta Numerica* (2013), 1–153.

J. Dick, F. Y. Kuo, Q. T. Le Gia, D. Nuyens and C. Schwab: Higher order QMC Galerkin discretization for parametric operator equations (submitted 2013).

J. Dick and F. Pillichshammer: *Digital Nets and Sequences*, Cambridge University Press, 2010.

H. W. Engl: Random fixed point theorems, in: *Nonlinear Equations in Abstract Spaces* (V. Lakshmikantham, ed.), Academic Press, New York, 1978, 67–80.

H. W. Engl and M. Z. Nashed: Stochastic projectional schemes for random linear operator equations of the first and second kinds, *Num. Funct. Anal. Optim.* 1 (1979), 451–473.

H. W. Engl and W. Römisch: Approximate solutions of nonlinear random operator equations: Convergence in distribution, *Pac. J. Math.* 120 (1985), 55–77.

H. W. Engl and W. Römisch: Weak covergence of approximate solutions of stochastic equations with applications to random differential and integral equations, *Numer. Funct. Anal. Optim.* 9 (1987), 61–104.

O. Fiedler and W. Römisch: Weak convergence of approximate solutions of random equations, *Numer. Funct. Anal. Optim.* 13 (1992), 495–511.

F. Y. Kuo, C. Schwab and I. H. Sloan: Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients, *SIAM J. Numer. Anal.* 50 (2012), 3351–3374.

A. Nowak: Random solutions of equations, in: Trans. 8th Prague Conference on Information Theory, Stat. Decision Functions, and Random Processes, Vol. B, Czechoslovak Acad. Sci., Prague, 1978, 77–82.