Multilevel Monte Carlo approximations of statistical solutions to the Navier-Stokes equations

A. Barth and Ch. Schwab and J. Sukys

MULTILEVEL MONTE CARLO APPROXIMATIONS OF STATISTICAL SOLUTIONS TO THE NAVIER–STOKES EQUATIONS

ANDREA BARTH, CHRISTOPH SCHWAB, AND JONAS ŠUKYS

ABSTRACT. We present Monte Carlo and multilevel Monte Carlo discretizations for the numerical approximation of the statistical solution to the viscous, incompressible Navier–Stokes equations in a bounded domain $D \subset \mathbb{R}^d$. We prove that Monte Carlo sampling produces sequences of sample averages of (Leray-Hopf) solutions to the Navier–Stokes equations which converge to a (generalized) moment of a (in two space dimensions unique) statistical solution (in the sense of Foiaş and Prodi), at the rate $M^{-1/2}$ in terms of the number of samples $M \in \mathbb{N}$. The Monte Carlo convergence rate $M^{-1/2}$ is shown to hold independently of the Reynolds number, with constant depending only on the mean kinetic energy of the initial velocity.

We discuss the effect of a space-time discretization on the Monte Carlo convergence with particular attention on the kinematic viscosity ν. For a multilevel Monte Carlo estimator, composed of ensembles of solutions with finite mean kinetic energy in $L^2(D)$, we give sufficient conditions for mean-square convergence to a (generalized) moment of the statistical solution. It is concluded that robust convergence rates are possible for multilevel Monte Carlo sample averages *provided* that the numerical solutions on the finest discretization are resolved or approximated with proper upscaling giving mean-square consistent bulk properties.

1. INTRODUCTION

In the bounded domain $D \subset \mathbb{R}^d$, for $d = 1, 2, 3$, with boundary ∂D and in the finite time interval $\bar{J} := [0, T]$, for $T < \infty$, we consider a viscous, incompressible flow subject to a prescribed divergence-free initial velocity field $u_0 : D \mapsto \mathbb{R}$ and subject to a body force f acting on the fluid particles in D. We assume that D is connected, and, in dimension $d = 1$ in addition, we assume that $D = (0, 1)$ (which could be achieved by scaling). The Navier–Stokes equations for viscous, incompressible flow of a Newtonian fluid are given in terms of the velocity field $u : \bar{J} \times D \mapsto \mathbb{R}^d$, and the pressure $p : \bar{J} \times D \mapsto \mathbb{R}$. The pressure takes the role of a Lagrange multiplier, enforcing the divergence-free constraint. They read, for $d \geq 2$ in $\bar{J} \times D$ (see, e.g., [35]),

$$\frac{\partial}{\partial t} u - \nu \Delta u + (u \cdot \nabla)u + \nabla p = f,$$
$$\nabla \cdot u = 0, \tag{1.1}$$

with the kinematic viscosity $\nu > 0$ and with a given initial velocity field $u(0) = u_0$.

We focus our exposition on Equation (1.1) equipped with either *no-slip* boundary conditions[1] or *periodic* boundary conditions. No-slip boundary conditions take the form

$$u = 0 \text{ on } \partial D. \tag{1.2}$$

When dealing with periodic boundary conditions of length $\mathcal{I} > 0$, in which case we denote the physical domain by $D = (0, \mathcal{I})^d$, we are interested in $\mathcal{I}$-*periodic* solutions which satisfy

$$u(t, x) = u(t, x + \mathcal{I}e_i), \tag{1.3}$$

for $i = 1, \ldots, d$ and $x \in D$. Here, $(e_i, i = 1, \ldots, d)$ denotes the set of unit vectors in the direction of the coordinate axes. Naturally, different periods $\mathcal{I}_i$ may be considered in the coordinate directions, all following results will hold in this case as well. To simplify notation and unless explicitly stated otherwise, we shall assume $\mathcal{I} = 1$.

For the sake of numerical analysis and for facilitating some of the numerical experiments in the second half of the present paper, we shall also consider Equation (1.1) in space dimension $d = 1$. In this case, we mean by Navier–Stokes equations the viscous Burgers' equation, with viscosity ν (which, in this particular case, of course takes on the meaning of a diffusion coefficient rather than the reciprocal Reynolds number). Then, the velocity field is scalar and Equation (1.1) simplifies to the viscous Burgers' equation, given by

$$\frac{\partial}{\partial t} u + \frac{1}{2} \frac{\partial}{\partial x} (u^2) = \nu \frac{\partial^2}{\partial x^2} u + f, \tag{1.4}$$

with initial condition $u(0) = u_0$. Here the incompressibility constraint $\nabla \cdot u = 0$ becomes void and the pressure variable (the Lagrange multiplier for the constraint) is discarded from the formulation. In particular, for Equation (1.4) the vanishing viscosity limit $\nu = 0$ of solutions is well-understood, and we refer for details to [24, Appendix B] and the references therein.

In this paper we consider numerical methods for generalized moments of statistical solutions to the Navier–Stokes equations, given in Equation (1.1) and, in space dimension $d = 1$, the Burgers' equation, given in Equation (1.4). The initial condition is then given by a probability measure and the solution is, therefore, at every time point also a measure. There are several, nonequivalent, mathematical definitions of a statistical solution for Equation (1.1). One way to define a statistical solution is, therefore, as one-parameter family of probability measures on the space of all velocity fields. This historically first definition is due to C. Foiaş and G. Prodi (see [12, 15]). This construction of statistical solutions is based on weak solutions of the Navier–Stokes equations in the sense of Leray–Hopf. Foiaş' statistical solution takes the form of a *one-parametric family of Borel probability measures* $(\mu_t, t \in [0, T])$ describing the distribution of the velocity fields of the flow at all times. The second approach to statistical solutions is due to M. I. Višik and A. V. Foursikov (see [37, 36]). Their construction takes mathematically the form of *one single ("Višik–Foursikov") measure on the space of flow trajectories*. Recently, in [16], a modification of the original Višik–Foursikov construction has been proposed to yield a Višik-Foursikov measure supported by the set of all Leray–Hopf weak solutions which is, moreover, under mild additional conditions, consistent with the earlier notion of statistical solution from C. Foiaş in [12]. The case of periodic boundary conditions is here of particular interest as it has been proved in [37, 18, 8] that any *homogeneous, statistical solution* of the Navier–Stokes equations is a limit (in the weak sense)

---

[1]Our results will hold with minor modifications also for the slip boundary condition $u \cdot n|_{\partial D} = 0$, where n denotes the unit outward normal to the Lipschitz boundary ∂D.

of homogeneous, statistical solutions which are concentrated on flow fields which are periodic in the space variables.

We adhere in our presentation closely to the notation and results in [14] and [16].

The paper is organized as follows: In Section 2 we introduce basic definitions and properties of probability measures on function spaces. In Section 3, we sum up the concept of statistical solutions of the Navier–Stokes equations. In Section 4, we introduce the Monte Carlo method for the approximation of (generalized) moments of Hilbert-space-valued random variables and establish a convergence estimate. We introduce space and time discretizations in Section 5. In Section 6, we investigate the effect of space and time discretization on the approximations, leading to a singlelevel Monte Carlo estimator. We then proceed to the analysis of a multilevel Monte Carlo discretization for the efficient approximation of statistical averages. As in our previous work on the multilevel Monte Carlo approximation of random PDEs (e.g. [4, 6, 7, 29, 30, 34] and the references there), we prove that a judicious combination of space and time discretization with Monte Carlo sampling, where the sample numbers depend on the discretization level and the Reynolds number, allow to approximate (generalized) moments of statistical solutions with work versus accuracy which, asymptotically, equals that for numerically computing one Leray–Hopf solution on the finest mesh. We give a convergence result for the approximation of the, possibly infinite dimensional, initial distribution in Section 7. The concluding Section 8 presents numerical experiments for a model problem, confirming the theoretical results of this manuscript.

2. MEASURES ON FUNCTION SPACES

As statistical solutions are by definition families of probability measures on the space of all flows with finite mean kinetic energy, and as our algorithms exploit the properties of such measures in an essential fashion, we recapitulate, for the convenience of the reader, basic definitions of such measures, from [11, Chap.1].

Let $(\Omega, \mathcal{F})$ be a measurable space. If $(E, \mathcal{G})$ denotes a second measurable space, then an E-valued random variable (or random variable taking values in E) is a mapping $X : \Omega \to E$ such that the set $\{\omega \in \Omega \colon X(\omega) \in A\} = \{X \in A\} \in \mathcal{F}$ for any $A \in \mathcal{G}$, i.e. such that X is a $\mathcal{G}$-measurable mapping from Ω into E.

Assume now that E is a metric space equipped with the Borel σ-algebra $\mathcal{B}(E)$, then $(E, \mathcal{B}(E))$ is a measurable space and we shall always assume that E-valued random variables X are $\mathcal{F}$-$\mathcal{B}(E)$-measurable. If E is a separable Banach space with norm $\| \cdot \|_E$ and (topological) dual E^*, then $\mathcal{B}(E)$ is the smallest σ-field of subsets of E containing all sets

$$\{x \in E : \varphi(x) \leq \alpha\},$$

for $\varphi \in E^*$ and $\alpha \in \mathbb{R}$. Hence, if E is a separable Banach space, then X is an E-valued random variable if and only if for every $\varphi \in E^*$, the mapping $\omega \mapsto \varphi(X(\omega))$ is an $\mathbb{R}$-valued random variable. Moreover, we have the following lemma (see [11]).

Lemma 2.1. *Let E be a separable Banach space and let X be an E-valued random variable on $(\Omega, \mathcal{F})$. Then the mapping $\omega \mapsto \|X(\omega)\|_E$ is measurable.*

The random variable X is called *Bochner integrable* if, for any probability measure $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$, we have

$$\int_\Omega \|X\|_E \, d\mathbb{P} < +\infty.$$

Here, a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a σ-additive set function from Ω into $[0, 1]$ such that $\mathbb{P}(\Omega) = 1$, and the resulting measure space $(\Omega, \mathcal{F}, \mathbb{P})$ is a *probability space*. We shall always assume, unless explicitly stated, that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, i.e. for every null-set $B \in \mathcal{F}$, every subset $A \subset B$ is measurable, meaning $A \in \mathcal{F}$.

We shall require for $1 \leq p \leq \infty$ Bochner spaces of p-summable random variables X taking values in the Banach space E. By $L^1(\Omega; E)$ we denote the set of all (equivalence classes of) integrable, E-valued random variables X and we equip it with the norm

$$\|X\|_{L^1(\Omega;E)} = \int_\Omega \|X\|_E \, d\mathbb{P} = \mathbb{E}(\|X\|_E).$$

More generally, for $1 \leq p < \infty$, we define $L^p((\Omega, \mathbb{P}); E)$ as the set of p-summable random variables (defined with respect to the measure $\mathbb{P}$) taking values in E and equip it with norm

$$\|X\|_{L^p((\Omega,\mathbb{P});E)} := \left(\mathbb{E}_\mathbb{P}(\|X\|_E^p) \right)^{1/p}.$$

Here, $\mathbb{E}_\mathbb{P}$ denotes the expectation defined with respect to the probability measure $\mathbb{P}$. For $p = \infty$, we denote by $L^\infty(\Omega; E)$ the set of all E-valued random variables which are essentially bounded. This set is a Banach space equipped with the norm

$$\|X\|_{L^\infty(\Omega;E)} := \operatorname*{ess\,sup}_{\omega \in \Omega} \|X(\omega)\|_E.$$

Likewise we define, for $1 \leq p < \infty$ (with the obvious modification for $p = \infty$) and for $T < +\infty$ the Bochner space $L^p(\bar{J}; E)$ of strongly measurable functions taking values in E, endowed with the norm

$$\|f\|_{L^p(\bar{J};E)} := \begin{cases} \left(\int_0^T \|f(t)\|_E^p \, dt \right)^{1/p}, & \text{for } 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{t \in (0,T)} \|f(t)\|_E, & \text{for } p = \infty. \end{cases}$$

For any separable Banach space E, and for any $r \geq p \geq 1$, we have

$$L^r(J; E), \ C^0(\bar{J}; E) \in \mathcal{B}\big(L^p(J; E)\big).$$

3. Statistical solutions

We recapitulate the concept and definition of a statistical solution of the Navier–Stokes equations in the sense of C. Foiaş and R. Prodi as introduced in [12, 15]. To this end, following [14], we first review the variational formulation of the Navier–Stokes equations in space, and its interpretation as dynamical system for $u(t) \in V$. Here, V is an appropriate function space of divergence-free velocity fields which is made precise later in this section. We review known existence and uniqueness results for weak solutions in the sense of Leray–Hopf. To distinguish these from statistical solutions, we shall call them subsequently "individual weak solutions". We then turn to the definition of a statistical solution for the Navier–Stokes equations as introduced by C. Foiaş in [12] and by C. Foiaş and G. Prodi in [15]; we shall refer to these statistical solutions as Foiaş–Prodi statistical solutions, for conciseness. Existence proofs in [12, 15] for Foiaş–Prodi statistical solutions were based on the classical "Faedo–Galerkin" approach due to J. Leray and H. Hopf. Foiaş–Prodi statistical solutions take the form of a one-parameter family of probability measures (the parameter being the time variable) on the space of divergence-free velocity fields in D.

A different notion of statistical solution was introduced by M. I. Višik and A. V. Foursikov in [37, 36]. We shall refer to these statistical solutions as Višik–Foursikov statistical solutions. Višik–Foursikov statistical solutions take the form of probability measures supported on sets

of *velocity trajectories generated by the family of individual weak solutions in the sense of Leray–Hopf.*

In [16], the relation between the Foiaş–Prodi and Višik–Foursikov statistical solutions was made more precise: Under slightly stronger assumptions than those imposed by M. I. Višik and A. V. Foursikov in [37, 36], it was shown that statistical solutions in the sense of Višik–Foursikov are, in fact, statistical solutions in the sense Foiaş–Prodi. More specifically, the 'time-t projection' of the measure constituting the Višik–Foursikov statistical solution is a Foiaş–Prodi statistical solution (see [16, Thm. 3.4]).

3.1. Function spaces.

In a bounded, connected domain $D \subset \mathbb{R}^d$, for $d = 2, 3$, we consider the incompressible, instationary Navier–Stokes equations, given in Equation (1.1), with either no–slip or periodic boundary conditions (see Equation (1.2) and Equation (1.3)).

The domain D is assumed to be open, bounded and connected and its boundary ∂D is assumed to be either C^2 or D is assumed to be convex, in order to ensure local $H^2(D)$ regularity of the velocity field (we refer to [28] and the references there for details on spatial regularity of velocity and pressure field in polyhedral domains).

We denote by H a subspace of divergence-free vector fields in $L^2(D)^d$ and by V a subspace of the closure of H in $H^1(D)^d$. For no-slip boundary conditions in space dimensions $d = 2, 3$

$$H = \{v \in L^2(D)^d : \nabla \cdot v = 0 \text{ in } H^{-1}(D), \ v \cdot n|_{\partial D} = 0 \text{ in } H^{-1/2}(\partial D)\},$$

which implies that $H \subset L^2(D)^d$ is a closed, linear subspace of $L^2(D)$. Here, n denotes the outward unit normal to the domain D which is defined almost everywhere on the Lipschitz boundary ∂D. The space V is defined for no-slip boundary conditions as

$$V = V_{dir} := \{v \in H^1(D)^d : \nabla \cdot v = 0 \text{ in } L^2(D), \ v|_{\partial D} = 0\}. \tag{3.1}$$

For periodic boundary conditions, with $D =]0, I]^d$ we choose

$$V = V_{per} := \{v \in H^1_{per}(D)^d : \nabla \cdot v = 0 \text{ in } L^2(D), \int_D v \, dx = 0\}, \tag{3.2}$$

$$H = \{v \in L^2(D)^d : \int_D v \, dx = 0\} \ .$$

Evidently, we have the dense inclusions $V \subset H$ and, throughout the following, we identify the Hilbert space H with its own dual, i.e., $H \simeq H^*$. On the spaces H and V we have the (canonical) inner products

$$(v, w)_H = \int_D v \cdot w \, d\boldsymbol{x} \quad \text{and} \quad (v, w)_V = \int_D \sum_{i=1}^d \frac{\partial v}{\partial x_i} \cdot \frac{\partial w}{\partial x_i} \, d\boldsymbol{x},$$

where $\boldsymbol{x} = (x_1, \dots, x_d)$ denotes a point in D, with associated norms

$$\|v\|_H = ((v, v)_H)^{1/2}, \text{ for } v \in H, \quad \|v\|_V = ((v, v)_V)^{1/2}, \text{ for } v \in V.$$

In the following, we shall use the symbols V and H in all statements which apply generically, i.e. to either choice of V and of H. For any subset $X \subseteq H$, we denote by X_w this subset endowed with the weak topology of H. For $X = H$, then, H_w denotes the space H endowed with its weak topology. For $R > 0$, we denote the closed ball of radius R in H by $B_H(R)$. Since H is separable, $B_H(R)_w$ is a complete, metrizable space.

3.2. **Leray–Hopf weak solutions.** With the function spaces just defined, the *weak formulation* of Equation (1.1) reads: given $T > 0$, $u_0 \in H$ and $f \in L^2(J; H)$, find $u \in L^\infty(J; H) \cap L^2(J; V)$, such that, for all $v \in V$

$$\frac{d}{dt}(u, v)_H + \nu(u, v)_V + b(u, u, v) = (f, v)_H. \tag{3.3}$$

Since our formulation is divergence-free there is no pressure term entering in the weak formulation. In Equation (3.3), the trilinear form b is defined by

$$b(u, v, w) = \sum_{i,j=1}^{d} \int_D u_i \frac{\partial v_j}{\partial x_i} w_j \, d\boldsymbol{x}.$$

The trilinear form b is continuous on V and

$$\forall u, v, w \in V: \quad b(u, v, v) = 0 \quad \text{and} \quad b(u, v, w) = -b(u, w, v) .$$

Alongside with the weak formulation we can achieve an equivalent formulation involving the Stokes operator

$$Au = -P\Delta u, \tag{3.4}$$

for all $u \in V \cap H^2(D)^d$. Here, P denotes the Leray projection onto H in $L^2(D)^d$. Defined like this, the Stokes operator A is a positive, self-adjoint operator and therefore we can define fractional powers of A. We denote the fractional powers by A^a, for $a \in \mathbb{R}$, and by $\mathcal{D}(A^a)$ the domain of A^a. We have $\mathcal{D}(A^{1/2}) = V$. Further, the trilinear form b induces, for fixed $u \in V$, a bilinear operator $B : \mathcal{D}(A^{1/2}) \times \mathcal{D}(A^{1/2}) \to \mathcal{D}(A^{-1/2})$ defined by

$$_{\mathcal{D}(A^{-1/2})}\langle B(u, v), w \rangle_{\mathcal{D}(A^{1/2})} = b(u, v, w),$$

for all $u, v, w \in \mathcal{D}(A^{1/2})$. This in hand, we introduce the functional formulation of the Navier–Stokes equations (1.1): given $T > 0$, $u_0 \in H$ and $f \in L^2(J; H)$, find $u \in L^\infty(J; H) \cap L^2(J; V)$ with $u' \in L^1(J; \mathcal{D}(A^{-1/2}))$ such that

$$u' + \nu Au + B(u, u) = f . \tag{3.5}$$

Then, we have the following definition.

Definition 3.1. On a time interval $J \subset \mathbb{R}$, a function $u : J \mapsto H$ is called a *Leray–Hopf weak solution* of Equation (1.1) if

  (i) $u \in L^\infty_{loc}(J; H) \cap L^2_{loc}(J; V)$,
  (ii) $(\partial_t u)(\cdot) \in L^{4/3}_{loc}(J; V')$,
  (iii) $t \mapsto u(t) \in \mathcal{C}_{loc}(J; H_w)$ (i.e. for every $v \in H$, $t \mapsto (u(t), v)_H$ is continuous from J to $\mathbb{R}$),
  (iv) u satisfies Equation (3.5) in the sense of distributions on J with values in V',
  (v) for almost all $t, t' \in J$, u satisfies the energy inequality

$$\frac{1}{2}\|u(t)\|_H^2 + \nu \int_{t'}^t \|u(s)\|_V^2 \, ds \leq \frac{1}{2}\|u(t')\|_H^2 + \int_{t'}^t (f(s), u(s))_H \, ds. \tag{3.6}$$

If further $J = [t_0, t_1]$ is closed and left-bounded then

  (vi) $u(t)$ is strongly right-continuous in H at t_0, i.e. $u_0 = \lim_{t \downarrow t_0} u(t)$ in H.

Under (v), Equation (3.6) holds for all time instances $t' \in [t_0, t_1]$ where the Leray–Hopf solution u is strongly (in H) right continuous. We note that the set of such points is of full measure in J. We may rewrite Equation (3.5) as nonlinear dynamical system

$$u'(t) = F(t, u(t)), \tag{3.7}$$

where the right hand side is given by $F(t, u) = f - \nu A u - B(u, u)$.

We call Leray–Hopf solutions from now on simply *weak solutions* of the Navier–Stokes equations. It is by now classical that for any $t_0 \in \mathbb{R}$ and for any $u_0 \in H$, there exists at least one global weak solution in $[t_0, \infty)$ such that $u(t_0) = u_0$ in H. In space dimension $d = 2$, this solution is, moreover, unique (see [35] and the references there). We denote by $S(t, 0)$ the solution operator that maps u_0 into $u(t)$. The solution operator is well defined in space dimension $d = 2$ thanks to the uniqueness of weak solutions (see, e.g. [35, Thm. 3.1] and the references there). It is, however, in general not a semigroup on H, since f could be time dependent. In space dimension $d = 3$, the definition of the solution operator is more involved, since in the presence of a time-dependent forcing function f, only local uniqueness has been shown to date [35, Thm. 3.2 ii)], albeit only for more regular initial conditions, such as $u_0 \in V$. However, additional regularity is required in any case for convergence rates estimates for the discretization methods to be discussed ahead.

3.3. A-priori estimates for Leray–Hopf solutions.
Here, we collect classical a-priori estimates on Leray–Hopf solutions. The Navier–Stokes equations is, at positive kinematic viscosity ν, formally a parabolic problem. Hence, various types of a-priori estimates and regularity results are available. As a rule, most available regularity estimates establish a-priori bounds on the solution in terms of Sobolev or Besov spaces without explicit dependence of constants on ν. The estimates which we require in our analysis are classical and elementary relative to recent developments. However, the dependence of constants on ν is explicit in these estimates.

First, for Leray–Hopf solutions u of Equation (1.1) with no-slip boundary conditions, we multiply with the testfunction $u(t)$ and integrate by parts, to obtain (compare Equation (3.6))

$$\frac{d}{dt}\|u(t)\|_H^2 + \nu\|u(t)\|_V^2 \leq \frac{1}{\nu}\|f(t)\|_{V^*}^2 \ .$$

Further, for $u_0 \in H$, this gives

$$\int_0^T \|u(t)\|_V^2 \, dt \leq K,$$

where K is defined by $K := \nu^{-1}\left(\|u_0\|_H^2 + \frac{1}{\nu}\int_0^T \|f(t)\|_{V^*}^2 \, dt\right)$. For $0 < s < T$, we may write

$$\|u(s)\|_H^2 \leq \|u_0\|_H^2 + \frac{1}{\nu}\int_0^s \|f(t)\|_{V^*}^2 \, dt,$$

so that

$$\sup_{s \in [0,T]} \|u(s)\|_H^2 \leq \nu K = \|u_0\|_H^2 + \frac{1}{\nu}\int_0^T \|f(t)\|_{V^*}^2 \, dt \ .$$

This bound is, in particular, independent of ν if $f \equiv 0$, i.e. if the flow is not externally forced.

Estimates in the V-norm (and in stronger norms) of $u(t)$ become poor as $\nu \to 0$. One exception from this are space dimensions $d = 1, 2$, and periodic boundary conditions. Here

(see, e.g., [35, Chapter 3.1, iii)]) holds the following energy equality

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_V^2 + \nu\|Au(t)\|_H^2 = (f(t), Au(t))_H,$$

which implies (recall the definition of V in Equation (3.1) and (3.2) in the periodic setting in space dimension $d = 2$)

$$\frac{d}{dt}\|u(t)\|_V^2 + \nu\|Au(t)\|_H^2 \leq \frac{1}{\nu}\|f(t)\|_H^2.$$

Further, we get

$$\sup_{0 \leq t \leq T}\|u(t)\|_V^2 \leq \|u_0\|_V^2 + \frac{1}{\nu}\|f(t)\|_H^2. \tag{3.8}$$

Equation (3.8) implies a ν-independent bound for $\|u\|_{L^\infty(J;V)}$ with V given in Equation (3.2) for $f \equiv 0$.

3.4. **Statistical solutions.** Contrary to the dynamical systems viewpoint of the Navier–Stokes equations, which associates a (unique in dimension $d = 2$) weak solution $u(t)$ to a given initial condition $u_0 \in H$, statistical solutions aim at desribing the evolution of *ensembles of solutions* through their probability distribution.

Specifically, we assume that we are given a probability measure on H, which is equipped with the Borel-σ-algebra $\mathcal{B}(H)$. Then, a statistical solution is a (parametric family of) probability measure(s). Rather than being restricted to one single initial condition, a Foiaş–Prodi statistical solution of the Navier–Stokes equations is a one-parameter family of probability measures which describes the evolution of velocity ensembles and their statistics. We remark that the initial-boundary value problem in Equation (3.3) can be interpreted as a special case of a statistical solution: In this case, the measure μ_0 places unit mass at one initial velocity $u_0 \in H$. In general, the initial distribution is defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and is assumed to be given as an image measure under an H-valued random variable with distribution μ_0. This random variable is defined as a mapping from the measurable space $(\Omega, \mathcal{F})$ into the measurable space $(H, \mathcal{B}(H))$ such that $\mu_0 = X \circ \mathbb{P}$. As a consequence we have a time-dependent family of measures $\mu = (\mu_t, t \geq 0)$ on H given by

$$\mu_t(E) = \mu_0(S(t,0)^{-1}E),$$

for all measurable ensembles of initial velocities $E \subset H$, i.e., for all $E \in \mathcal{B}(H)$. In other words, the probability $\mu_t(E)$ for $u(t)$ to be in $E \in \mathcal{B}(H)$ is the same as for u_0 to be in $S(t,0)^{-1}E$, i.e., $\mu_0(S(t,0)^{-1}E)$. For any time $t \geq 0$, we then define the generalized moment

$$\int_H \Phi(v)\,d\mu_t(v)$$

for a μ_t-integrable function Φ on H. The evolution of statistical moments of the flow in time may be formally derived then as

$$\frac{d}{dt}\int_H \Phi(v)\,d\mu_t(v) = \frac{d}{dt}\int_H \Phi(S(t,0)v)\,d\mu_0(v)$$

$$= \int_H (F(t, S(t,0)v), \Phi'(S(t,0)v))_H\,d\mu_0(v)$$

$$= \int_H (F(t,v), \Phi'(v))_H\,d\mu_t(v),$$

for suitable testfunctionals Φ. The expression

$$\frac{d}{dt} \int_H \Phi(v) \, d\mu_t(v) = \int_H (F(t,v), \Phi'(v))_H \, d\mu_t(v) \tag{3.9}$$

with F as in Equation (3.7) is meaningful even if the solution operator is not defined, as in the general case for $d = 3$. A suitable class of testfunctionals is given by the following definition.

Definition 3.2. Let $\mathcal{C}$ be the space of cylindrical test functionals Φ on H which are real-valued and depend only on a finite number of components of $v \in H$, i.e. for $k < \infty$

$$\Phi(v) = \phi((v, g_1)_H, \dots, (v, g_k)_H),$$

where ϕ is a compactly supported C^1 scalar function on $\mathbb{R}^k$ and $g_1, \dots, g_k \in V$.

For $\Phi \in \mathcal{C}$ we denote by Φ' its differential in H, which is given by

$$\Phi'(v) = \sum_{i=1}^{k} \partial_i \phi((v, g_1)_H, \dots, (v, g_k)_H) g_i \,.$$

As a linear combination of elements in V, $\Phi'(v)$ belongs to V.

Energy-type inequalities are central to the notion of statistical solutions of Equation (1.1). Suppose for now that the mapping

$$t \mapsto \int_V \|v\|_V^2 \, d\mu_t$$

is integrable on $J = (0, T)$ (resp. locally integrable on $(0, \infty)$). This implies that the family of measures $(\mu_t, t \in J)$ does not carry any mass on $H \backslash V$ and that, for any $\Phi \in \mathcal{C}$ the mapping

$$v \mapsto (F(t,v), \Phi'(v))_V$$

is continuous in V and that the right hand side of Equation (3.9) is well defined. Then, we may integrate Equation (3.9), which gives

$$\int_H \Phi(v) \, d\mu_t(v) = \int_H \Phi(v) \, d\mu_0(v) + \int_0^t \int_H (F(s,v), \Phi'(v))_H \, d\mu_s(v) \, ds.$$

This leads to an energy-type inequality, given, for all $t \in [0, T]$, by

$$\int_H \|v\|_H^2 \, d\mu_t(v) + 2\nu \int_0^t \int_V \|v\|_V^2 \, d\mu_s(v) \, ds$$

$$\leq \int_0^t \int_H (f(s), v)_H \, d\mu_s(v) \, ds + \int_H \|v\|_H^2 \, d\mu_0(v). \tag{3.10}$$

We remark that for $d = 2$ we have equality in Equation (3.10) (see, e.g., [14, Eqn. (V.1.9)]). Equation (3.9) and Equation (3.10) motivate the definition of statistical solutions of (1.1).

Definition 3.3. A one-parameter family $\mu = (\mu_t, t \in J)$ of Borel probability measures on H is called *statistical solution* of Equation (1.1) on $J \subset \mathbb{R}$ if

(i) the initial Borel probability measure μ_0 on H has *finite mean kinetic energy*, i.e.,

$$\int_H \|v\|_H^2 \, d\mu_0(v) < \infty,$$

(ii) $f \in L^2(J; H)$ and the Borel probability measures μ_t satisfy Equation (3.9) for all $\Phi \in \mathcal{C}$ and Equation (3.10) holds,

(iii) the mapping

$$J \ni t \mapsto \int_H \varphi(v) \, d\mu_t(v)$$

is measurable on J for every bounded, continuous, real-valued function $\varphi : H \mapsto \mathbb{R}$ and the Borel probability measures $(\mu_t, t \in J)$ satisfies (compare [14, (V.1.12), (V.1.13)])

$$t \mapsto \int_V \|v\|_V^2 \, d\mu_t(v) \in L^1(J), \quad t \mapsto \int_H \|v\|_H^2 \, d\mu_t(v) \in L^\infty(J)$$

(iv) (Liouville Equation) for every cylindrical testfunction Φ as in Definition 3.2, and for every $t, t' \in J$, μ_t satisfies

$$\int_H \Phi(v) \, d\mu_t(v) = \int_H \Phi(v) \, d\mu_{t'}(v)$$
$$+ \int_{t'}^t \int_H (f, \Phi'(v)) - \nu(Av, \Phi'(v)) - (B(v,v), \Phi'(v)) \, d\mu_s(v) \, ds.$$

(v) (strengthened mean energy inequality (3.10)) on the time interval $J \subset \mathbb{R}$ there exists a subset $J' \subset J$ of full measure such that, for every nonnegative continuously differentiable function $\psi : [0, \infty) \to \mathbb{R}$ with $\|\psi'\|_{L^\infty((0,\infty))} < \infty$, there holds

$$\frac{1}{2} \int_H \psi(\|u\|_H^2) \, d\mu_t(u) + \nu \int_{t'}^t \int_H \psi'(\|u\|_H^2) \|u\|_V^2 \, d\mu_s(u) \, ds$$
$$\leq \frac{1}{2} \int_H \psi(\|u\|_H^2) \, d\mu_{t'}(u) + \int_{t'}^t \int_H \psi'(\|u(s)\|_H^2)(f(s), u(s))_H \, d\mu_s(u) \, ds$$

for every $t' \in J'$ and every $t \in J$ with $t' < t$.

Having the definition of the statistical solution in place, we state the following result (see [12, Thm. 1 and Prop.1] and [16]).

Theorem 3.4. *Let μ_0 be a Borel probability measure on H with finite mean kinetic energy,*

$$\int_H \|v\|_H^2 \, d\mu_0(v) < +\infty \ .$$

Let, moreover, $f \in L^2(J; H)$ be a forcing term. Then, for either the no-slip case (see Equation (3.1)) or the periodic case (see Equation (3.2)), there exists a statistical solution $(\mu_t, t \in J)$ of the Navier–Stokes equations on H in the sense of Definition 3.3.

In dimension $d = 2$, if μ_0 is supported in $B_H(R)$ for some $0 < R < \infty$, and if the forcing term $f \in H$ is time-independent, the statistical solution is unique and explicitly given by $\mu_t = S(t,0)\mu_0$, for $t \geq t_0$, i.e. by μ_0 transported under the flow $(S(t,0), t \in J)$ of (unique in space dimension $d = 2$) Leray–Hopf solution operators of Equation (3.3).

The proof of the existence (and uniqueness in dimension $d = 2$) result can be found in [14, Thms. V.1.1, V.1.2] in the case of no-slip boundary conditions and in [14, Thm. V.1.3-V.1.5] in the case of periodic boundary conditions. We also remark that in space dimension $d = 3$, there is in general no uniqueness, and only a local energy inequality holds. We refer to [9], where the notion of *homogeneous* statistical solutions on all of $\mathbb{R}^3$ is extended to Navier–Stokes equations including the pressure variable.

4. Monte Carlo method

In this section we prove that, for every $M \in \mathbb{N}$, Monte Carlo sample averages built from a finite ensemble of M individual weak solutions in the sense of Leray–Hopf approximate generalized moments of a Foiaş–Prodi statistical solution with finite mean kinetic energy *in mean square sense* at rate $M^{-1/2}$ with a constant which is independent of the Reynolds number. This Monte Carlo convergence result assumes, however, availability of M *exact solutions in the sense of Leray–Hopf* in the Monte Carlo estimator. In practice, the M Leray-Hopf solutions which constitute the sample average need to be approximated numerically. This is addressed in Section 5. The Monte Carlo convergence analysis in the present section sets the stage for discretization error bounds which have to be satisfied by discretizations of individual weak solutions of the Navier–Stokes equations in order to achieve consistent approximations of statistical solutions.

4.1. **Monte Carlo method.** Our goal is the numerical approximation of (generalized) moments of a Foiaş–Prodi statistical solution $(\mu_t, t \in J)$ for a given initial distribution μ_0 on H. More precisely, we are interested in approximating, for given $\Phi \in \mathcal{C}$ (with $\mathcal{C}$ as in Definition 3.2) and for μ_0 with finite mean kinetic energy, we approximate

$$\mathbb{E}_{\mu_t}(\Phi) = \int_H \Phi(v)\, d\mu_t(v)\,, \quad t \in J\,.$$

For the first approach, we assume that we can sample from the exact initial distribution μ_0. Since μ_0 is a distribution on the infinite-dimensional space H, this is, in general, a simplifying assumption. However, if μ_0 is given by a finite-dimensional measure the assumption is no constraint. We discuss an appropriate approximation of the initial distribution in Section 7. We generate $M \in \mathbb{N}$ independent copies $(v^i, i = 1, \ldots, M)$ of u_0, where u_0 is μ_0-distributed. We assume further that for each sample, distributed according to μ_0, we can solve $u(t) = S(t,0)u_0$ exactly. Further, we suppose that we can evaluate the real-valued functional $\Phi(u(t))$ exactly. Then, we have the approximation

$$\mathbb{E}_{\mu_t}(\Phi) \approx E_{\mu_t}^M(\Phi) := \frac{1}{M} \sum_{i=1}^{M} \Phi(S(t,0)v^i), \tag{4.1}$$

where we denoted by $(E_{\mu_t}^M, M \in \mathbb{N})$ the sequence of Monte Carlo estimators which approximate the (generalized) expectation $\mathbb{E}_{\mu_t}(\Phi)$.

To state the error bound on the variance of the Monte Carlo estimator, given in Equation (4.1), we assume for simplicity that the right hand side of Equation (1.1) is equal to zero, i.e., $f \equiv 0$ (all results that follow have an analog for nonzero forcing). All test functions in $\mathcal{C}$ fulfill, for some constant $C > 0$, the *linear growth condition*

$$\forall v \in H: \quad |\Phi(v)| \leq C(1 + \|v\|_H). \tag{4.2}$$

Proposition 4.1. *Let $\Phi \in \mathcal{C}$ be a testfunction. Then, an error bound on the mean-square error of the Monte Carlo estimator $E_{\mu_t}^M$, for $M \in \mathbb{N}$, is given by*

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^M(\Phi)\|_{L^2((H,\mu_t);\mathbb{R})} = \frac{1}{\sqrt{M}}\left(\mathrm{Var}_{\mu_t}(\Phi)\right)^{1/2} \leq C\frac{1}{\sqrt{M}}\left(1 + \left(\int_H \|v\|_H^2\, d\mu_0(v)\right)^{1/2}\right).$$

For $\nu > 0$, the latter inequality is strict. Here we used the notation $\mathrm{Var}_{\mathbb{P}}(X) = \mathbb{E}_{\mathbb{P}}(\|\mathbb{E}_{\mathbb{P}}(X) - X\|_E^2)$ for a square-integrable E-valued random variable X.

Proof. We have, with the independence of the copies $(v^i, i = 1, \ldots, M)$ and the fact that they are identically distributed

$$\mathbb{E}_{\mu_t}\big(|\mathbb{E}_{\mu_t}(\Phi) - E^M_{\mu_t}(\Phi)|^2\big) = \mathbb{E}_{\mu_t}\Big(\frac{1}{M}\big|\int_H \Phi(v)\,d\mu_t(v) - \Phi(S(t,0)v)\big|^2\Big).$$

Now, we may write, by the linearity of the expectation,

$$\frac{1}{M}\mathbb{E}_{\mu_t}\Big(\big|\int_H \Phi(v)\,d\mu_t(v) - \Phi(S(t,0)v)\big|^2\Big) = \frac{1}{M}\,\mathrm{Var}_{\mu_t}(\Phi(v)).$$

Further we have with the linear growth condition in Equation (4.2)

$$\frac{1}{M}\,\mathrm{Var}_{\mu_t}(\Phi(v)) \leq \frac{1}{M}\|\Phi\|^2_{L^2((H,\mu_t);\mathbb{R})} \leq C\frac{1}{M}\big(1 + \int_H \|v\|^2_H\,d\mu_t(v)\big).$$

Equation (3.10) reads (under the assumption that $f \equiv 0$)

$$\int_H \|v\|^2_H\,d\mu_t(v) + 2\nu\int_0^t\int_V \|v\|^2_V\,d\mu_s(v)\,ds \leq \int_H \|v\|^2_H\,d\mu_0(v),$$

and since all terms are positive we have

$$\int_H \|v\|^2_H\,d\mu_t(v) \leq \int_H \|v\|^2_H\,d\mu_0(v),$$

which proves the assertion. $\qquad\square$

We remark that the error estimate in Proposition 4.1 does not contain any implicit constant. We therefore conclude that the (mean-square over all flow configurations) *convergence rate of Monte Carlo sample averages is uniform with respect to the physical parameters of the flow* but depends, of course, on the second moment of μ_0, i.e. on the mean kinetic energy of the initial probability measure μ_0.

4.2. **Convergence of Monte Carlo sample averages to Višik–Foursikov solutions.**
The significance of Monte Carlo sampling in the approximation of generalized moments of a statistical solution $(\mu_t, t \in J)$ is highlighted by the fact that for any finite number M of samples, the Monte Carlo estimator $E^M_{\mu_t}$ converges in distribution to the expectation of a Višik–Foursikov statistical solution.

Using the Krein–Milman argument from [14, 16], Monte Carlo sample averages, being convex combinations of Višik–Foursikov statistical solutions which are Dirac measures supported on individual Leray–Hopf solutions, can be expected to converge (in the sense of measures) as $M \to \infty$, to a Višik–Foursikov statistical solution, which is the (unique in space dimension $d = 2$) Foiaş–Prodi statistical solution. We refer to [17, Sec. 4.3] for details.

5. SPACE AND TIME DISCRETIZATION

The convergence bound in Proposition 4.1 indicates that numerical ensemble averaging over many flow configurations can yield mean-square convergent approximations to generalized moments of statistical solutions with convergence rates which are uniform with respect to the kinematic viscosity of the flow under the sole (physically meaningful) assumption that the ensemble of initial velocities has finite mean kinetic energy. The Monte Carlo error bounds in Proposition 4.1 are semi-discrete in the sense that they assume the availability of an exact Leray–Hopf solution of the Navier–Stokes equations for each initial velocity sample drawn from μ_0, and they pertain to bulk properties of the flow in the sense that they depend on the

H-norm of the individual flows. We have, therefore, to perform additional space and time discretizations in order to obtain computationally feasible approximations of (generalized) moments of statistical solutions. In this section, we address the effect of additional discretization errors incurred by space- and time-discretizations on the accuracy of the Monte Carlo sample averages.

For viscous, incompressible flows, there are by now numerous space and time discretization methods with convergence rate estimates available. We refer to [35, 23, 22, 21, 20] and the references there. While [35] contains the classical (due to Hopf and Leray) Faedo–Galerkin discretization and establishes convergence without any rate (which suffices to claim existence of Leray–Hopf solutions), for the multilevel Monte Carlo approach we require *error bounds with convergence rates* such as those provided in [23, 22, 21, 20]. As is customary in numerical analysis, these convergence rates depend on the smoothness s in scales of Sobolev or other function spaces of the Leray–Hopf solutions which arise as individual solutions in the Monte Carlo sampling. It is common in the discretization error analysis to state error bounds with constants which depend on the viscosity parameter $\nu > 0$ in an unspecified fashion. This precludes, in our ensuing multilevel Monte Carlo error analysis, to obtain convergence rates which are robust with respect to the Reynolds number. In order to do so, we work under the *hypothesis* that *robust convergence rates*, i.e. error bounds with constants independent of ν, require *scale resolving discretizations*, i.e. spatial resolution of the flow to the scale of viscous cut-off. All results which follow will hold under this assumption which is, however, prohibitive in most practical applications. In multilevel Monte Carlo sampling strategies such as those proposed in the next section, we consider a sequence of (space and time) discretizations which are indexed by a level index ℓ. We shall allow a *level-dependent sample number M_ℓ* to estimate (generalized) moments of statistical solutions. In this approach, the coarse level samples entail, inevitably, numerical simulations on scale-underresolving meshes, even in the (ideal) setting when the finest discretization level reaches the viscous cut-off. On the other hand, the convergence rate $M^{-1/2}$ shown in Proposition 4.1 is in "mean kinetic energy" of the solution ensemble. The analysis in the next section indicates that, as long as bulk properties and ensemble averages of the flows are of interest, scale-underresolution can be compensated to some extent with multilevel Monte Carlo, in the sense of mean-square kinetic energy, by sufficiently large sample numbers on coarse discretization levels. More precisely, the result shows that, as long as the finest level is scale-resolved (i.e. resolves the viscous cut-off), we can compensate spacial-temporal underresolution on coarser mesh levels by additional Monte Carlo sampling. Our multilevel error analysis indicates in particular that some form of turbulence modeling must be used in fully discrete multilevel Monte Carlo approximations of statistical solutions to generate numerical sample averages whose accuracy is uniform with respect to the Reynolds number. In this way the ν-independence of the error bounds in Proposition 4.1 is preserved under discretization.

5.1. **Space semi-discrete formulation.** For the space semi-discrete formulation, we consider a dense, nested family of finite dimensional subspaces $\mathcal{V} = (V_\ell, \ell \in \mathbb{N}_0)$ of V and therefore of H. Associated to the subspaces V_ℓ, we have the refinement levels $\ell \in \mathbb{N}_0$ and the refinement sizes $(h_\ell, \ell \in \mathbb{N}_0)$ and the H-orthogonal projections $(P_\ell, \ell \in \mathbb{N}_0)$. Furthermore, we endow the finite dimensional spaces in $\mathcal{V}$ with the norm induced by H. For $\ell \in \mathbb{N}_0$ the sequence is supposed to be dense in H in the sense that

$$\lim_{\ell \to +\infty} \|v - P_\ell v\|_H = 0.$$

Then the *space semi-discrete weak formulation* of the initial-boundary value problem Equation (1.1) reads: given $T > 0$, $u_0 \in H$ and $f \in L^2((0,T);H)$, find $u_\ell \in L^\infty((0,T);H) \cap L^2((0,T);V_\ell)$, such that, for all $v_\ell \in V_\ell$, there holds

$$\frac{d}{dt}(u_\ell, v_\ell)_H + \nu(u_\ell, v_\ell)_V + b(u_\ell, u_\ell, v_\ell) = (f, v_\ell)_H . \tag{5.1}$$

We assume that the spaces in $\mathcal{V}$ are chosen such that the following asymptotic discretization error bound holds.

Assumption 5.1. The sequence of semi-discrete solutions $u_\ell = (u_\ell(t), t \geq 0)$ converges to the solution $u = (u(t), t \geq 0)$ of Equation (1.1). Then the discretization error is bounded asymptotically, as $h_\ell \to 0$, for $t \in [0,T]$, by

(a)
$$\|u(t) - u_\ell(t)\|_H \leq Ch_\ell^s, \tag{5.2}$$
for $d = 1$ and for some $s \in [0,1)$;

(b)
$$\|u(t) - u_\ell(t)\|_H \leq C\frac{h_\ell^\sigma}{\nu}, \tag{5.3}$$
for $d \geq 2$ and for some $\sigma > 0$.

In both cases $C > 0$ is independent of ν, ℓ and h_ℓ.

Let us comment on Assumption 5.1. The convergence estimates (5.2) and (5.3) are explicit in the discretization parameter h_ℓ (equal to, for example, a meshwidth of Finite Volume mesh, or to N^{-1} where N denotes the spectral order of a spectral method) and in the kinematic viscosity ν. The assumption that Equation (5.3) holds with some $\sigma \geq 0$ implies *robust convergence*, i.e. a rate of convergence that holds for large Reynolds numbers. We emphasize that standard space-time discretizations of the Navier–Stokes equations in space dimension $d = 2, 3$, such as those in [35, 23, 22, 21, 20], will *not* satisfy Equation (5.3) with $\sigma > 0$, *unless some form of upscaling or turbulence modeling is included*. Establishing error bounds like Equation (5.3) with $\sigma > 0$ and where $C > 0$ depends polynomially on ν is in general an open problem.

In spatial dimension $d = 1$, it is shown in [25] that Equation (5.2) holds, with $s = 1/2$ and for some constant $C > 0$ independent of ν (strictly speaking, the bound is shown in $L^1(D)$, but the result in $H = L^2(D)$ can be derived with the $L^\infty(D)$-stability of the discretization scheme and with an application of Hölder's inequality). In space dimension $d \geq 2$, there is a rather substantial body of numerical analysis of space (and time) discretization schemes for the Navier–Stokes equations; we mention only [23, 22, 21, 20, 32, 33]. All error bounds, however, of the type of Equation (5.3) which are proved in these works have constants $C > 0$ which implicitly depend (exponentially) on T/ν and which are, therefore, not suitable to infer statements on the performance of the multilevel Monte Carlo approximation of statistical solutions for small values of ν.

For the case $d \geq 2$, in order to achieve asymptotic convergence under Assumption 5.1, it is required, that there exists $\ell^* \in \mathbb{N}_0$ such that $h_{\ell^*}^\sigma \leq \nu$. In this case we have convergence for all $\ell > \ell^*$. If we are in the regime $\ell < \ell^*$ (implying $h_\ell^\sigma > \nu$), then the convergence requirement is not fulfilled.

5.2. Fully-discrete formulation.
The approximation of the weak formulation in Equation (5.1) is space semi-discrete. In order to obtain a computationally feasible method, we introduce a sequence of time discretizations $\Theta = (\Theta_\ell, \ell \in \mathbb{N}_0)$ of the time interval $[0, T]$, for $T < +\infty$, each of equidistant/maximum time steps of size $\Delta_\ell t$. The time discretization at

level $\ell \in \mathbb{N}_0$, Θ_ℓ, is the partition of $[0, T]$ which is given by

$$\Theta_\ell = \{t_\ell^i \in [0, T] : t_\ell^i = i \cdot \Delta_\ell t, \ i = 0, \ldots, \frac{T}{\Delta_\ell t}\} .$$

We view the fully-discrete solution to Equation (3.5) as the solution to a nonlinear dynamical system according to

$$D_t(u_{\ell,\ell}) = F_\ell(t, u_{\ell,\ell}),$$

where D_t denotes the weak derivative with respect to time and the right hand side is given by (compare Equation (3.7))

$$F_\ell(t, v) = f - \nu A_\ell v - B_\ell(v, v) .$$

Here, A_ℓ denotes the discrete Stokes operator and B_ℓ the associated bilinear form.

We denote by $S_\ell = (S_\ell(t_\ell^i, 0), \ i = 0, \ldots, T/\Delta_\ell t)$ the fully-discrete solution operator that maps u_0 into $u_{\ell,\ell} = (u_{\ell,\ell}(t_\ell^i), \ i = 0, \ldots, T/\Delta_\ell t)$.

We assume that the spaces in $\mathcal{V}$ and the time discretizations Θ are chosen such that the following error bound holds.

Assumption 5.2. The sequence of fully-discrete solutions $(u_{\ell,\ell}, \ell \in \mathbb{N}_0)$ converges to the (unique, in space dimension $d \leq 2$) solution u of Equation (1.1). The space and time discretization error is bounded, for $\ell \in \mathbb{N}$ and $t \in \Theta_\ell$, by

(a) $$\|u(t) - u_{\ell,\ell}\|_H = \|S(t, 0)u_0 - S_\ell(t, 0)u_0\|_H \leq C \, (h_\ell^s + (\Delta_\ell t)^s),$$
for $d = 1$ and for some $s \in [0, 1)$;

(b) $$\|u(t) - u_{\ell,\ell}\|_H = \|S(t, 0)u_0 - S_\ell(t, 0)u_0\|_H \leq C \, (\frac{h_\ell^\sigma}{\nu} + \frac{(\Delta_\ell t)^\sigma}{\nu}),$$
for $d \geq 2$ and for some $\sigma > 0$.

Here, the constant $C > 0$ is independent of ℓ. With the choice $h_\ell \simeq \Delta_\ell t$ this reduces to

(a) $$\|u(t) - u_{\ell,\ell}\|_H \leq C \, h_\ell^s, \qquad\qquad (5.4)$$
for $d = 1$ and for some $s \in [0, 1)$;

(b) $$\|u(t) - u_{\ell,\ell}\|_H \leq C \, \frac{h_\ell^\sigma}{\nu}, \qquad\qquad (5.5)$$
for $d \geq 2$ and for some $\sigma > 0$.

Here, as in Assumption 5.1, for $d \geq 2$ we obtain the scale resolution requirement $\ell > \ell^*$, where $\ell^* \in \mathbb{N}_0$ such that $\nu \geq h_{\ell^*}^\sigma$.

Remark 5.3. The assumption of a space and time discretization with the convergence bounds (5.4) and (5.5) where the constant $C > 0$ is independent of the fluid viscosity in the norm $L^\infty(J; H)$ is strong. It amounts to saying that, essentially, the numerical scheme resolves the "bulk properties" (i.e. properties that are continuous in the norm of H) of the flow consistent to order $s, \sigma > 0$ independent of the small scale features of the flow. For example, flows which satisfy a ν-independent bound such as Equation (3.8) for $\|u\|_{L^\infty(J;V)}$ with V as in Equation (3.2) could afford Equation (5.4) with $s \simeq 1$.

In practice, therefore, Assumptions 5.1 and 5.2 imply that, for flows with large Reynolds number, some form of upscaling (such as, e.g. a turbulence model) which provides ν-independent convergence rates in the norm of H must be used for discretization levels above the viscous cut-off of the flow.

6. Multilevel Monte Carlo method

The space and time discretization introduces a bias in the error bound of the variance of the (discrete) Monte Carlo estimator. In this section, we first introduce the singlelevel approximation, where all samples of the Monte Carlo estimator are approximated with one common space and time discretization. Then we extend this approach to a multilevel discretization.

6.1. **Singlelevel Monte Carlo method.** With the discretization in hand we can combine the error in the spatial and temporal domain with the statistical sampling by the Monte Carlo method, leading to what we shall refer to as the singlelevel Monte Carlo approach.

We define, for $\ell \in \mathbb{N}_0$ and $t \in \Theta_\ell$, the Monte Carlo estimator with M_ℓ samples

$$E_{\mu_t}^{M_\ell}(\Phi_\ell) := \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \Phi(S_\ell(t,0)v^i) \ .$$

Here, S_ℓ denotes the fully-discrete solution operator, defined in Section 5. We assume further that the testfunction $\Phi \in \mathcal{C}$ satisfies a *Lipschitz condition*: there exists $C > 0$ such that

$$\forall u, v \in H : \quad |\Phi(u) - \Phi(v)| \leq C \|u - v\|_H \ . \tag{6.1}$$

We remark that Equation (6.1) follows from ϕ being continuously differentiable and with compact support. The constant depends on the maximum of ϕ' and on the H norms of $g_1, \ldots, g_k$. Under Equation (6.1), the Monte Carlo estimator admits the following mean-square error bound.

Proposition 6.1. *If, for $\Phi \in \mathcal{C}$ fulfilling Eq. (6.1) and $\ell \in \mathbb{N}_0$, the generalized moment of the statistical solution fulfills Assumption 5.2, for some $s \in [0,1)$ or some $\sigma > 0$ and $h_\ell \simeq \Delta_\ell t$, then the fully-discrete Monte Carlo estimator $E_{\mu_t}^{M_\ell}(\Phi_\ell)$ admits, for $t \in \Theta_\ell$, the bound*

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \frac{1}{\sqrt{M_\ell}} \big(\text{Var}_{\mu_t}(\Phi) \big)^{1/2} + \|\Phi - \Phi_\ell\|_{L^2((H,\mu_t);\mathbb{R})}$$

$$\leq C \big(\frac{1}{\sqrt{M_\ell}} + \rho(h_\ell) \big).$$

Here, $\rho(z) = z^s$ for $d = 1$ and $\rho(z) = \frac{z^\sigma}{\nu}$ for $d \geq 2$ and $z \in [0,1]$. The constant $C > 0$ is independent of ℓ, h_ℓ and of ν.

Proof. We use the triangle inequality and may write

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})}$$
$$\leq \|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi)\|_{L^2((H,\mu_t);\mathbb{R})} + \|E_{\mu_t}^{M_\ell}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})}.$$

The first term on the right hand side is bounded with Proposition 4.1 by

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \frac{1}{\sqrt{M_\ell}} (\text{Var}_{\mu_t}(\Phi))^{1/2} \leq C \frac{1}{\sqrt{M_\ell}} \Big(1 + \big(\int_H \|v\|_H^2 \, d\mu_0(v) \big)^{1/2} \Big).$$

Further, we have that

$$\|E_{\mu_t}^{M_\ell}(\Phi)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \|\Phi\|_{L^2((H,\mu_t);\mathbb{R})} = \big(\int_H |\Phi(v)|^2 \, d\mu_t(v) \big)^{1/2}.$$

Then the second term is bounded, with Assumption 5.2, by

$$
\begin{aligned}
\|E_{\mu_t}^{M_\ell}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})} &\leq \|\Phi - \Phi_\ell\|_{L^2((H,\mu_t);\mathbb{R})} \\
&\leq \mathbb{E}_{\mu_t}\big(|\Phi(S(t,0)v) - \Phi(S_\ell(t,0)v)|^2\big)^{1/2} \\
&\leq C\,\mathbb{E}_{\mu_t}\big(\|S(t,0)v - S_\ell(t,0)v\|_H^2\big)^{1/2} \\
&\leq C\,\big(\rho(h_\ell) + \rho(\Delta_\ell t)\big).
\end{aligned}
$$

Here, and in further calculations C denotes a varying constant which is independent of ν and the discretization parameters. Overall this leads to the bound

$$
\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})} \leq C\,\Big(\frac{1}{\sqrt{M_\ell}} + \rho(h_\ell) + \rho(\Delta_\ell t)\Big).
$$

In the case that $h_\ell \simeq \Delta_\ell t$, we have

$$
\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^{M_\ell}(\Phi_\ell)\|_{L^2((H,\mu_t);\mathbb{R})} \leq C\,\Big(\frac{1}{\sqrt{M_\ell}} + \rho(h_\ell)\Big).
$$

This proves the proposition. $\qquad\square$

The error bound for the singlelevel Monte Carlo estimator consists of two additive components, the approximation of the spatial and temporal discretization and of the Monte Carlo sampling. Although we only established an upper bound, one can show that this error is, indeed, of additive nature. This, in turn, indicates that the lack of scale-resolution in the spatial and temporal approximation, i.e. if the discretization underresolves the scale of viscous cut-off, can partly (in a mean-square sense) be offset by increasing the number of samples, on the mesh-level in the Monte Carlo approximation. This is in line with similar findings for multilevel Monte Carlo Galerkin discretizations for elliptic homogenization problems in [3]. In order to ensure that the total error in Proposition 6.1 is smaller than a prescribed tolerance $\epsilon > 0$, we require

$$
\frac{1}{\sqrt{M_\ell}}\big(\operatorname{Var}_{\mu_t}(\Phi)\big)^{1/2} + \|\Phi - \Phi_\ell\|_{L^2((H,\mu_t);\mathbb{R})} \leq \epsilon.
$$

This is the case if, for some $\eta \in (0,1)$

$$
\frac{1}{\sqrt{M_\ell}}\big(\operatorname{Var}_{\mu_t}(\Phi)\big)^{1/2} \leq \eta \cdot \epsilon \quad \text{and} \quad \|\Phi - \Phi_\ell\|_{L^2((H,\mu_t);\mathbb{R})} \leq (1 - \eta)\epsilon .
$$

In the case of a discretization, which does not fulfill the convergence requirement, the second term will be large (and η will be small), and an increase of the number of samples may achieve an error of order ϵ. In this case we can only achieve a tolerance ϵ that is bigger than $\|\Phi - \Phi_\ell\|_{L^2((H,\mu_t);\mathbb{R})}$.

6.2. **Multilevel Monte Carlo method.** The idea of the multilevel Monte Carlo estimator is to expand the expectation of the approximation of the solution on some discretization level L, for $t \in \Theta_L$, as the expectation of the solution on the (initial) discretization level 0 and a sum of correcting terms on all discretization levels $\ell = 1, \ldots, L$, i.e.,

$$
\mathbb{E}_{\mu_t}(\Phi_\ell) = \mathbb{E}_{\mu_t}(\Phi_0) + \sum_{\ell=1}^{L} \mathbb{E}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) .
$$

Then we approximate the expectation in each term on the right hand side with a Monte Carlo estimator with a level dependent number of samples, so that we may write

$$E_{\mu_t}^L(\Phi_L) = E_{\mu_t}^{M_0}(\Phi_0) + \sum_{\ell=1}^{L} E_{\mu_t}^{M_\ell}(\Phi_\ell - \Phi_{\ell-1}).$$

We call $E_{\mu_t}^L$ the multilevel Monte Carlo estimator for discretization level $L \in \mathbb{N}_0$. The multilevel Monte Carlo estimator has the following mean-square error bound.

Proposition 6.2. *If, for $\Phi \in \mathcal{C}$ fulfilling Eq. (6.1) and $L \in \mathbb{N}_0$, the generalized moment of the statistical solution fulfills Assumption 5.2, for all $\ell = 0, \ldots, L$ with $s \in [0, 1)$ or $\sigma > 0$ and $h_\ell \simeq \Delta_\ell t$, then the error of the fully-discrete multilevel Monte Carlo estimator $E_{\mu_t}^L(\Phi_L)$ admits, for $t \in \Theta_L$, the bound*

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^L(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} + \sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \big(\mathrm{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2}$$

$$\leq C \left(\rho(h_L) + \frac{1}{\sqrt{M_0}} \big(1 + \rho(h_0)\big) + \sum_{\ell=1}^{L} \frac{1}{\sqrt{M_\ell}} \big(\rho(h_\ell) + \rho(h_{\ell-1}) \big) \right),$$

where $\Phi_{-1} \equiv 0$ and $\rho(z) = z^s$ for $d = 1$ and $\rho(z) = \frac{z^\sigma}{\nu}$ for $d \geq 2$ and $z \in [0, 1]$.

Proof. We can bound the mean-square error of the multilevel Monte Carlo error by

$$\|\mathbb{E}_{\mu_t}(\Phi) - E_{\mu_t}^L(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})}$$
$$\leq \|\mathbb{E}_{\mu_t}(\Phi) - \mathbb{E}_{\mu_t}(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})} + \|\mathbb{E}_{\mu_t}(\Phi_L) - E_{\mu_t}^L(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})}$$
$$\leq \|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} + \|\mathbb{E}_{\mu_t}(\Phi_L) - E_{\mu_t}^L(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})}.$$

The first term satisfies (as in the proof of Proposition 6.1)

$$\|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} \leq C \big(\rho(h_L) + \rho(\Delta_L t) \big).$$

The second term is bounded, with the convention that $\Phi_{-1} \equiv 0$, by

$$\|\mathbb{E}_{\mu_t}(\Phi_L) - E_{\mu_t}^L(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \| \sum_{\ell=0}^{L} \big(\mathbb{E}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) - E_{\mu_t}^{M_\ell}(\Phi_\ell - \Phi_{\ell-1}) \big) \|_{L^2((H,\mu_t);\mathbb{R})}$$

$$\leq \sum_{\ell=0}^{L} \| \mathbb{E}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) - E_{\mu_t}^{M_\ell}(\Phi_\ell - \Phi_{\ell-1}) \|_{L^2((H,\mu_t);\mathbb{R})}$$

$$\leq \sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \big(\mathrm{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2}.$$

Further, we have, for $\ell = 1, \ldots, L$, with Equation (6.1) and Assumption 5.2,

$$\frac{1}{\sqrt{M_\ell}} \big(\mathrm{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2} \leq \frac{1}{\sqrt{M_\ell}} \| \Phi_\ell - \Phi_{\ell-1} \|_{L^2((H,\mu_t);\mathbb{R})}$$

$$\leq \frac{1}{\sqrt{M_\ell}} \big(\| \Phi - \Phi_\ell \|_{L^2((H,\mu_t);\mathbb{R})} + \| \Phi - \Phi_{\ell-1} \|_{L^2((H,\mu_t);\mathbb{R})} \big)$$

$$\leq C \frac{1}{\sqrt{M_\ell}} \big(\rho(h_\ell) + \rho(\Delta_\ell t) + \rho(h_{\ell-1}) + \rho(\Delta_{\ell-1} t) \big).$$

For $\ell = 0$ we can bound the term as follows

$$\frac{1}{\sqrt{M_0}} \big(\operatorname{Var}_{\mu_t}(\Phi_0) \big)^{1/2} \leq \frac{1}{\sqrt{M_0}} \big(\|\Phi_0 - \Phi\|_{L^2((H,\mu_t);\mathbb{R})} + \|\Phi\|_{L^2((H,\mu_t);\mathbb{R})} \big)$$

$$\leq C \frac{1}{\sqrt{M_0}} \big(\rho(h_0) + \rho(\Delta_0 t) \big) + \frac{1}{\sqrt{M_0}} \Big(\int_H |\Phi(v)|^2 \, d\mu_t(v) \Big)^{1/2}$$

$$\leq C \frac{1}{\sqrt{M_0}} \big(\rho(h_0) + \rho(\Delta_0 t) + 1 \big) .$$

With the assumption $h_\ell \simeq \Delta_\ell t$, this proves the assertion. $\qquad\square$

By adjustment of the constant, we get from Proposition 6.2 the following proposition

Proposition 6.3. *Assume the setting of Proposition 6.2 and that, for all $\ell = 1, \ldots, L$, it holds $h_\ell \simeq \Delta_\ell t$ and that $h_{\ell-1} \leq \varrho h_\ell$, with some reduction factor $0 < \varrho < 1$ independent of ℓ. Then, there exists $C(\varrho) > 0$ independent of L, such that there holds the error bound*

$$\|\mathbb{E}_{\mu_t}(\Phi) - E^L_{\mu_t}(\Phi_L)\|_{L^2((H,\mu_t);\mathbb{R})} \leq \|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} + \sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \big(\operatorname{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2}$$

$$\leq C(\varrho) \Big(\rho(h_L) + \frac{1}{\sqrt{M_0}} + \sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \rho(h_\ell) \Big).$$

This leads again to the question of the sample numbers $(M_\ell, \ell = 1, \ldots, L)$ that would yield a given (mean kinetic energy) error threshold ϵ. We require

$$\|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} + \sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \big(\operatorname{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2} \leq \epsilon,$$

which leads, if we assume that $\eta_L \in (0,1)$, to the requirement

$$\|\Phi - \Phi_L\|_{L^2((H,\mu_t);\mathbb{R})} \leq (1 - \eta_L) \, \epsilon,$$

and to

$$\sum_{\ell=0}^{L} \frac{1}{\sqrt{M_\ell}} \big(\operatorname{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2} \leq \eta_L \, \epsilon.$$

If we have that for some $\theta_\ell > 0$, for $\ell = 0, \ldots, L$,

$$\big(\operatorname{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1}) \big)^{1/2} \leq \theta_\ell,$$

then, to equilibrate the error for each level $\ell = 1, \ldots, L$, we choose the sample sizes

$$M_\ell = \theta_\ell^2 \alpha_\ell (\eta_L \epsilon)^{-2} \tag{6.2}$$

for a sequence $(\alpha_\ell, \ell = 1, \ldots, L)$ with $\alpha_\ell \in [0,1]$, subject to the constraint $\sum_{\ell=1}^{L} \alpha_\ell = 1$. We determine the required number M_ℓ of Monte Carlo samples on each discretization level $\ell = 0, \ldots, L$ based on equilibration of the errors arising from each term $\operatorname{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1})$ such that the total mean-square error from Proposition 6.3 is bounded by the prescribed tolerance $\epsilon > 0$. This is only possible if the convergence requirement is fulfilled for level L, since then we can choose η_L according to satisfy a preset error bound. However, the convergence requirement might not be fulfilled for all $\ell < L$, hence, for those levels we have to sample accordingly. Especially for the first level ℓ^* where the solution is scale-resolved, the term $\operatorname{Var}_{\mu_t}(\Phi_{\ell^*} - \Phi_{\ell^*-1})$ will be large, as will be θ_{ℓ^*}; thus α_{ℓ^*} has to be chosen accordingly. Since

it is difficult to determine the values $\mathrm{Var}_{\mu_t}(\Phi_\ell - \Phi_{\ell-1})$ we estimate sample numbers from the second (more general) bound in Proposition 6.3. We refer to [4] for an analysis of the computational complexity of multilevel estimators in the case of weak or strong errors.

We proceed to determine the numbers M_ℓ of Monte Carlo samples. To this end, we continue to work under Assumption 5.2. We start our derivation with the case $d \geq 2$ and in the case that at least on the finest level the convergence requirement is fulfilled, i.e., $h_L^\sigma < \nu$. We consider the case where the convergence requirement is *not* fulfilled for all levels up to level $\ell^*(\nu)$. In this case, for $0 \leq \ell^*(\nu) < L$ (meaning $h_{\ell^*(\nu)}^\sigma \geq \nu$ and $h_{\ell^*(\nu)+1}^\sigma < \nu$), we choose on the first level the sample number

$$M_0 = \mathrm{O}\left(\left(\frac{\nu}{h_L^\sigma}\right)^2\right)$$

to equilibrate the statistical and the discretization error contributions. Here, and in what follows, all constants implied in the Landau symbols $\mathrm{O}(\cdot)$ are independent of ν. According to this convergence analysis, the multilevel Monte Carlo sample numbers M_ℓ, for discretization levels $\ell = 1, \ldots, \ell^*(\nu), \ldots, L$ should be chosen according to

$$M_\ell = \mathrm{O}\left(\left(\frac{h_\ell^\sigma}{h_L^\sigma}\right)^2 \ell^{2(1+\gamma)}\right),$$

for $\gamma > 0$ arbitrary (with the constant implied in O depending on γ). Note that h_ℓ^σ might be large for underresolved discretization levels if no turbulence model is used. For $d = 1$ we optimize the sample numbers by choosing

$$M_0 = \mathrm{O}\left(\left(\frac{1}{h_L^s}\right)^2\right)$$

and, for $\ell = 1, \ldots, L$ and $\gamma > 0$,

$$M_\ell = \mathrm{O}\left(\left(\frac{h_\ell^s}{h_L^s}\right)^2 \ell^{2(1+\gamma)}\right).$$

This choice of sample numbers is in line with Equation (6.2) for one particular sequence $(\alpha_\ell, \ell = 1, \ldots, L)$. We remark that the Lipschitz condition (see Equation (6.1)) could be substituted by a so-called *Hölder condition*: there exists a constant $C > 0$ and some $0 < \alpha \leq 1$ such that

$$\forall u, v \in H : \quad |\Phi(u) - \Phi(v)| \leq C\|u - v\|_H^\alpha.$$

While the convergence analysis stays verbatim the same, the convergence rate and the dependence of the sample numbers on the discretization parameters in Proposition 6.1, Proposition 6.2 and Proposition 6.3 would include the additional parameter α.

7. Discretization of the initial distribution μ_0

The numerical simulation of a statistical solution requires sampling from the measure μ_0 defined on the generally infinite dimensional space H. In this section we give a convergence result for finite dimensional, "principal component" approximations of this initial measure μ_0. We follow here closely the approach in [5].

The initial distribution μ_0 is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and is assumed to be given as an image measure under an H-valued random variable with distribution μ_0. This random variable is defined as a mapping from the measurable space $(\Omega, \mathcal{F})$ into the measurable space $(H, \mathcal{B}(H))$ such that $\mu_0 = X \circ \mathbb{P}$. We assume throughout the numerical experiments

that μ_0 is a Gaussian measure supported on H or on a subspace of H (see, e.g., [13, Sec. 5] for a detailed discussion of this case).

Gaussian measures on a separable, infinite-dimensional Hilbert space H are completely characterized by the mean $m \in H$ and covariance operator Q defined on H, being a symmetric, nuclear trace-class operator. Any Gaussian random variable $X \in L^2(\Omega, \mathbb{P}; H)$ can then be represented by its Karhunen–Loève expansion

$$X = m + \sum_{i \in \mathbb{N}} \sqrt{\lambda_i} \beta_i w_i,$$

where $((\lambda_i, w_i), i \in \mathbb{N})$ is a complete orthonormal system in H and consists of eigenvalues and eigenfunctions of Q. The sequence $(\beta_i, i \in \mathbb{N})$ consists of real-valued, independent, (standard) normal-distributed random variables. With κ-term truncations of Karhunen–Loève expansions we define a sequence of random variables $(X^\kappa, \kappa \in \mathbb{N})$ given by

$$X^\kappa = m + \sum_{i=1}^{\kappa} \sqrt{\lambda_i} \beta_i w_i \;,$$

with mean $m \in H$ and covariance operator Q^κ. The sequence truncated sums X^κ converge $\mathbb{P}$-a.s. to X for $\kappa \to +\infty$. We control the $L^2(\Omega; H)$-error of this truncation by the decay of the eigenvalues. To this end, we write

$$\|X - X^\kappa\|_{L^2(\Omega;H)} = \|\sum_{i=\kappa+1}^{\infty} \sqrt{\lambda_i} \beta_i w_i\|_{L^2(\Omega;H)} \;.$$

Then, we have the following lemma.

Lemma 7.1. *If the eigenvalues $(\lambda_i, i \in \mathbb{N})$ of the covariance operator Q of the Gaussian random variable X on H have a rate of decay of $\lambda_i \leq C\, i^{-\gamma}$, then the sequence $(X^\kappa, \kappa \in \mathbb{N})$ converges to X in $L^2(\Omega; H)$ and the error is bounded by*

$$\|X - X^\kappa\|_{L^2(\Omega;H)} \leq C\, \frac{1}{\sqrt{\gamma - 1}} \kappa^{-\frac{\gamma - 1}{2}} \;.$$

As an example we expand data and solutions in terms of eigenfunctions of the Stokes operator A defined in Equation (3.4). The Stokes operator A is a closed, unbounded, self-adjoint operator on its domain $\mathcal{D}(A)$. By the spectral theorem, A has for either set of boundary conditions a discrete spectrum $\Sigma \subset \mathbb{R}_+$, which consists of real eigenvalues $\lambda_j \in \Sigma$ which accumulate only at infinity, and which admits a countable sequence of eigenfunctions $(w_j, j \in \mathbb{N})$ which are dense in H and in V. We assume that the sequence $(w_j, j \in \mathbb{N})$ constitutes an orthonormal basis of H. Then, for every $v \in H$ we may write

$$v = \sum_{j \in \mathbb{N}} v_j w_j \;, \qquad v_j = (v, w_j)_H$$

and $\|v\|_H^2 = \sum_{\lambda \in \Sigma} |v_\lambda|^2$. Domains of fractional powers A^a for any $a \in \mathbb{R}$ can be characterized in terms of the Fourier coefficients v_λ:

$$v \in \mathcal{D}(A^a) \iff \sum_{j \in \mathbb{N}} \lambda_j^a |v_j|^2 < \infty \;.$$

Naturally, the $(v_\lambda, \lambda \in \Sigma)$ depend on the boundary conditions and on the domain D. The Stokes operator being a second order, strongly elliptic differential operator, classical results on spectral asymptotics (see, e.g., [14, Eqn. (II.6.25)]) imply that for any boundary conditions the

eigenvalues $\lambda \in \Sigma$, enumerated in increasing magnitude and repeated according to multiplicity by λ_j, grow asymptotically as

$$\lambda_j \sim \lambda_1 j^{2/d} \qquad \text{as} \qquad j \to \infty. \tag{7.1}$$

In the particular case of periodic boundary conditions, as given in Equation (1.3), and for the domain $D = (0, \mathcal{I})^d$, fully explicit expressions are available (see, e.g. [35, Chap. I.2.2] or [14, Eqns. (II.6.16) - (II.6.18)]): Denoting by A the Stokes operator in Equation (3.4), we have

$$A w_{k,\alpha} = \lambda_k w_{k,\alpha} \,, \quad w_{k,\alpha} \in D(A), \alpha = 1, ..., d,$$

where

$$w_{k,\alpha} = \left(e_\alpha - \frac{k_\alpha k}{|k|^2} \right) \exp(2i\pi k \cdot x/\mathcal{I}) \,, \qquad \lambda_k = 4\pi^2 |k|^2 \mathcal{I}^{-2}. \tag{7.2}$$

Here, we set $k = (k_1, ..., k_d) \in \mathbb{Z}^d$, $\alpha = 1, 2, ..., d$, and e_α denotes the unit vector on coordinate axis α in $\mathbb{R}^d$. It is easily verified that Equation (7.2) is consistent with Equation (7.1).

For the numerical realization of the (multilevel) Monte Carlo method, and in particular for the numerical experiments in Section 8 ahead, we need to draw samples from the initial distribution. As an example we therefore introduce a Gaussian distribution on $H = L^2_{\text{per}}(D)$, where $D = (0, 1)$ (with periodic boundary conditions). In the univariate case, a basis of $L^2_{\text{per}}(D)$ is given by $(w_i, i \in \mathbb{N})$, where we have $w_i(x) = \sin(2i\pi x)$. Then the covariance operator Q is with Mercer's theorem defined, for $\phi \in L^2_{\text{per}}(D)$, as

$$Q\phi(x) = \int_D q(x, y)\phi(y) dy$$

where the kernel q is

$$q(x, y) = \sum_{i \in \mathbb{N}} \lambda_i w_i(x) w_i(y) = \sum_{i \in \mathbb{N}} \lambda_i \sin(2i\pi x) \sin(2i\pi y).$$

Now, we can choose any series $(\lambda_i, i \in \mathbb{N})$ with $\sum_{i \in \mathbb{N}} \lambda_i < \infty$ to define a covariance operator Q on H which is trace class. One possible choice would be $\lambda_i \simeq i^{-\alpha}$, for $\alpha > 2$.

More generally, in any space dimension $d \geq 2$, we can parametrize a Gaussian measure on the divergence-free initial velocity fields, i.e. on $H = \{v \in L^2(D)^d : \text{div } v = 0\}$, by prescribing a divergence-free mean velocity field $\langle u_0 \rangle = \mathbb{E}_{\mu_0}(H)$ and the Gaussian covariance operator Q on H. We shall choose

$$Q = A^{-\delta} \tag{7.3}$$

where $2/d\,\delta = \alpha$ and A denotes the Stokes operator defined in Equation (3.4). Then draws of the random initial velocity u_0 with law μ_0 can be obtained from the Karhunen–Loève expansion

$$u_0(x, \omega) = \langle u_0 \rangle + \sum_{i \in \mathbb{N}} \sqrt{\mu_i} \xi_i(\omega) w_i(x)$$

where $w_i \in V$ denote the eigenfunctions of the Stokes operator A, where $\xi_i \sim \mathcal{N}(0, 1)$ are independent standard normal random variables taking values in $\mathbb{R}$ and where μ_i are the Karhunen–Loève eigenvalues.

For the convergence in Lemma 7.1, for the "Stokes"-covariances A^a in Equation (7.3), the spectral mapping theorem implies

$$\mu_i = \lambda_i^{-\delta}$$

with the asymptotic behavior of λ_i as $i \to \infty$ given by Equation (7.1).

8. Numerics

We describe numerical experiments for the one dimensional case of the Navier–Stokes equations with stochastic initial data. As described in the introduction, the Navier–Stokes equation in one space dimension reduces to the, so called, viscous Burgers' equation (see Equation (1.4)) with viscosity $\nu > 0$. Despite being a gross simplification of the Navier–Stokes equations, statistical solutions of Burgers' equation have been proposed as models for (some features of) turbulent flow (e.g. [19]). As emphasized in Section 5.1, in the case $d = 1$, i.e. for scalar problems in one spatial dimension, the bound in Equation (5.2) in Assumption 5.1 holds with $s = 1/2$ and with a constant $C > 0$ independent of ν (see [25]). If the mesh used for the space discretization is sufficiently fine, then the first order Finite Volume method converges with rate $s = 1$ in $L^1(D)$ due to the high spatial regularity of the solution u, albeit with constants which blow up as the viscosity $\nu \to 0$. Convergence bounds with constants that are independent of the viscosity (such as those stipulated in Assumption 5.1) seem to be presently available only in one space dimension. Specifically, we consider periodic boundary conditions and the physical domain $D = [0,1] \subset \mathbb{R}$. Then, the viscous Navier-Stokes equations simplify to the scalar, viscous Burgers' equation

$$\frac{\partial}{\partial t}u + \frac{1}{2}\frac{\partial}{\partial x}(u^2) = \nu\frac{\partial^2}{\partial x^2}u + f, \quad \forall x \in D,\ t \in [0,T],\ \omega \in \Omega, \tag{8.1}$$

which is completed with the random initial condition $u(0) = u_0 \in L^2(\Omega, L^1(D) \cap L^\infty(D))$, inducing an initial measure μ_0 on $L^1(D) \cap L^\infty(D) \subset H = L^2(D)$ with finite second moments.

The stochastic initial data u_0 is assumed to be given by its Karhunen–Loève expansion,

$$u_0(x,\omega) = \langle u(x)\rangle + \sum_{m=1}^{\infty} \sqrt{\lambda_m}w_m(x)Y_m(\omega), \tag{8.2}$$

with eigenvalues $(\lambda_m, m \in \mathbb{N}) \in \ell^{\frac{1}{2}}(\mathbb{N})$, eigenfunctions $w_m \in L^2(D)$, $\|w_m\|_{L^2(D)} = 1$, mean field $\langle u\rangle \in L^2(D)$, and independent, identically distributed random variables $Y_m \in L^2(\Omega)$ with zero mean and finite variance.

8.1. **Karhunen–Loève expansion with normally distributed $(Y_m, m \in \mathbb{N})$.** In particular, we fix the domain $D = [0,1]$ and choose *normally* distributed random variables $Y_m \sim \mathcal{N}(0,1)$, i.e. with zero mean and unit variance, eigenvalues $\lambda_m = m^{-2.5}$ for $m \leq 8$ and zero otherwise, eigenfunctions $w_m(x) = \sin(2\pi mx)$, and the mean field $\langle u(x)\rangle \equiv 0$, i.e.

$$u_0(x,\omega) = \sum_{m=1}^{8} \frac{1}{m^{5/4}}\sin(2\pi mx)Y_m(\omega). \tag{8.3}$$

The kinematic viscosity is chosen to be $\nu = 10^{-3}$.

All simulations reported below were performed on Cray XE6 in CSCS [2] with the recently developed massively parallel code ALSVID-UQ [1, 34, 30].

The initial data in Equation (8.3) and the reference solution u_{ref} at time $t = 2$ are depicted in Figure 1. The solid line represents the mean $\mathbb{E}_{\mu_t}(u_{\text{ref}})$ and the dashed lines represent the mean plus/minus the standard deviation $\sqrt{\text{Var}_{\mu_t}(u_{\text{ref}})}$ of the (random) solution u_{ref} at every point $x \in D$. The variance and therefore the standard deviation can easily be calculated by

$$\text{Var}_{\mu_0}(u_0(x)) = \sum_{m=1}^{8} (\frac{1}{m^{5/4}}\sin(2\pi mx))^2, \quad x \in D.$$

The solution is computed with a standard first-order Finite Volume scheme using the Rusanov HLL solver on a spatial grid in D of size 32768 cells and the explicit forward Euler time stepping (see [27]) with the CFL number set to 0.9. The number of levels of refinement is 9 (the coarsest level has 64 cells). The number of samples is chosen according to the analysis in Section 6 with $s = 1$, i.e.

$$M_\ell = M_L 2^{2(L-\ell)}, \quad \ell = 0, \dots, L,$$

where the number of samples on the finest mesh set to $M_L = 4$ (this leads to $M_0 = 262144$). The simulation took 50 minutes (wall-clock time) on 256 cores.
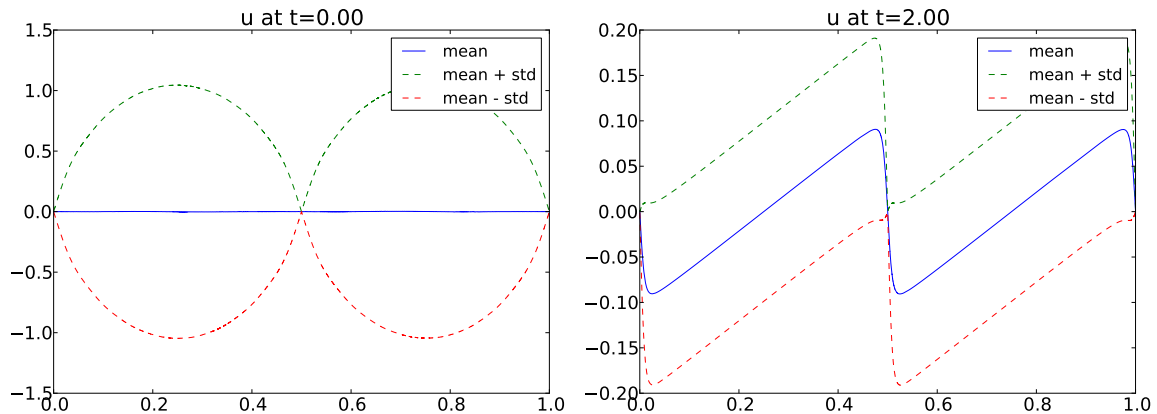
FIGURE 1. Reference solution computed using a multilevel Monte Carlo Finite Volume method with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with normally distributed Y_m.

Numerical error convergence analysis in the fully resolved case. Next, following Definition 3.2, for $k = 1$, $\phi(\xi) = \xi$ and a given kernel $g_1 \in L^\infty(D)$, we define a continuous, linear functional Φ on $L^1(D) \cap L^\infty(D)$ by

$$\Phi(u)(t,\omega) = \int_D u(x,t,\omega)g_1(x)dx, \quad \forall t \in [0,T] \ \omega \in \Omega. \tag{8.4}$$

Note, that formally the function $\phi(\xi) = \xi$ is not compactly supported. However, for one-dimensional problems ($d = 1$), there holds an energy bound (we refer to the results in [27]) with respect to the initial data $u_0(\cdot,\omega)$,

$$\|u(\cdot,t,\omega)\|_{L^2(D)} \leq \|u_0(\cdot,\omega)\|_{L^2(D)}.$$

Since the values of the inner product can be bounded for every t and ω by

$$|(u(\cdot,t,\omega), g_1)_H| \leq \|u(\cdot,t,\omega)\|_{L^2(D)}\|g_1\|_{L^2(D)} \leq \|u_0(\cdot,\omega)\|_{L^2(D)}\|g_1\|_{L^2(D)} < \infty,$$

the function $\phi(\xi)$ may be modified for large values of ξ, enforcing the required compact support of $\phi(\xi)$ in the Definition 3.2. We note, that such modification is ω-dependent, and hence a more stringent bound in the $\|u_0\|_{L^\infty(\Omega, L^2(D))}$-norm on the initial data is required instead, i.e. we require that $\|u_0(\cdot,\omega)\|_{L^2(D)} < C$ holds $\mathbb{P}$-a.s. for some constant $C < \infty$. Such a bound holds for the *uniformly* distributed initial condition presented in the forth-coming

Section 8.2, however, it does *not* hold for the *normally* distributed initial condition considered in the current section.

In the following numerical experiment, the function g_1 in Equation (8.4) is $g_1(x) = (x - 0.5)^3$. The function g_1 formally does not give rise to a cylindrical test function $\Phi \in \mathcal{C}$. However, since in the present example in space dimension $d = 1$ there holds a maximum principle for all $\nu \geq 0$, g_1 may be modified for large values of x in order to fit into the theoretical setting *for the particular example considered here* (this argument may *not* apply for the Navier–Stokes equations in $d = 3$).

Using multilevel Monte Carlo Finite Volume approximations for the mean $\mathbb{E}_{\mu_t}(\Phi_{\mathrm{ref}})$ and the variance $\mathrm{Var}_{\mu_t}(\Phi_{\mathrm{ref}})$ from Figure 1 as a reference solution, we compute approximate solutions u_ℓ using Monte Carlo Finite Volume and multilevel Monte Carlo Finite Volume methods on a family of meshes with spatial resolutions ranging from $n_0 = 64$ cells up to $n_L = 2048$ cells. We monitor the convergence of the errors in $E_{\mu_t}^L(\Phi_L)$ and $\mathrm{Var}_{\mu_t}^L(\Phi_L)$,

$$\varepsilon_L^{\mathbb{E}} = \left| \mathbb{E}_{\mu_t}(\Phi_{\mathrm{ref}}) - E_{\mu_t}^L(\Phi_L) \right|, \quad \varepsilon_L^{\mathbb{V}} = \left| \mathrm{Var}_{\mu_t}(\Phi_{\mathrm{ref}}) - \mathrm{Var}_{\mu_t}^L(\Phi_L) \right| .$$

The number of samples on the finest mesh is set to $M_L = 4$. The number of levels for the multilevel Monte Carlo Finite Volume method is chosen so that the coarsest level contains 64 cells. Since $1/64 \approx 0.015 < \sqrt{\nu} = 10^{-1.5} \approx 0.03$, the "viscous cut-off" scale (which, in the present problem in one space dimension $d = 1$ coincides with the scale of the viscous shock profile) of the solution u_ℓ is *resolved* on *every* mesh resolution level $\ell = 0, \ldots, L$.

Error estimator. Since the solution is a random field, the discretization error ε_ℓ is a random quantity as well. For error convergence analysis we, therefore, compute a statistical estimator by averaging estimated discretization errors from several independent runs. We compute the error in Proposition 6.3 by approximating the $L^2(H, \mathbb{R})$-norm by Monte Carlo sampling. Let Φ_{ref} denote the reference solution and $(\Phi_L^{(k)}, k = 1, \ldots, K)$ be a sequence of independent approximate solutions obtained by running the Monte Carlo Finite Volume or multilevel Monte Carlo Finite Volume solver K times corresponding to K independent, identically distributed realizations of the initial velocity, drawn from μ_0. The $L^2(\Omega; H)$-based relative percentage error estimator is defined to be

$$\mathcal{R}\varepsilon_L^{\mathbb{E}} = 100 \times \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(\frac{\varepsilon_L^{\mathbb{E},(k)}}{|\mathbb{E}_{\mu_t}(\Phi_{\mathrm{ref}})|} \right)^2}, \quad \mathcal{R}\varepsilon_L^{\mathbb{V}} = 100 \times \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(\frac{\varepsilon_L^{\mathbb{V},(k)}}{|\mathrm{Var}_{\mu_t}(\Phi_{\mathrm{ref}})|} \right)^2}.$$

In order to obtain an accurate estimate of $\mathcal{R}\varepsilon_L^{\mathbb{E}}$ and $\mathcal{R}\varepsilon_L^{\mathbb{V}}$, the number K must be large enough to ensure a sufficiently small (< 0.1) relative variance $\sigma^2(\mathcal{R}\varepsilon_L^{\mathbb{E}})$ and $\sigma^2(\mathcal{R}\varepsilon_L^{\mathbb{V}})$, which can be estimated, for the expectation, by

$$\sigma^2(\mathcal{R}\varepsilon_L^{\mathbb{E}}) \approx \sigma_K^2(\mathcal{R}\varepsilon_L^{\mathbb{E}}) = \frac{1}{K-1} \frac{E_{\mu_t}^K \left((\mathcal{R}\varepsilon_L^{\mathbb{E}})^2 - E_{\mu_t}^K(\mathcal{R}\varepsilon_L^{\mathbb{E}})^2 \right)}{E_{\mu_t}^K(\mathcal{R}\varepsilon_L^{\mathbb{E}})} .$$

For the variance we have an equivalent expression. We found $K = 30$ to be sufficient for our numerical experiments. Next, we analyze the relative percentage error convergence plots of mean and variance.

In Figure 2, we plot the relative percentage error $\varepsilon_L^{\mathbb{E}}$ against the number of cells on discretization level L in the left subplot and versus the computational work (runtime) in the right subplot. Both multilevel Monte Carlo and Monte Carlo methods give similar relative percentage errors for the same spatial resolution. However, there is a significant difference in

the runtime: multilevel Monte Carlo methods are two orders of magnitude faster than plain
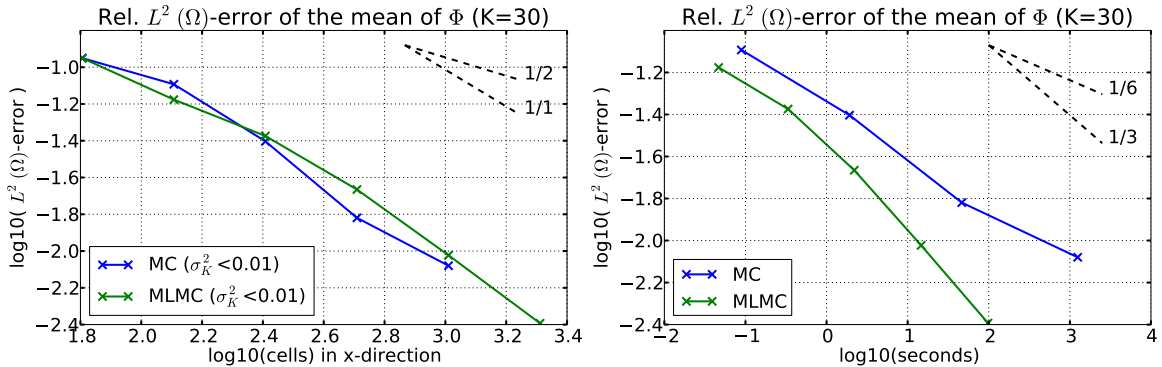Monte Carlo methods.

FIGURE 2. Convergence of the relative percentage error $\varepsilon_L^{\mathbb{E}}$ of the mean $\mathbb{E}_{\mu_t}(\Phi)$
of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu = 10^{-3}$ and
random initial data u_0 with normally distributed Y_m.

The lower dashed line in the top-right corner of each plot in Figure 2 (and all subsequent
figures) indicates the *expected* convergence rate of the multilevel Monte Carlo method obtained
in Proposition 6.3. These expected convergence rates coincide with the observations in the
numerical experimental data.

In Figure 3, we plot the relative percentage error $\varepsilon_L^{\mathbb{V}}$ versus the number of cells on dis-
cretization level L in the left subplot and versus the computational work (runtime) in the
right subplot. Analogously as in the plots for the expectation, both multilevel Monte Carlo
and Monte Carlo methods give similar errors for the same spatial resolution. In terms of the
required computational work for one percent error, multilevel Monte Carlo methods are, in
this example, two orders of magnitude faster than plain Monte Carlo methods.
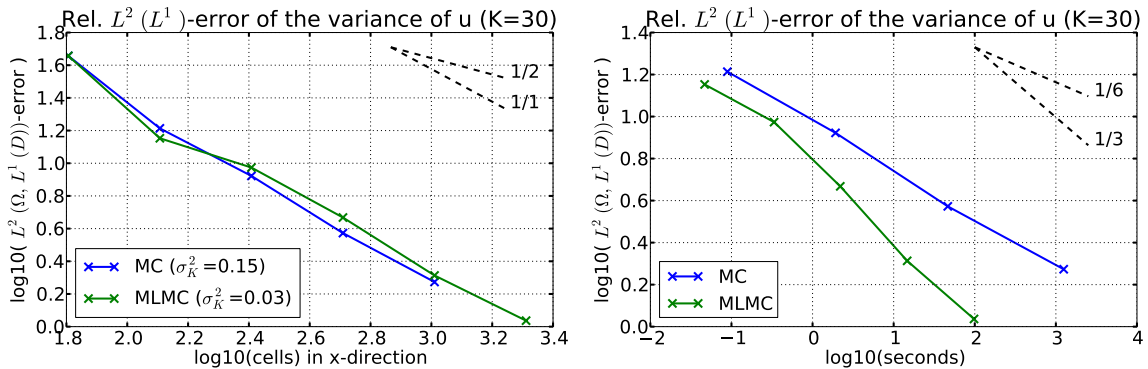

FIGURE 3. Convergence of the relative percentage error $\varepsilon_L^{\mathbb{V}}$ of the variance
$\mathbb{V}_{\mu_t}(\Phi)$ of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu =
10^{-3}$ and random initial data u_0 with normally distributed Y_m.

Numerical error convergence analysis in the under-resolved case. Next, we repeat the error convergence analysis for Burgers' equation, but this time with much *fewer* cells on the coarsest mesh resolution in the multilevel Monte Carlo Finite Volume estimator. In particular, instead of taking 64 cells on the coarsest mesh resolution, we will take only 8 cells by adding three more levels of mesh refinement. Since in this case $1/8 > \sqrt{\nu} = 10^{-1.5} \approx 0.03$, the viscous cut-off length scale of the solution u_ℓ is *not* resolved on every mesh resolution level, in particular, it is resolved only on the mesh resolution levels $\ell = 3, \ldots, L$, and it is under-resolved on $\ell = 0, 1, 2$. Notice, that the number of cells on the finer mesh resolutions stays the same as in the previous experiment, where $n_3 = 64, \ldots, n_L = 2048$. Note also that by the theory in [25], the presently used numerical scheme converges *robustly* in H with order $s \simeq 1/2$, meaning that the constant in the convergence bound is independent of ν, and that this computational example is within Assumption 5.2 (a).
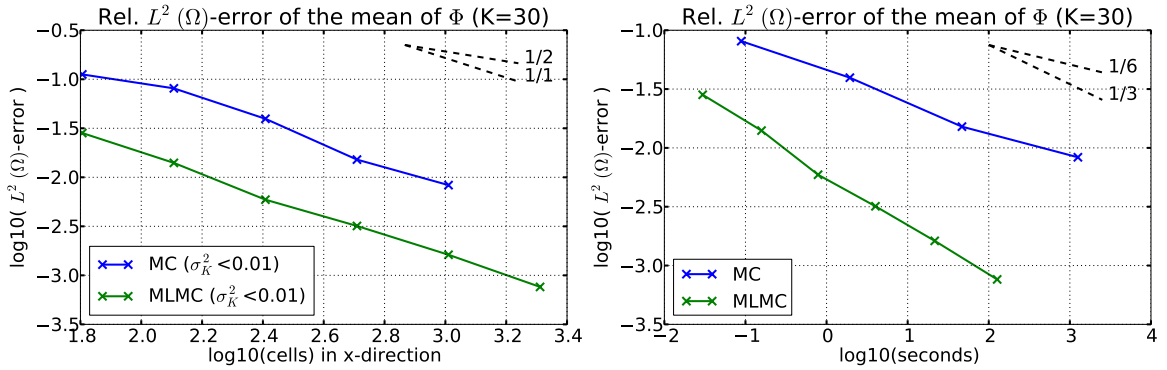

FIGURE 4. Convergence of the relative percentage error $\varepsilon_L^{\mathbb{E}}$ of the mean $\mathbb{E}_{\mu_t}(\Phi)$ of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with normally distributed Y_m.
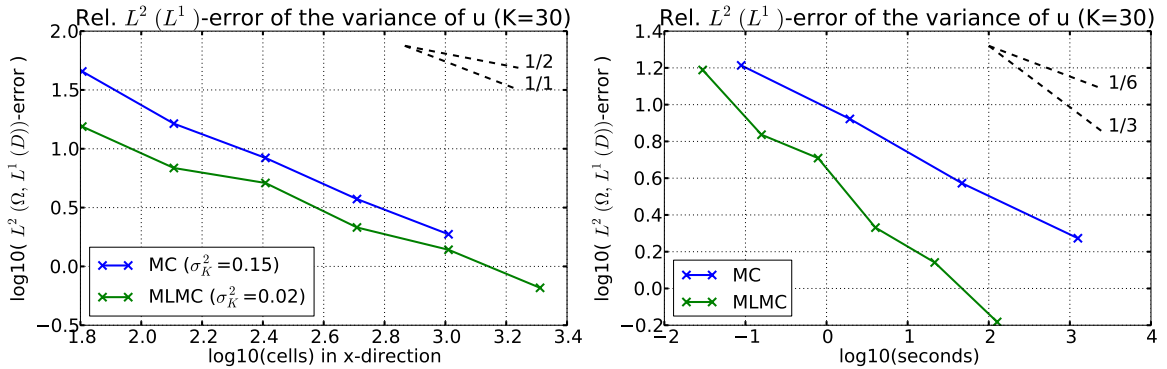

FIGURE 5. Convergence of the relative percentage error $\varepsilon_n^{\mathbb{V}}$ of the variance $\mathbb{V}_{\mu_t}(\Phi)$ of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with normally distributed Y_m.

In Figure 4, we plot the relative percentage error $\varepsilon_L^{\mathbb{E}}$ against the number of cells n_L in the left subplot and versus computational work (runtime) in the right subplot for the case

of 8 cells on the coarsest resolution. Even in the presence of multiple under-resolved levels, the error convergence of the multilevel Monte Carlo Finite Volume method is faster than the previous setup (compared to Figure 2). In Figure 5, we plot the relative percentage error $\varepsilon_L^{\mathbb{V}}$ versus the number of cells n_L in the left subplot and versus the computational work (runtime) in the right subplot for case of 8 cells on the coarsest resolution. Again, even in the presence of multiple under-resolved levels, the error convergence of the multilevel Monte Carlo Finite Volume method is faster than the previous setup (compared to Figure 3).

8.2. **Karhunen–Loève expansion with uniformly distributed** $(Y_m, m \in \mathbb{N})$**.** In this section, we perform the same set of numerical experiments as in Section 8.1, but this time with *uniformly* distributed random variables $(Y_m, m \in \mathbb{N})$ in Equation (8.2), i.e. $Y_m \sim \mathcal{U}([-1, 1])$.
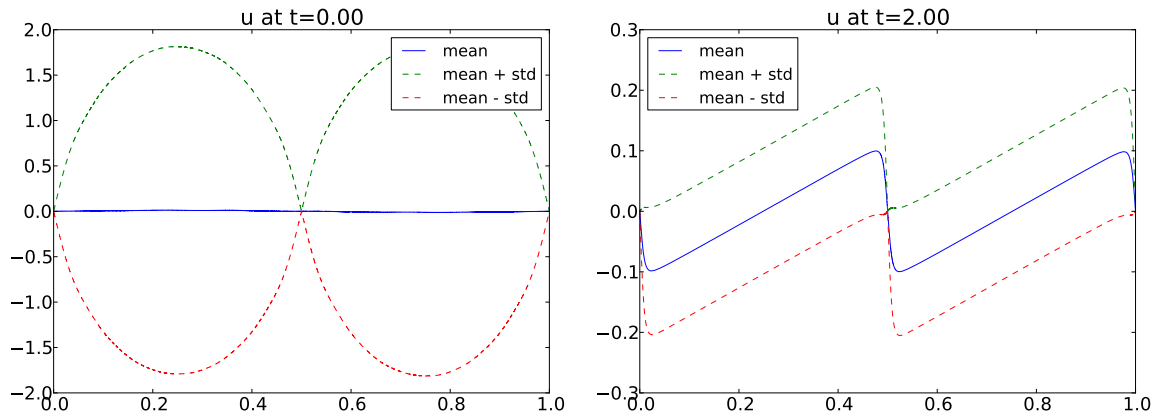

FIGURE 6. Reference solution computed using a multilevel Monte Carlo Finite Volume method with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with uniformly distributed Y_m.

The reference solution is depicted in Figure 6 and the relative percentage error convergence analysis in Figure 7 - 8. The setup is equivalent to the previous section. The reference solution is computed with a standard first-order Finite Volume scheme using the Rusanov HLL solver on a spatial grid in D of size 16384 cells and the explicit Forward Euler time stepping with CFL number set to 0.9. The number of levels of refinement is 8 (the coarsest level has 64 cells). The number of samples on the finest mesh is $M_L = 4$ and on the coarsest mesh is $M_0 = 65536$.

9. CONCLUSIONS

We have developed and analyzed a novel computational approach for the numerical simulation of statistical solutions to the Navier–Stokes equations, governing the evolution of viscous, incompressible flow. The approach consists in a multilevel Monte Carlo sampling strategy for the efficient numerical approximation of ensemble averages and of bulk properties of the statistical solution of these equations. The proposed approach is based on the computational approximation of generalized moments of the statistical solution to the Navier–Stokes equations by sampling, combined with the use of standard space and time discretization methods (and flow codes) for each sample.
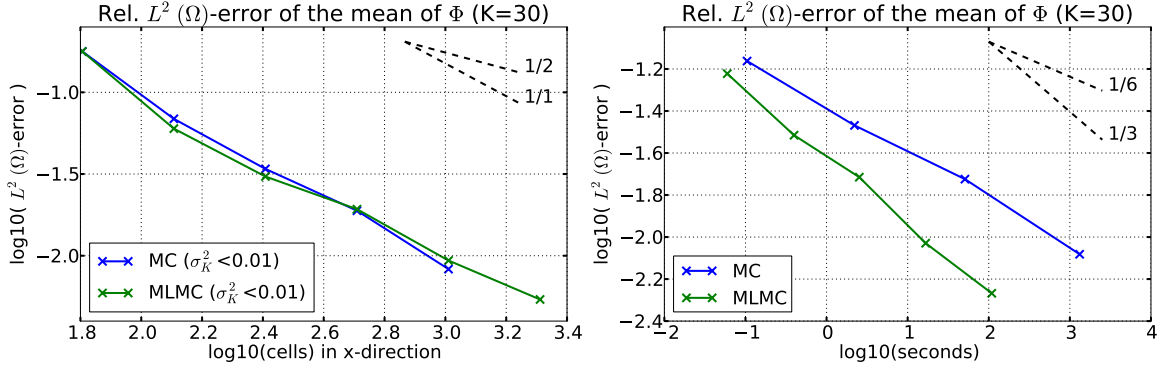
FIGURE 7. Convergence of the relative percentage error $\varepsilon_L^{\mathbb{E}}$ of the mean $\mathbb{E}_{\mu_t}(\Phi)$ of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with uniformly distributed Y_m.
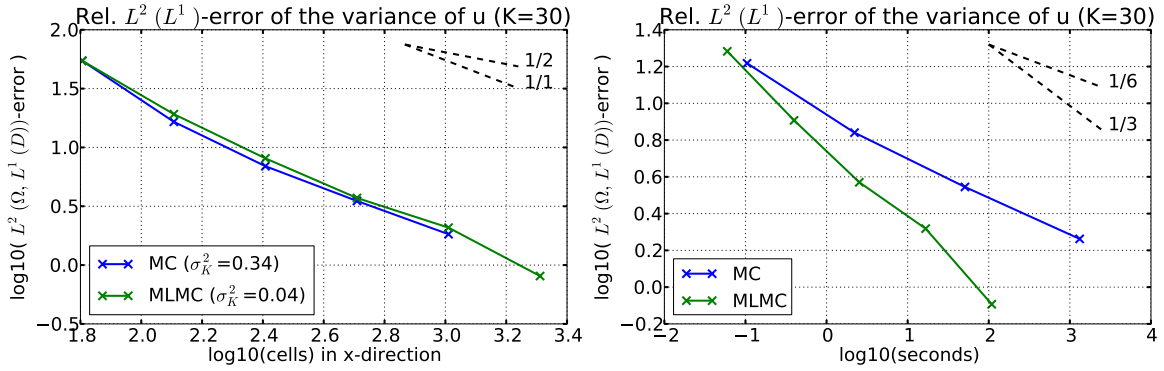

FIGURE 8. Convergence of the relative percentage error $\varepsilon_L^{\mathbb{V}}$ of the variance $\mathbb{V}_{\mu_t}(\Phi)$ of the viscous Burgers' equation (8.1) with kinematic viscosity $\nu = 10^{-3}$ and random initial data u_0 with uniformly distributed Y_m.

Under the (physically meaningful) assumption of a statistical solution with *finite mean kinetic energy*, a convergence rate of $1/2$ in $L^\infty((0,T); L^2(D))$, and in quadratic mean over the ensemble of all velocity fields is proved. The constants in the error bound only depend on the mean kinetic energy of the velocity ensemble constituting the statistical solution, but are independent of the kinematic viscosity. Our analysis delineates conditions to be satisfied by space and time discretizations and turbulence models in numerical flow solvers for the numerical approximation of individual (Leray–Hopf) solutions in order to ensure robust simulations. Besides the numerical analysis of approximations of statistical solutions ($\mu_t, t \in J$) by Monte Carlo sampling and multilevel Monte Carlo discretization, the presently proposed multilevel Monte Carlo approximation for generalized moments of these solutions offers the perspective of increased computational stability in capturing ensemble averages and so-called bulk properties of viscous, incompressible flows. We established the convergence estimate $CM^{-1/2}$ for the mean-square error of sample-averages over M Monte Carlo samples, with the constant $C > 0$ being independent of the Reynolds number. We also presented an analysis of the impact of discretization errors in space-, time- and for the initial probability measure

μ_0 on the accuracy of the Monte Carlo sample average. Consistency conditions on these discretizations to preserve the Monte Carlo convergence estimate, were identified and shown to hold in the particular case of the scalar, viscous Burgers' equation, with periodic boundary conditions. This analysis of discretization errors is then used to derive a multilevel Monte Carlo strategy for the cost-efficient numerical estimation of statistical solutions; in particular, of a selection of sample numbers M_ℓ in dependence on the level of discretization in the spacial and temporal variables. The rather large number M_ℓ of Monte Carlo samples on coarse space and time grids naturally allows for massively parallel computation strategies and, unlike singlelevel, direct numerical simulations, has some degree of robustness towards hardware failure, i.e. of 'fault tolerance', built in. We refer to [31] for details on fault tolerance of multilevel Monte Carlo algorithms. The mathematical analysis of using turbulence models in the multilevel Monte Carlo sampling for simulations of individual solutions on underresolved space and time discretization levels is the topic of ongoing research. For some indications on this, we refer to [3] where the combination of (numerical upscaling) with multilevel Monte Carlo simulations for random elliptic homogenization problems is analyzed. Parts of the analysis in the present work require uniqueness and a bounded domain D which, essentially, confines the results to space dimension $d = 2$. The approximation of *homogeneous* statistical solutions which are considered in dimension $d = 2, 3$, e.g. in [10, 9], will be dealt with elsewhere.

References

[1] *ALSVID-UQ*. Version 3.0, http://www.sam.math.ethz.ch/alsvid-uq.

[2] *Rosa (Cray XE6)*. Swiss National Supercomputing Center (CSCS), Lugano, www.cscs.ch.

[3] A. Abdulle, A. Barth, and C. Schwab, *Multilevel monte carlo methods for stochastic elliptic multiscale pdes*, Multiscale Model. Simul., 11 (2013), pp. 1033–1070.

[4] A. Barth and A. Lang, *Multilevel Monte Carlo method with applications to stochastic partial differential equations*, Int. J. Comput. Math., 89 (2012), pp. 2479–2498.

[5] ———, *Simulation of stochastic partial differential equations using finite element methods*, Stochastics, 84 (2012), pp. 217–231.

[6] A. Barth, A. Lang, and C. Schwab, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT, 53 (2013), pp. 3–27.

[7] A. Barth, C. Schwab, and N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.

[8] A. Basson, *Homogeneous statistical solutions and local energy inequality for 3D Navier-Stokes equations*, Communications in Mathematical Physics, 266 (2006), pp. 17–35.

[9] A. Basson, *Homogeneous statistical solutions and local energy inequality for 3D Navier-Stokes equations*, Comm. Math. Phys., 266 (2006), pp. 17–35.

[10] D. Chae and C. Foias, *On the homogeneous statistical solutions of the* 2-*D Navier-Stokes equations*, Indiana Univ. Math. J., 43 (1994), pp. 177–185.

[11] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, vol. 44 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.

[12] C. Foias, *Statistical study of Navier-Stokes equations. I, II*, Rend. Sem. Mat. Univ. Padova, 48 (1972), pp. 219–348 (1973); ibid. 49 (1973), 9–123.

[13] C. Foias, L. Hoang, and B. Nicolaenko, *On the helicity in 3D-periodic Navier-Stokes equations. II. The statistical case*, Comm. Math. Phys., 290 (2009), pp. 679–717.

[14] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes equations and turbulence*, vol. 83 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2001.

[15] C. Foias and G. Prodi, *Sur les solutions statistiques des équations de Navier-Stokes*, Ann. Mat. Pura Appl. (4), 111 (1976), pp. 307–330.

[16] C. Foiaş, R. M. S. Rosa, and R. Temam, *A note on statistical solutions of the three-dimensional Navier-Stokes equations: the stationary case*, C. R. Math. Acad. Sci. Paris, 348 (2010), pp. 347–353. References in text to extended version ArXiV 1111.6257v2.

[17] ———, *Properties of time-dependent statistical solutions of the three-dimensional Navier-Stokes equations*, Annales de l'Institute Fourier (to appear), (2013). ArXiV 1111.6257v2.

[18] C. Foiaş and R. Temam, *Homogeneous statistical solutions of Navier-Stokes equations*, Indiana Univ. Math. J., 29 (1980), pp. 913–957.

[19] U. Frisch and J. Bec, *"Burgulence"*, in Turbulence: nouveaux aspects/New trends in turbulence (Les Houches, 2000), EDP Sci., Les Ulis, 2001, pp. 341–383.

[20] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier-Stokes problem. I. Regularity of solutions and second-order error estimates for spatial discretization*, SIAM J. Numer. Anal., 19 (1982), pp. 275–311.

[21] ———, *Finite element approximation of the nonstationary Navier-Stokes problem. II. Stability of solutions and error estimates uniform in time*, SIAM J. Numer. Anal., 23 (1986), pp. 750–777.

[22] ———, *Finite element approximation of the nonstationary Navier-Stokes problem. III. Smoothing property and higher order error estimates for spatial discretization*, SIAM J. Numer. Anal., 25 (1988), pp. 489–512.

[23] ———, *Finite-element approximation of the nonstationary Navier-Stokes problem. IV. Error analysis for second-order time discretization*, SIAM J. Numer. Anal., 27 (1990), pp. 353–384.

[24] H. Holden and N. H. Risebro, *Front tracking for hyperbolic conservation laws*, vol. 152 of Applied Mathematical Sciences, Springer, New York, 2011. First softcover corrected printing of the 2002 original.

[25] K. H. Karlsen, U. Koley, and N. H. Risebro, *An error estimate for the finite difference approximation to degenerate convection-diffusion equations*, Numer. Math., 121 (2012), pp. 367–395.

[26] M. Lesieur, *Turbulence in fluids*, vol. 40 of Fluid Mechanics and its Applications, Kluwer Academic Publishers Group, Dordrecht, third ed., 1997.

[27] R. LeVeque, *Numerical solution of hyperbolic conservation laws*, Cambridge U. Press, 2005.

[28] V. Maz'ya and J. Rossmann, *Mixed boundary value problems for the stationary Navier-Stokes system in polyhedral domains*, Arch. Ration. Mech. Anal., 194 (2009), pp. 669–712.

[29] S. Mishra and C. Schwab, *Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data*, Math. Comp., 81 (2012), pp. 1979–2018.

[30] S. Mishra, C. Schwab, and J. Šukys, *Multi-level monte carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions*, J. Comp. Phys., 231(8), pp. 3365–3388, (2012).

[31] S. Pauli, P. Arbenz, and C. Schwab, *Intrinsic fault tolerance of multi level monte carlo methods*, Tech. Report 2012-24, Seminar for Applied Mathematics, ETH Zürich, 2012.

[32] J. Shen, *On error estimates of projection methods for Navier-Stokes equations: first-order schemes*, SIAM J. Numer. Anal., 29 (1992), pp. 57–77.

[33] ———, *Projection methods for time-dependent Navier-Stokes equations*, Appl. Math. Lett., 5 (1992), pp. 35–37.

[34] J. Šukys, S. Mishra, and C. Schwab, *Static load balancing for multi-level Monte Carlo finite volume solvers*, PPAM 2011, Part I, LNCS 7203, pp. 245–254. Springer, Heidelberg, (2012).

[35] R. Temam, *Navier-Stokes equations and nonlinear functional analysis*, vol. 41 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1983.

[36] M. I. Višik and A. V. Foursikov, *L'équation de Hopf, les solutions statistiques, les moments correspondants aux systèmes des équations paraboliques quasilinéaires*, J. Math. Pures Appl. (9), 56 (1977), pp. 85–122.

[37] ———, *Solutions statistiques homogènes des systèmes differentiels paraboliques et du système de Navier-Stokes*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 4 (1977), pp. 531–576.

(Andrea Barth)
ETH, Seminar für Angewandte Mathematik
Rämistrasse 101
8092 Zürich
    E-mail address: `andrea.barth@math.ethz.ch`

(Christoph Schwab)
ETH, Seminar für Angewandte Mathematik
Rämistrasse 101
8092 Zürich
    E-mail address: `schwab@math.ethz.ch`

(Jonas Šukys)
ETH, Seminar für Angewandte Mathematik
Rämistrasse 101
8092 Zürich
    E-mail address: `jonas.sukys@sam.math.ethz.ch`

Recent Research Reports

Nr.	Authors/Title

2013-23   P. Grohs and M. Sprecher
Projection-based Quasiinterpolation in Manifolds

2013-24   P. Grohs and S. Keiper and G. Kutyniok and M. Schaefer
α-Molecules: Curvelets, Shearlets, Ridgelets, and Beyond

2013-25   A. Cohen and A. Chkifa and Ch. Schwab
Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs

2013-26   A. Lang
Isotropic Gaussian random fields on the sphere

2013-27   Ch. Schwab and C. Schillings
Sparse Quadrature Approach to Bayesian Inverse Problems

2013-28   V. Kazeev and I. Oseledets
The tensor structure of a class of adaptive algebraic wavelet transforms

2013-29   J. Dick and F.Y. Kuo and Q.T. Le Gia and D. Nuyens and Ch. Schwab
Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs

2013-30   R. Hiptmair and A. Paganini and S. Sargheini
Comparison of Approximate Shape Gradients

2013-31   R. Hiptmair and A. Moiola and I. Perugia
Plane Wave Discontinuous Galerkin Methods: Exponential Convergence of the hp-version

2013-32   U. Koley and N. Risebro and Ch. Schwab and F. Weber
Multilevel Monte Carlo for random degenerate scalar convection diffusion equation