Referee's report on the D.Sc. dissertation *Weak convergence of Galerkin finite element approximations of stochastic evolution equations with additive noise* by Mihály Kovács

The main topic of the dissertation concerns weak error estimates, that is, bounds on quantities of the form

$$|\mathbb{E}(\varphi(X^n) - \varphi(X))|$$

where X is the solution of an infinite-dimensional stochastic evolution equation

$$\partial_t X - LX = f(X) + \xi, \tag{1}$$

X^n is a temporal/spatial/full discretisation of it, and φ is a functional on the target space of X. The main examples of the linear operator L are the Laplacian, bi-Laplacian, wave operator, and time-convolution operator of Volterra equations. The latter example adds a twist of a non-Markovian setting. The nonlinearity f appears only in Chapter 2, where it is of 0 order (i.e. contains no derivatives), and the forcing ξ is either a Wiener or a Lévy noise. When f is not present, X is not actually an unknown quantity: it is given exactly by the Duhamel formula

$$X(t) = S(t)X(0) + \int_0^t S(t-s)\text{“}\xi(s)ds\text{”}, \tag{2}$$

where $S(t)$ is the solution operator process of L, which in most examples admits a closed form, and “$\xi(s)ds$” is interpreted as an Itô stochastic integral. The thesis is composed of 3 chapters, with varying assumptions of the above objects and various choices of approximations. In addition to the weak error estimates, in some of the cases the strong error is also discussed.

In Chapter 1, $f = 0$ and ξ is a Wiener noise. The class of φ considered here are of the form $\varphi(X) = g(X(T))$, for some sufficiently regular function g. The error estimates are based on the use of a drift-free backward Kolmogorov equation associated to a transformed process, a nice trick introduced by Debussche and Printemps. The main tool is summarised in a general setting in Theorem 1.2.1. It reduces controlling the weak error of approximating stochastic convolutions like the one above to the control of the solution of the Kolmogorov equation and the difference of the integrands. The former is automatic by posing sufficient assumptions on g (bounded second derivative typically suffices), while the for the latter one can invoke the deterministic bounds between the solution operator process $S(t)$ and its discretised counterparts. This is the strategy employed throughout the majority of the dissertation in Chapters 1 and 3.

First this method is applied to the wave equation. For the temporal approximation a family of discretisation is considered, parametrised by $p \in \mathbb{N}$, including standard methods like the implicit Euler scheme or the Crank-Nicholson scheme. The spatial discretisation

is a Galerkin finite element method parametrised by $\mathbb{N} \ni r \geq 2$, although as pointed out in Remark 1.3.10., the practical implementation for $r > 2$ gets quite involved: the existence and the construction of the finite elements becomes nontrivial and poses a number of assumptions on the domain. A final parameter β describes the pointwise in time L_2-based Sobolev regularity of the solution - which at the end of the day parametrises the regularity assumptions on the noise and the initial condition. With these notations in hand, Theorem 1.3.13. establishes a weak error rate of $\min(2\beta\frac{r}{r+1}, r)$ (spatially) and $\min(2\beta\frac{p}{p+1}, 1)$ (temporally).

Next the linearised CHC equation is considered. This corresponds to $L = -\Delta^2$ - I admit I do not understand the reason for the terminology: not only is the nonlinear term omitted from the true CHC equation, but the Laplacian as well, and so one in fact ends up with the biharmonic heat equation. The parameters are more specific in this case: the $p = 1$, $r = 2, 3$, $\beta \in (0, r/2]$ case is considered. The weak rate is shown to be 2β spatially and $\beta/2$ temporally in Theorem 1.4.1.

Finally, for Volterra equations, the spatial discretisation is like above, with $r = 2$, but the temporal discretisation requires some additional care due to the time-convolution involved in $LX(t) = \int_0^t b(t-s)AX(s)\,ds$. This is done by a quadrature rule due to Lubich. The weak rate is established in Theorem 1.5.33. to be 2β spatially and $\rho\beta$ temporally, in the regularity range $\beta \in (0, 1/\rho]$, where $\rho \in (1, 2)$ is the parameter for describing the smoothing effect of the equation. The latter appears in many forms, for instance via the bounds $\|A^s S(t)\| \lesssim t^{-s\rho}$ or $\|A^s \dot{S}(t)\| \lesssim t^{-s\rho-1}$, cf. the $\rho = 1$ in the Markovian case formally obtained by taking $b = \delta_0$. In particular the larger ρ is, the more regular the noise is required to be to guarantee X taking values in H^β.

Chapter 3 is close in spirit to Chapter 1: it considers still the linear case $f = 0$, functionals φ of the same form as above, but ξ this time is a Lévy process. The methodology is also very similar: once the weak error is represented via a drift-free Kolmogorov equation (Theorem 3.2.6.), one can use deterministic estimates between the true and approximate solution operator processes. The results are analogous to the Gaussian case: for the heat equation (with $p = 1$, $r = 2$, $\beta \in (0, 1]$) Theorem 3.3.4. shows the weak rate to be 2β and β, for Volterra equations (with $r = 2$, $\beta \in (0, 1/\rho]$) Theorem 3.4.1. shows the weak rate to be 2β and $\rho\beta$, and for the wave equation (with general parameters) Theorem 3.5.3. shows the weak rate to be $\min(2\beta\frac{r}{r+1}, r)$ and $\min(2\beta\frac{p}{p+1}, 1)$, spatially and temporally, respectively.

Chapter 2 moves away from the linear setting. The scope of the semilinear term f includes relevant examples like Nemytskii operators $X \to F \circ X$ with sufficiently regular F. It however excludes nonlinearities with more than linear growth or depending also on derivatives - for example in the true CHC equation the nonlinearity Δu^3 violates both

conditions. The class of functionals is quite a bit larger than above: they are assumed to be of the form

$$Z \mapsto \prod_{i=1}^{k} \varphi_i\big(\int_0^T Z(t)\, d\nu_i(t)\big),$$

with some sufficiently regular functions φ_i and finite Borel measures ν_i. The strategy is also different: loosely speaking it makes use of the fact that very rough estimates like

$$\begin{aligned}\mathbb{E}\big(g(X(T)) - g(X^n(T))\big) &= \mathbb{E}\Big(\int_0^1 g'(X(T) + \lambda(X^n(T) - X(T)))\, d\lambda\Big)(X^n(T) - X(T)) \\ &\lesssim \sup_x |g'(x)| \mathbb{E}|X(T) - X^n(T)|\end{aligned}$$

can be much improved via Malliavin calculus, using instead appropriate dual Malliavin norms in the final inequality. Denoting by $\mathbf{M}^{1,p,q}$ the space of L^p_ω random variables with $L^p_\omega L^q_t$ Malliavin derivative and by $(\mathbf{M}^{1,p,q})^*$ its dual in the Gelfand triple $\mathbf{M}^{1,p,q} \subset L^2_\omega \subset (\mathbf{M}^{1,p,q})^*$, Proposition 2.2.1. first establishes the uniform $\mathbf{M}^{1,p,q}$-regularity of the true solution and its approximation. Then Lemma 2.3.6. proves the desired rate of convergence for the strong error in the weak norm $(\mathbf{M}^{1,p,q})^*$. Putting the two together, the weak rate follows easily, this is formulated in the main result of the chapter Theorem 2.3.7.

All of this is set up in an general abstract framework, with the assumptions on the equation and its approximations formulated in Sections 2.1. and 2.3.1. The chapter concludes with Section 2.4., where the abstract assumptions are briefly verified for two concrete examples.

While the topic of numerical analysis is typically motivated with computational considerations in mind, the thesis does not detail applications or simulations, and therefore I view it as a primarily theoretical one. Nevertheless, one natural application-related question that arises:

Q1. What is the practical motivation behind studying weak convergence for infinite-dimensional systems? What is the typical φ that appears?

The error representation Theorem 1.2.1. is set up in an elegant abstract way to simultaneously handle different leading order operators, as illustrated with parabolic, hyperbolic, as well as some non-Markovian examples. On the other hand, linear equations with additive noise are rather special in that an exact solution formula (2) is available, which seems to be used in a crucial way, so the robustness of the methods is not clear. The author himself notes in the introduction that the method of Chapters 1 and 3 no longer work when $f \neq 0$, citing the issue of Markovianity for Volterra equations.

Q2. Let us consider only, say, parabolic equations, where Markovianity is not an problem. Is there an example where the bounds and/or the strategy in the linear additive case can be used in either a nonlinear or a multiplicative (w.r.t. the noise) equation?

Q3. When applying the error representation formula, the bounds used on the derivatives of u are the ones that immediately follow from differentiating (1.1.21) (or (3.2.4)). Are there more sophisticated regularity results available for the Kolmogorov equation considered here that could relax certain assumptions? For instance could analogues of the finite dimensional bounds

$$\partial_t u + \Delta u = 0, \quad u(T,x) = g(x) \qquad \Rightarrow \qquad |u_{xx}(T-t,x)| \lesssim t^{-1/2} \sup_y |g'(y)|$$

relax the condition $g \in C^2$?

In Chapter 2 the Malliavin regularity of the solution is established in the $\mathbf{M}^{1,p,q}$ spaces. This is sort of 'orthogonal' to the question of spatial regularity, which is made use of in the other chapters.

Q4. The variation of constants formula (2.1.2) together with the smoothing effect (2.1.1) readily implies some Sobolev regularity of the solution. Could this spatial regularity be used to allow f to lose regularity, i.e. map from H to $H^{-\delta'}$ for some $\delta' > 0$ (perhaps different from δ)?

Q5. How important is the role of boundary conditions? In some examples Dirichlet, in others Neumann is chosen, is there a particular reason for these choices?

The results of the dissertation are novel and due to the author. Upon a successful defense I recommend awarding the title Doctor of the Hungarian Academy of Sciences.

Máté Gerencsér, Ph.D.
16 April 2019, Klosterneuburg