Regular Article - Theoretical Physics

SMD-based numerical stochastic perturbation theory

Mattia Dalla Brida[1,a], **Martin Lüscher**[2,3,b]

[1] Dipartimento di Fisica, Università di Milano-Bicocca and INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milan, Italy
[2] Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland
[3] AEC, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

Abstract The viability of a variant of numerical stochastic perturbation theory, where the Langevin equation is replaced by the SMD algorithm, is examined. In particular, the convergence of the process to a unique stationary state is rigorously established and the use of higher-order symplectic integration schemes is shown to be highly profitable in this context. For illustration, the gradient-flow coupling in finite volume with Schrödinger functional boundary conditions is computed to two-loop (i.e. NNL) order in the SU(3) gauge theory. The scaling behaviour of the algorithm turns out to be rather favourable in this case, which allows the computations to be driven close to the continuum limit.

1 Introduction

Numerical stochastic perturbation theory (NSPT) [1–3] is a powerful tool that allows many interesting calculations in QCD and other quantum field theories to be performed to high order in the interactions. For technical reasons, the computations proceed in the framework of lattice field theory, but results for renormalized quantities in the continuum theory can then be obtained through an extrapolation to vanishing lattice spacing. NSPT can be highly automated and the application of the method in finite volume and to correlation functions of complicated composite fields gives rise to hardly any additional difficulties.

Reliable extrapolations to the continuum limit require accurate data at several lattice spacings in the scaling region. NSPT calculations can therefore rapidly become large-scale projects, where computational efficiency is all-important. Traditionally, NSPT is based on the Langevin equation, but the success of the HMC algorithm [4] in lattice QCD suggests that the inclusion of a molecular-dynamics update step in the underlying stochastic process might be beneficial. Smaller

autocorrelation times and an improved scaling behaviour towards the continuum limit could perhaps be achieved in this way. Moreover, through the use of highly efficient symplectic integration schemes, the systematic errors deriving from the discretization of the simulation time may conceivably be reduced.

NSPT based on the SMD (stochastic molecular dynamics, or generalized HMC) algorithm [5–7] has recently been briefly looked at in Ref. [8] and was found to perform well. Here we establish the convergence of the algorithm to a unique stationary state and study its efficiency in the case of the gradient-flow coupling in the SU(3) gauge theory. Various technical problems are addressed along the way, among them the modifications required to ensure that the stochastic process does not run away in the gauge directions.

2 Stochastic molecular dynamics

In order to bring out the basic structure of the SMD-variant of NSPT most clearly, a generic system described by a set $q = (q_1, \ldots, q_n)$ of real coordinates and an action $S(q)$ is considered in this and the following two sections.

2.1 Preliminaries

The action $S(q)$ is assumed to be differentiable and to have an expansion in powers of a coupling g of the form

$$S(q) = \sum_{r=0}^{\infty} g^r S_r(q), \tag{2.1}$$

where $S_r(q)$ is a polynomial in q of degree $d_r \geq 2$. Moreover, it is taken for granted that the leading-order term

$$S_0(q) = \frac{1}{2}(q, \Delta q) = \frac{1}{2} \sum_{k,l=1}^{n} q_k \Delta_{kl} q_l \tag{2.2}$$

a e-mail: mattia.dalla.brida@desy.de
b e-mail: Martin.Luescher@cern.ch

is a strictly positive quadratic form in the coordinates.

The observables $\mathcal{O}(q)$ of interest are assumed to be similarly expandable and their expectation values

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}_S} \int dq_1 \ldots dq_n \, \mathcal{O}(q) e^{-S(q)} \qquad (2.3)$$

then have a well-defined perturbation expansion with coefficients given by Feynman diagrams as usual.

2.2 SMD algorithm

The SMD algorithm operates in the phase space of the theory and thus updates both the coordinates q and their canonical momenta $p = (p_1, \ldots, p_n)$. An SMD update cycle consists of a momentum rotation followed by a molecular-dynamics evolution and, optionally, an acceptance–rejection step.

The momenta are rotated in a random direction according to

$$p \rightarrow c_1 p + c_2 \upsilon, \qquad (2.4)$$

where the momentum υ is randomly chosen from a Gaussian distribution with mean zero and unit variance. The coefficients

$$c_1 = e^{-\gamma \epsilon}, \quad c_2 = (1 - c_1^2)^{1/2}, \qquad (2.5)$$

depend on the simulation time step $\epsilon > 0$ and a parameter $\gamma > 0$ that controls the rotation angle.

In the second step, the molecular-dynamics equations

$$\partial_t p = -\nabla S(q), \qquad \partial_t q = p, \qquad (2.6)$$

are integrated from the current simulation time t to $t + \epsilon$ using a reversible symplectic integration scheme (see Sect. 2.3). The algorithm (momentum rotation followed by the molecular-dynamics evolution) simulates the canonical distribution

$$\frac{1}{\mathcal{Z}_H} e^{-H(p,q)}, \qquad H(p, q) = \frac{1}{2}(p, p) + S(q), \qquad (2.7)$$

provided the integration errors are negligible or an acceptance–rejection step is included in the update cycle which corrects for these [9]. Stochastic estimates of the expectation values (2.3) of the observables of interest are then obtained by averaging their values over a range of simulation time.

2.3 Integration schemes

The molecular-dynamics equations may be integrated by applying a sequence of the elementary steps

$$I_{p,h} : \ p \rightarrow p - h\nabla S(q), \qquad (2.8)$$
$$I_{q,h} : \ q \rightarrow q + hp. \qquad (2.9)$$

to the current momenta and coordinates, with step sizes h proportional to ϵ. A well-known example is the leapfrog integrator $I_{p,\epsilon/2} I_{q,\epsilon} I_{p,\epsilon/2}$, and several highly efficient schemes are described in Ref. [10].

Integrators I_ϵ of this kind are symplectic and they can be (and are here) required to be reversible, i.e. to be such that

$$I_\epsilon P I_\epsilon = P, \qquad (2.10)$$

where P stands for the momentum reflection $p \rightarrow -p$.

3 Stochastic perturbation theory

Stochastic perturbation theory [11,12] is usually derived from the Langevin equation by expanding the stochastic variables and the driving forces in powers of the coupling. In this section, another (although probably closely related) form of stochastic perturbation theory is discussed, which is obtained by expanding the SMD algorithm in the same way.

3.1 SMD algorithm at weak coupling

Since the acceptance–rejection step is not smooth in the coupling, its effects would be difficult to take into account in perturbation theory. In the following, the acceptance–rejection step is therefore omitted, without further notice, and one is thus left with an algorithm that simulates the system only up to integration errors.

The histories $p(t)$, $q(t)$ of the momenta and coordinates generated by the SMD algorithm depend on the coupling g through the force term in the integration step (2.8). In particular, they are smooth functions of the coupling and may consequently be expanded in the asymptotic series

$$p(t) = \sum_{r=0}^{\infty} g^r \hat{p}_r(t), \qquad q(t) = \sum_{r=0}^{\infty} g^r \hat{q}_r(t), \qquad (3.1)$$

where the leading-order histories $\hat{p}_0(t)$, $\hat{q}_0(t)$ coincide with the ones generated by the algorithm in the free theory with action (2.2).

In terms of the coefficients $\hat{p}_r$, $\hat{q}_r$, the momentum rotation (2.4) becomes

$$\hat{p}_r \rightarrow c_1 \hat{p}_r + \delta_{r0} c_2 \upsilon \qquad (3.2)$$

and the molecular-dynamics integration steps (2.8), (2.9) assume the form

$$I_{\hat{p}_r,h} : \ \hat{p}_r \rightarrow \hat{p}_r - h\hat{F}_r(\hat{q}_0, \ldots, \hat{q}_r), \qquad (3.3)$$
$$I_{\hat{q}_r,h} : \ \hat{q}_r \rightarrow \hat{q}_r + h\hat{p}_r. \qquad (3.4)$$

The forces $\hat{F}_r$ in Eq. (3.3) are given by

$$\nabla S(q) = \sum_{r=0}^{\infty} g^r \hat{F}_r(\hat{q}_0, \ldots, \hat{q}_r) \qquad (3.5)$$

and it is understood that all momenta and all coordinates are updated alternately so that the variables on the right of Eqs. (3.3), (3.4) are always the current ones. For any given initial data, these rules completely determine the histories $\hat{p}_r(t), \hat{q}_r(t)$.

3.2 Perturbation expansion of observables

Similarly to the gradient of the action, any observable

$$\mathcal{O}(q) = \sum_{r=0}^{\infty} g^r \hat{\mathcal{O}}_r(\hat{q}_0, \dots, \hat{q}_r) \tag{3.6}$$

may be expanded in powers of the coupling. The coefficients $k_r(\mathcal{O})$ in the perturbation expansion

$$\langle \mathcal{O} \rangle = \sum_{r=0}^{\infty} k_r(\mathcal{O}) g^r \tag{3.7}$$

of its expectation value (2.3) then coincide with the averages of $\hat{\mathcal{O}}_r(\hat{q}_0(t), \dots, \hat{q}_r(t))$ over the simulation time t up to statistical (and integration) errors.

4 Convergence to a stationary state

Stochastic processes can run away or do not converge to a stationary distribution for other reasons. In the case of the stochastic perturbation theory described in Sect. 3, the asymptotic stationarity of the underlying process can be rigorously shown if the simulation step size ϵ is sufficiently small. The range of step sizes, where convergence is guaranteed, depends on the chosen integration scheme for the molecular-dynamics equations and the matrix Δ in the leading-order part (2.2) of the action.

4.1 Molecular-dynamics evolution in the free theory

If the coupling g is turned off, the molecular-dynamics equations become linear and their (approximate) integration from time t to $t + \epsilon$ amounts to a linear transformation

$$\begin{pmatrix} p \\ q \end{pmatrix} \rightarrow M \begin{pmatrix} p \\ q \end{pmatrix} \tag{4.1}$$

of the current momenta and coordinates. The $2n \times 2n$ matrix M in this equation has a block structure,

$$M = \begin{pmatrix} M_{pp} & M_{pq} \\ M_{qp} & M_{qq} \end{pmatrix}, \tag{4.2}$$

with $n \times n$ blocks that are polynomials in ϵ and the matrix Δ with some numerical coefficients. In particular, they are commuting real symmetric matrices. In the case of the leapfrog integrator,

$$M_{pp} = M_{qq} = 1 - \frac{1}{2}\epsilon^2 \Delta, \quad M_{pq} = -\epsilon \Delta \left(1 - \frac{1}{4}\epsilon^2 \Delta\right),$$
$$M_{qp} = \epsilon, \tag{4.3}$$

an explicit expression for the blocks can be obtained for other popular integrators as well (see Appendix A).

The symplecticity and reversibility of the chosen integration scheme imply

$$M_{pp} = M_{qq}, \tag{4.4}$$
$$M_{pp}M_{qq} - M_{pq}M_{qp} = 1. \tag{4.5}$$

It follows from these relations that the Hamilton function

$$\hat{H}(p, q) = \frac{1}{2}(p, p) + \frac{1}{2}(q, \hat{\Delta} q), \tag{4.6}$$
$$\hat{\Delta} = -M_{pq}(M_{qp})^{-1} = \Delta + \mathrm{O}(\epsilon), \tag{4.7}$$

is exactly conserved by the integrator. In the following, ϵ and Δ are assumed to be such that M_{qp} is non-singular and $\hat{\Delta}$ positive definite. Both conditions are met in the case of the leapfrog integrator if $\epsilon^2 \|\Delta\| < 4$, where $\|\Delta\|$ is the largest eigenvalue of Δ. The probability distribution

$$\hat{P}(p, q) \propto \mathrm{e}^{-\hat{H}(p,q)} \tag{4.8}$$

is then well-behaved and preserved by the SMD algorithm, since it is preserved both by the momentum rotation and the molecular-dynamics evolution.

4.2 Convergence of the leading-order process

For any initial distribution of the momenta and coordinates, the SMD algorithm produces a sequence of distributions, which converges to the stationary distribution (4.8) in the free theory. One can show this by working out the action of an SMD update cycle on a given distribution, but the convergence of the algorithm may be established more easily starting from the identity

$$\begin{pmatrix} p(t) \\ q(t) \end{pmatrix} = \tilde{M}^{t/\epsilon} \begin{pmatrix} p(0) \\ q(0) \end{pmatrix} + c_2 \sum_{u=0}^{t-\epsilon} \tilde{M}^{(t-u)/\epsilon-1} M \begin{pmatrix} \upsilon(u) \\ 0 \end{pmatrix}, \tag{4.9}$$

where $\upsilon(u)$ is the momentum chosen randomly in the momentum rotation (2.4) at simulation time $u = 0, \epsilon, 2\epsilon, \dots$ and

$$\tilde{M} = M \begin{pmatrix} c_1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.10}$$

The long-time behaviour of the momenta and coordinates thus depends on the properties of the matrix $\tilde{M}$.

The blocks M_{pp}, M_{pq}, M_{qp} and $\hat{\Delta}$ are commuting real symmetric matrices. Since

$$M_{pp}^2 = 1 - M_{qp}\hat{\Delta}M_{qp} \tag{4.11}$$

and since $\hat{\Delta}$ is assumed to be positive definite, the eigenvalues of M_{pp} have magnitude strictly less than 1. The eigenvalues of $\tilde{M}$ are then

$$\lambda_{\pm} = \frac{1}{2}\left\{(1+c_1)\mu \pm \sqrt{(1+c_1)^2\mu^2 - 4c_1}\right\}, \qquad (4.12)$$

where μ runs through the eigenvalues of M_{pp}. In particular, $|\lambda_{\pm}| < 1$ and $\tilde{M}$ is thus a contraction matrix.

The first term on the right of Eq. (4.9) consequently dies away exponentially with increasing simulation time t. Since the random momenta are normally distributed, the momenta and coordinates at large t are then normally distributed as well, with mean zero and variances equal to their two-point autocorrelation functions. In the large-time limit, the qq autocorrelation function, for example, is given by

$$\langle q(t)q(s)\rangle_{\upsilon} \underset{t \geq s}{=} \left\{\tilde{M}^{(t-s)/\epsilon}K\right\}_{qq}, \qquad (4.13)$$

$$K = c_2^2 \sum_{u=0}^{\infty} \tilde{M}^{u/\epsilon}MP_+\left\{\tilde{M}^{u/\epsilon}M\right\}^T, \qquad P_+ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (4.14)$$

and the other two-point functions by the pp, pq and qp blocks of the matrix on the right of Eq. (4.13). The kernel K satisfies

$$\tilde{M}K\tilde{M}^T = K - c_2^2 MP_+M^T, \qquad (4.15)$$

i.e. an inhomogeneous linear equation that has a unique solution, since $\tilde{M}$ is a contraction matrix. A few lines of algebra then show that

$$K = \begin{pmatrix} 1 & 0 \\ 0 & \hat{\Delta}^{-1} \end{pmatrix} \qquad (4.16)$$

solves the equation. In particular, the variances of the momenta and coordinates at any fixed simulation time coincide with the ones of the stationary distribution (4.8), which proves that the latter coincides with the distribution simulated by the SMD algorithm.

4.3 Convergence beyond the leading order

The assumed structure of the action $S(q)$ implies that the force in the molecular-dynamics integration step (3.3) is of the form

$$\hat{F}_r(\hat{q}_0, \ldots, \hat{q}_r) = \Delta\hat{q}_r + \hat{F}'_r(\hat{q}_0, \ldots, \hat{q}_{r-1}), \qquad (4.17)$$

where the second term on the right is a polynomial in the coordinates up to order $r-1$. If the history of the latter and the associated momenta is already known (including their values at the intermediate integration steps), the integration of the molecular-dynamics equations at order r thus amounts to solving an inhomogeneous linear recursion. A moment of thought then reveals that

$$\begin{pmatrix} \hat{p}_r(t) \\ \hat{q}_r(t) \end{pmatrix} = \tilde{M}^{t/\epsilon}\begin{pmatrix} \hat{p}_r(0) \\ \hat{q}_r(0) \end{pmatrix}$$
$$+ \sum_{u=\epsilon}^{t} \tilde{M}^{(t-u)/\epsilon}\hat{V}_r(\hat{p}_0(u), \ldots, \hat{q}_{r-1}(u)) \qquad (4.18)$$

for all $r \geq 1$. The "vertices" $\hat{V}_r(\hat{p}_0, \hat{q}_0; \ldots; \hat{p}_{r-1}, \hat{q}_{r-1})$ in this formula are polynomials in their arguments, whose exact form depends on both the integration scheme and the forces (4.17).

Recalling Eq. (4.9) and the fact that $\tilde{M}$ is a contraction matrix, the convergence of the autocorrelation functions of the momenta and coordinates at large times t may now be shown recursively from order 0 to any finite order r. Equation (4.18) actually allows the highest-order variables in any correlation function to be expressed through lower-order ones up to an exponentially decaying contribution. Clearly, in the large-time limit, the autocorrelation functions do not depend on the initial distribution of the variables and are stationary, i.e. invariant under time translations.

The expectation values of the coefficients in the perturbation expansion (3.6) of the observables coincide with a sum of autocorrelation functions of the coordinates at equal times. Their convergence at large times is therefore guaranteed as well.

4.4 Summary

The discussion in this section shows that the SMD algorithm converges to all orders of perturbation theory if the matrix Δ is strictly positive and if $\epsilon^2\|\Delta\| < \kappa$, where κ depends on the molecular-dynamics integrator. In the case of the leapfrog, the second-order OMF and the fourth-order OMF integrators, κ is equal to 4, 6.51 and 9.87, respectively (cf. Appendix A).

5 Stochastic perturbation theory in lattice QCD

With respect to the generic system considered so far, the situation in lattice QCD is complicated by the gauge symmetry and the quark fields. In this section, stochastic perturbation theory is first set up for the pure SU(N) gauge theory. The modifications required for the damping of the gauge modes are then discussed and the section ends with a brief description of how the quarks can be included in the simulations.

5.1 Lattice fields

The lattice theory is set up on a $T \times L^3$ lattice with periodic boundary conditions in the space directions and Schrödinger functional (SF) [13,14] or open-SF [15] boundary conditions in the time direction. In both cases the link variables $U(x, \mu)$ satisfy

$$U(x, k) = 1, \quad k = 1, 2, 3, \tag{5.1}$$

at time $x_0 = T$ and additionally at $x_0 = 0$ if SF boundary conditions are chosen. The notation used in the following coincides with the one previously employed in Ref. [15]. In particular, all dimensionful quantities are expressed in units of the lattice spacing.

The momenta $\pi(x, \mu)$ of the link variables $U(x, \mu)$ take values in the Lie algebra $\mathfrak{su}(N)$ of the gauge group. They vanish at the boundaries of the lattice, where the link variables are frozen to unity. The scalar product of any two momentum fields,

$$(\pi, \upsilon) = \sum_{x,\mu} w_{x,\mu}^{-1} \pi^a(x, \mu) \upsilon^a(x, \mu), \tag{5.2}$$

includes a conventional weight factor

$$w_{x,u} = \begin{cases} 2 & \text{if } \mu > 0 \text{ and } x_0 = 0 \text{ or } x_0 = T \\ 1 & \text{otherwise} \end{cases} \tag{5.3}$$

which will reappear below in various expressions.

Infinitesimal gauge transformations are fields $\omega(x)$ of $\mathfrak{su}(N)$ elements defined on the sites x of the lattice. They vanish at $x_0 = T$ and must, furthermore, be constant at $x_0 = 0$ if SF boundary conditions are imposed. In the latter case, the fields may be split in two parts according to

$$\omega(x) = (1 - x_0/T) \omega(0) + \hat{\omega}(x), \tag{5.4}$$

where $\hat{\omega}(x)$ vanishes at $x_0 = 0, T$ and is otherwise unconstrained. A possible choice of scalar product is then

$$(\omega, v) = TL^3 \omega^a(0) v^a(0) + \sum_x \hat{\omega}^a(x) \hat{v}^a(x), \tag{5.5}$$

while in the case of open-SF boundary conditions

$$(\omega, v) = \sum_x w_x^{-1} \omega^a(x) v^a(x) \tag{5.6}$$

with $w_x = 2$ at $x_0 = 0, T$ and $w_x = 1$ elsewhere.

Independently of the boundary conditions imposed on the gauge field, the quark fields are required to vanish at both $x_0 = 0$ and $x_0 = T$. No weight factor is included in the scalar product of these fields.

5.2 Basic stochastic process

In the absence of quark fields, the SMD algorithm proceeds along the lines of Sect. 2. For the action $S(U)$ of the gauge field any of the frequently used ones may be taken. The Hamilton function from which the SMD algorithm derives is then given by

$$H(\pi, U) = \frac{1}{2}(\pi, \pi) + S(U). \tag{5.7}$$

In particular, the random field in the momentum rotation

$$\pi(x, \mu) \to c_1 \pi(x, \mu) + c_2 \upsilon(x, \mu) \tag{5.8}$$

must be distributed with probability density proportional to $e^{-\frac{1}{2}(\upsilon, \upsilon)}$.

If the weight factor $w_{x,\mu}$ in the scalar product (5.2) is also included in the symplectic structure (i.e. the Poisson bracket), the molecular-dynamics equations assume the form

$$\partial_t \pi(x, \mu) = -g_0 w_{x,\mu} (\partial_{x,\mu}^a S)(U) T^a,$$
$$\partial_t U(x, \mu) = g_0 \pi(x, \mu) U(x, \mu). \tag{5.9}$$

For later convenience, the expressions on the right of the equations have been scaled by the bare gauge coupling g_0, an operation that could be undone by rescaling the simulation time. The integration steps (2.8) and (2.9) are then given by

$$I_{\pi,h} : \pi(x, \mu) \to \pi(x, \mu) - hg_0 w_{x,\mu} (\partial_{x,\mu}^a S)(U) T^a, \tag{5.10}$$

$$I_{U,h} : U(x, \mu) \to e^{hg_0 \pi(x,\mu)} U(x, \mu). \tag{5.11}$$

Since it will be omitted in perturbation theory, the acceptance–rejection step needed to correct for the integration errors is not described here.

5.3 Perturbation expansion

In perturbation theory, the link variables are represented by a gauge potential $A_\mu(x)$ through

$$U(x, \mu) = \exp\{g_0 A_\mu(x)\} = 1 + \sum_{r=0}^{\infty} g_0^{r+1} \hat{U}_r(x, \mu),$$
$$\hat{U}_0(x, \mu) = A_\mu(x). \tag{5.12}$$

The gauge potential takes values in $\mathfrak{su}(N)$ and satisfies the same boundary conditions as the momentum field

$$\pi(x, \mu) = \sum_{r=0}^{\infty} g_0^r \hat{\pi}_r(x, \mu). \tag{5.13}$$

When the gauge and momentum fields are replaced by these expansions, the SMD algorithm leads to a hierarchy of stochastic equations as in the case of the generic system considered in Sect. 3.

At lowest order in the coupling, the momentum is rotated according to Eq. (5.8) (with $\hat{\pi}_0$ in place of π) and the integration steps (5.10), (5.11) become

$$I_{\hat{\pi}_0,h} : \hat{\pi}_0(x, \mu) \to \hat{\pi}_0(x, \mu) - h(\Delta \hat{U}_0)(x, \mu), \tag{5.14}$$

$$I_{\hat{U}_0,h} : \hat{U}_0(x, \mu) \to \hat{U}_0(x, \mu) + h\hat{\pi}_0(x, \mu), \tag{5.15}$$

where Δ is the symmetric linear operator in the leading-order expression

$$S_0(U) = \frac{1}{2}(A, \Delta A) \tag{5.16}$$

for the gauge action.

5.4 Damping of the gauge modes

The space $\mathcal{H}_1$ of gauge potentials may be split into the subspace $\mathcal{H}_1^L$ of gauge modes and its orthogonal complement $\mathcal{H}_1^T$. There is a one-to-one correspondence between the infinitesimal gauge transformations, introduced in Sect. 5.1, and the gauge modes through the forward difference operator

$$(d\nu)(x, \mu) = \partial_\mu \nu(x), \qquad (5.17)$$

which maps any field ν in the space $\mathcal{H}_0$ of infinitesimal gauge transformations to a field $d\nu \in \mathcal{H}_1^L$. The adjoint operator d^* goes in the opposite direction and is defined by the requirement that

$$(d\nu, A) = (\nu, d^*A) \qquad (5.18)$$

for all $\nu \in \mathcal{H}_0$ and $A \in \mathcal{H}_1$. In particular, the subspace $\mathcal{H}_1^T$ coincides with the space of gauge potentials A satisfying $d^*A = 0$ (d^* is given explicitly in Appendix B).

Since the gauge modes are annihilated by Δ, the $\mathcal{H}_1^L$-component of the leading-order gauge field $\hat{U}_0$ performs a random walk and thus slowly runs away in the course of a simulation. Stability can be regained by applying a gauge transformation

$$\pi(x, \mu) \to \Lambda(x)\pi(x, \mu)\Lambda(x)^{-1},$$
$$\Lambda(x) = \exp\{\epsilon g_0 \omega(x)\}, \qquad (5.19)$$
$$U(x, \mu) \to \Lambda(x)U(x, \mu)\Lambda(x + \hat{\mu})^{-1}, \qquad (5.20)$$

to the fields after each SMD update cycle, where $\omega \in \mathcal{H}_0$ is a new field that is evolved together with the other fields. There are two update steps for this field, the first,

$$\omega(x) \to c_1 \omega(x), \qquad (5.21)$$

being applied together with the momentum rotation and the second,

$$\omega(x) \to \omega(x) + \epsilon \lambda_0 (d^*C)(x), \qquad (5.22)$$

$$C_\mu(x) = \frac{1}{2g_0}\left\{ U(x, \mu) - U(x, \mu)^{-1} - \frac{1}{N}\mathrm{tr}\left[U(x, \mu) - U(x, \mu)^{-1} \right] \right\}, \qquad (5.23)$$

at the end of the molecular-dynamics evolution. The parameter $\lambda_0 > 0$ controls the feedback from the current gauge field to the gauge-damping field $\omega(x)$ and can, in principle, be set to any value. Clearly, the history of gauge-invariant observables is not affected by these modifications of the SMD algorithm, but as discussed below (in Sect. 5.5), they have the desired damping effect on the gauge modes.[1]

---

[1] In the continuous-time limit $\epsilon \to 0$, the time-dependence of the gauge-damping field is governed by a first-order differential equation and the modified algorithm integrates a form of the stochastic molecular-

Like the other fields, the gauge-damping field is expanded in a series

$$\omega(x) = \sum_{r=0}^{\infty} g_0^r \hat{\omega}_r(x) \qquad (5.24)$$

in perturbation theory. At leading order, the new update steps are then

$$\hat{\omega}_0(x) \to c_1 \hat{\omega}_0(x), \qquad (5.25)$$
$$\hat{\omega}_0(x) \to \hat{\omega}_0(x) + \epsilon \lambda_0 (d^*\hat{U}_0)(x), \qquad (5.26)$$
$$\hat{U}_0(x, \mu) \to \hat{U}_0(x, \mu) - \epsilon(d\hat{\omega}_0)(x, \mu), \qquad (5.27)$$

and the higher-order rules have the familiar hierarchical structure.

5.5 Long-time stationarity of the process

Although the stochastic process now includes variables without associated momenta, the discussion in Sect. 4 applies here too with only minor changes. In particular, the convergence of the process to a unique stationary state is guaranteed to all orders of the coupling, if the matrix $\tilde{M}$ describing the evolution of the fields $(\hat{\pi}_0, \hat{U}_0, \hat{\omega}_0)$ at leading order is a contraction matrix.

Since the gauge modes are zero modes of Δ, the subspace of field vectors

$$(\hat{\pi}_0, \hat{U}_0, \hat{\omega}_0) = (d\hat{v}_0, d\hat{\sigma}_0, \hat{\omega}_0), \qquad \hat{v}_0, \hat{\sigma}_0 \in \mathcal{H}_0, \qquad (5.28)$$

is invariant under the action of $\tilde{M}$. When restricted to its orthogonal complement (which is invariant too), the matrix describes the evolution of the $\mathcal{H}_1^T$-components of $\hat{\pi}_0, \hat{U}_0$. The chosen boundary conditions (SF or open-SF) imply the strict positivity of Δ in this subspace and $\tilde{M}$ is therefore a contraction matrix there if

$$\epsilon^2 \|\Delta\| < \kappa, \qquad (5.29)$$

where κ depends on the molecular-dynamics integrator (cf. Sect. 4.4).

In the subspace (5.28) of the gauge modes, on the other hand, the action of $\tilde{M}$ amounts to applying the linear transformation

$$\begin{pmatrix} \hat{v}_0 \\ \hat{\sigma}_0 \\ \hat{\omega}_0 \end{pmatrix} \to \begin{pmatrix} c_1 & 0 & 0 \\ \epsilon c_1 & 1 - \epsilon^2 \lambda_0 d^*d & -\epsilon c_1 \\ 0 & \epsilon \lambda_0 d^*d & c_1 \end{pmatrix} \begin{pmatrix} \hat{v}_0 \\ \hat{\sigma}_0 \\ \hat{\omega}_0 \end{pmatrix} \qquad (5.30)$$

to the fields. The associated eigenvalues are equal to c_1 or to

$$\frac{1}{2}\left\{ b \pm \sqrt{b^2 - 4c_1} \right\}, \qquad b = 1 + c_1 - \epsilon^2 \lambda_0 \mu^2, \qquad (5.31)$$

Footnote 1 continued

dynamics equations, which coincides with the standard one up to a time-dependent gauge transformation (see Appendix C).

where μ^2 is any eigenvalue of d^*d. This operator has no zero modes and some simple estimates then show that $\tilde{M}$ is a contraction matrix in the subspace of gauge modes, provided the bound

$$\epsilon^2 \lambda_0 \|d^*d\| < 2(1 + c_1) \tag{5.32}$$

holds.

Convergence of the stochastic process to equilibrium is thus guaranteed if the inequalities (5.29) and (5.32) are both satisfied. With the chosen boundary conditions,

$$\|\Delta\| \le 16k, \qquad \|d^*d\| \le 16, \tag{5.33}$$

where k is equal to 1, 5/3 and 3.648 in the case of the Wilson [16], tree-level Symanzik improved [17,18] and Iwasaki [19] gauge action, respectively. In practice, the constraints (5.29) and (5.32) therefore tend to be fairly mild and the main concern is to ensure that the integration errors are sufficiently small at the chosen values of ϵ.

5.6 Inclusion of the quark fields

As in the case of the HMC algorithm [4], the effects of the quarks can be included in SMD simulations by adding a multiplet of pseudo-fermion fields to the theory with the appropriate action. The details are not important here and it suffices to know that the action is a sum of terms, one per pseudo-fermion field ϕ, of the form

$$S_{\text{pf}}(U, \phi) = (R(U)\phi, R(U)\phi), \tag{5.34}$$

with $R(U)$ being some gauge-covariant linear operator (see Ref. [20], for example).

Only few modifications of the SMD update cycle are required in the presence of the pseudo-fermion fields. Similarly to the momentum field, each field is rotated according to[2]

$$\phi(x) \to c_1 \phi(x) + c_2 (R(U)^{-1}\eta)(x) \tag{5.35}$$

at the beginning of the cycle, where η is a Gaussian random field with mean zero and unit variance. The molecular-dynamics evolution of the gauge and momentum field then proceeds at fixed pseudo-fermion fields, with the contribution of the pseudo-fermion action to the driving force properly taken into account. Clearly, the gauge transformation (5.19), (5.20) applied at the end of the update cycle must be extended to the pseudo-fermion fields.

When the algorithm is expanded in powers of the coupling g_0, the renormalization of the quark masses should be taken into account so that the masses in the leading-order stochastic equations are the renormalized ones. The pseudo-fermion fields decouple from the other fields at lowest order

and are simulated exactly by the random rotation (5.35). Convergence of the stochastic process to a unique stationary state is then again guaranteed to all orders, if the bounds (5.29) and (5.32) are satisfied.

6 Computation of the gradient-flow coupling

The gradient-flow coupling in finite volume with SF boundary conditions has recently been used in step scaling studies of three-flavour QCD [22,23]. Such calculations serve to relate the low-energy properties of the theory to the high-energy regime, where contact with the standard QCD parameters and matrix elements can be made in perturbation theory [24] (see Ref. [25] for a review).

In the following, the perturbation expansion of the gradient-flow coupling is determined in the SU(3) Yang–Mills theory up to two-loop order, using the SMD-variant of NSPT described in the previous section. To one-loop order, the expansion coefficient in the $\overline{\text{MS}}$ scheme of dimensional regularization is known in infinite volume since a while [26], but a huge effort plus the best currently available techniques were required to be able to extend this calculation to the next order [27]. In finite volume with SF boundary conditions, these techniques do not apply and a similar analytical computation may be practically infeasible at present.

6.1 Definition of the coupling

The renormalized coupling considered in this paper belongs to a family of couplings based on the Yang–Mills gradient flow. Explicitly, it is given by [21]

$$\bar{g}^2 = k \left\{ t^2 \langle E(t, x) \rangle \right\}_{T=L, x_0=L/2, \sqrt{8t}=cL}, \tag{6.1}$$

where $E(t, x)$ denotes the Yang–Mills action density at flow time t and position x, c is a parameter of the scheme and the proportionality constant k is determined by the requirement that $\bar{g}^2$ coincides with g_0^2 to lowest order of perturbation theory. Most of the time c will be set to 0.3, which implies a localization range of the action density of about $0.3 \times L$.

On the lattice, the gradient flow is implemented as in Ref. [15]. The Wilson action, with boundary terms so as to ensure the absence of O(a) lattice effects in the flow equation, is thus used to generate the flow. For the action density $E(t, x)$ in Eq. (6.1) either the Wilson action density or the square of the familiar symmetric "clover" expression for the gauge-field tensor is inserted. Furthermore, alternative couplings, where the full action density is replaced by its spatial or time-like part, are considered as well.

All in all this makes six different action densities and couplings, labeled w, ws, wt, c, cs, and ct, where the letters stand for Wilson, clover, space and time, respectively (E^{ct},

---

[2] There is no reason other than simplicity to set the parameter γ that determines the rotation angle to the same value for all fields.

for example, denotes the clover expression for the time-like part of the action density).

6.2 Expansion in powers of α_s

The gradient-flow coupling has an expansion

$$\bar{g}^2 = 4\pi \left\{ \alpha_s + k_1 \alpha_s^2 + k_2 \alpha_s^3 + \cdots \right\} \tag{6.2}$$

in powers of the running coupling α_s in the $\overline{\text{MS}}$ scheme of dimensional regularization at momentum scale q, with coefficients $k_1, k_2, \ldots$ depending on q and L. If q is scaled with L like

$$q = (cL)^{-1}, \tag{6.3}$$

the coefficients are of order one and independent of L in the continuum theory [28].

In the following k_1 and k_2 are computed by combining the expansion

$$t^2 \langle E(t,x) \rangle = E_0 g_0^2 + E_1 g_0^4 + E_2 g_0^6 + \cdots \tag{6.4}$$

obtained in NSPT with the relation

$$\alpha_s = \alpha_0 + d_1 \alpha_0^2 + d_2 \alpha_0^3 + \cdots, \quad \alpha_0 = \frac{g_0^2}{4\pi}, \tag{6.5}$$

between α_s and the bare coupling. For the Wilson gauge action (which is the action used in NSPT), the coefficients d_1 and d_2 are accurately known [29]. The coefficients k_1, k_2 determined in this way depend on the spacing of the simulated lattices so that an extrapolation to the continuum limit is then still required.

6.3 Computation of the coefficients E_k in NSPT

In order to cancel the O(a) lattice effects in E_k, the action of the theory must include boundary counterterms at $x_0 = 0$ and $x_0 = T$ with a tuned coupling [13][3]

$$c_{\text{t}} = 1 + c_{\text{t}}^{(1)} g_0^2 + c_{\text{t}}^{(2)} g_0^4 + \cdots . \tag{6.6}$$

In the case of the Wilson action,

$$c_{\text{t}}^{(1)} = -0.08900(5), \quad c_{\text{t}}^{(2)} = -0.0294(3), \tag{6.7}$$

as was shown long ago [30–32]. The counterterms lead to further terms in the forces that drive the molecular-dynamics evolution, but do not require a modification of the general framework described in Sect. 5. Alternatively, the effects of the counterterms can be included in the calculations by treating $c_{\text{t}} - 1$ as a second coupling.

Once a representative sample of gauge configurations (5.12) has been generated, the stochastic estimation of the gradient-flow coupling requires $E(t,x)$ to be expanded in powers of g_0 for each of these configurations. To this end,

---

[3] In Ref. [15] the improvement coefficient c_{t} is denoted by c_{G}'.

Table 1 Simulation parameters

L	γ	ϵ	$\Delta t_{\text{ms}}/\epsilon$	N_{ms}
10	5.0	0.168	190	59400
12	5.0	0.168	190	59880
14	5.0	0.168	190	59400
16	5.0	0.168	190	59880
18	4.5	0.168	240	58800
20	4.0	0.168	290	59400
24	3.0	0.190	270	71880
28	3.0	0.144	340	79200
32	3.0	0.126	480	88400
40	3.0	0.100	950	80100

the gauge field $V_t(x,\mu)$ at flow time t is represented by a gauge potential $B_\mu(t,x)$ according to

$$V_t(x,\mu) = \exp\{g_0 B_\mu(t,x)\} = 1 + \sum_{r=0}^{\infty} g_0^{r+1} \hat{V}_{t,r}(x,\mu). \tag{6.8}$$

The numerical integration of the flow equation, using a Runge–Kutta integrator, for example, then amounts to applying a sequence of integration steps to the expansion coefficients of the field as in the case of the integration of the molecular-dynamics equations. Gauge damping is however not required here, since the linearized flow is transversal and leaves the gauge modes unchanged.

6.4 Simulation parameters and tables of results

The parameters of the NSPT runs performed for the "measurement" of E_0, E_1, E_2 are listed in Table 1. In all cases, the gauge-damping parameter λ_0 was set to 2 and the fourth-order OMF integrator was employed for the molecular-dynamics equations (cf. Appendix A). Measurements were made after every $\Delta t_{\text{ms}}/\epsilon$ update cycles using a third-order Runge–Kutta integrator for the gradient flow [26], with step size varying from 0.002 at small flow times to 0.1 at large times. With this scheme, the gradient-flow integration errors are guaranteed to be completely negligible with respect to the statistical errors. The number N_{ms} of measurements made is listed in the last column of Table 1 (the programs that were used in these simulations can be downloaded from http://luscher.web.cern.ch/luscher/NSPT).

At the chosen values of the parameters, the bounds (5.29) and (5.32) are satisfied by a wide margin so that the convergence of the SMD algorithm is rigorously guaranteed. The results obtained on the simulated lattices for the expansion coefficients k_1 and k_2 and their statistical errors are listed in Appendix D, for $c \in \{0.2, 0.3, 0.4\}$ and all choices w,c,ws,... of the action density (cf. Sect. 6.1).

7 Statistical and systematic errors

The values of k_1 and k_2 obtained in NSPT depend on the scheme parameter c, the lattice size L (in lattice units), the simulation step size ϵ and the SMD parameter γ. No attempt is made here to determine all these dependencies in detail. Instead some basic facts and empirical results are discussed that helped controlling the errors in the case of the simulations reported in this paper (see Refs. [8,33] for related complementary studies of the ϕ^4 theory).

7.1 Autocorrelations and statistical variances

Usually the variances of the observables are a property of the simulated field theory and hence independent of the simulation algorithm. In NSPT this is not so, because the algorithm is not exact, but mainly because the square of the order-r coefficient $\hat{\mathcal{O}}_r$ of an observable $\mathcal{O}$ in general does not coincide with the order-$2r$ coefficient of another observable. The time average of $\hat{\mathcal{O}}_r^2$ and thus the variance of $\hat{\mathcal{O}}_r$ are then not necessarily determined by the theory alone. For illustration, the dependence on γ of the variances of the coefficients in

$$\frac{t^2}{L^3} \sum_x E(t,x) = \hat{E}_0 g_0^2 + \hat{E}_1 g_0^4 + \hat{E}_2 g_0^6 + \dots \qquad (7.1)$$

(where $x_0 = L/2$ and $\sqrt{8t} = cL$ as before) is shown in Fig. 1. As will be discussed in Sect. 7.3, the integration errors are negligible with respect to the rapid growth of the variances of the one- and two-loop coefficients seen at small γ, which is therefore entirely an effect of the change of algorithm.

In practice one would like to minimize the computational work required to obtain the calculated coefficients to a given statistical precision. The simulation algorithm should thus be tuned so as to minimize the products

$$\tau_{\text{int}}(\hat{\mathcal{O}}_r) \times \text{var}(\hat{\mathcal{O}}_r) \qquad (7.2)$$
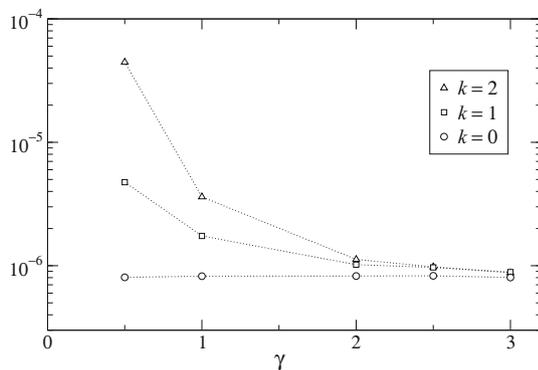
Fig. 1 Variances of $\hat{E}_k^c$ at fixed $L = 16, c_t = 1, c = 0.3$ and $\epsilon = 0.238$ versus γ. The *dotted lines* connecting the data points are drawn to guide the eye. Beyond $\gamma = 3$ the variances are practically constant

Table 2 Autocorrelation times of $\hat{E}_k^c$ and associated products (7.2)[†]

γ	$k = 0$		$k = 1$		$k = 2$	
0.5	2.0	1.6×10^{-6}	2.0	9.4×10^{-6}	2.0	8.8×10^{-5}
1.0	2.7	2.2×10^{-6}	3.0	5.2×10^{-6}	2.9	1.0×10^{-5}
2.0	5.9	4.9×10^{-6}	6.7	6.8×10^{-6}	7.3	8.1×10^{-6}
2.5	6.7	5.6×10^{-6}	8.4	8.1×10^{-6}	9.2	9.1×10^{-6}
3.0	7.7	6.2×10^{-6}	9.4	8.4×10^{-6}	11.2	9.8×10^{-6}
5.0	10.9	8.7×10^{-6}	13.0	1.1×10^{-5}	16.6	1.3×10^{-5}
9.0	14.8	1.2×10^{-5}	17.7	1.4×10^{-5}	21.5	1.5×10^{-5}

† All lattice and algorithm parameters are as in Fig. 1. The autocorrelation times are given in units of simulation time

of the integrated autocorrelation times and variances of the order-r coefficients $\hat{\mathcal{O}}_r$ of the observables $\mathcal{O}$ of interest. Empirical studies show that the two factors often work against each other, i.e. algorithms tuned for small autocorrelations tend to give large variances and vice versa.

The autocorrelation times of the coefficients $\hat{E}_k^c$, for example, grow monotonically with γ (see Table 2). At the chosen point in parameter space, the associated products (7.2) are then minimized at values of γ around 0.5, 1.0 and 2.0 for $k = 0$, 1 and 2. The example shows that there may be no uniformly best choice of γ, but a range of values, where all coefficients of interest are obtained reasonably efficiently.

7.2 Critical slowing down

The behaviour of the autocorrelation times and variances near the continuum limit $L \rightarrow \infty$ depends on the simulation algorithm and the observables considered. When NSPT is based on the Langevin equation, the autocorrelations of the coefficients of multiplicatively renormalizable quantities can be shown to diverge proportionally to L^2 [34,35], while their variances grow at most logarithmically [36].

At large γ, the variant of NSPT studied here is expected to behave similarly, since the SMD algorithm then effectively integrates the Langevin equation. Choosing γ to depend on L in some particular way may, however, conceivably lead to an improved scaling behaviour. In the free theory, for example, the autocorrelation times grow proportionally to L rather than L^2 if γ is scaled like $1/L$ [5–7]. Beyond the leading order, the situation is complicated by the algorithm-dependence of the variances and the effects of the parameter tuning are then not easy to predict.

Considering the fact that the computational cost of the measurements of $\hat{E}_k$ tends to be larger than the one of the SMD update cycles, the parameters of the runs on the large lattices (those with $L = 24, \dots, 40$ in Table 1) were chosen so that subsequent measurements are practically uncorrelated. At fixed $\gamma = 3$ the required increase with L of the measurement time separation Δt_{ms} then turned out to be quite

moderate. Moreover, from $L = 24$ to $L = 40$, the variances of $\hat{E}_k$ grow only slowly (at $c = 0.3$ and for $k = 0$, 1 and 2 by about 2, 19 and 30 percent).

7.3 Integration errors

As already noted in Appendix A, the theory is very accurately simulated at leading order if the fourth-order OMF integrator is used for the molecular-dynamics equations. The expectation values of the coefficient $\hat{E}_0$ calculated in the runs listed in Table 1 in fact all agree with the known analytic formula [15,21] within a relative statistical precision of about 2×10^{-4}.

Beyond the leading order, the integration errors remain difficult to detect in empirical studies (see Fig. 2). The stability bounds (5.29), (5.32) are respected in all these runs and the coefficients $\hat{H}_0, \ldots, \hat{H}_4$ of the Hamilton function H are accurately conserved, with deficits decreasing like ϵ^5 (as has to be the case in the asymptotic regime of a fourth-order integrator). It thus seems safe to conclude that the integration errors in the tests reported in Fig. 2 are smaller than the statistical errors.

Apart from adjustments to match the target statistics, the step sizes in the runs listed in Table 1 were, for $L \geq 24$, scaled proportionally to $1/L$ so that the integration errors fall off like $1/L^4$ at large L and thus much more rapidly than the leading lattice effects. Since the statistical errors are approximately the same in all runs, the scaling of the step sizes may be a luxury, but was applied here as a safeguard measure against an enhancement of systematic errors through the continuum extrapolation.

7.4 Extrapolation to the continuum limit

With O(a)-improvement in place, the L-dependence of the one-loop coefficient k_1^{w} is asymptotically given by [37,38]
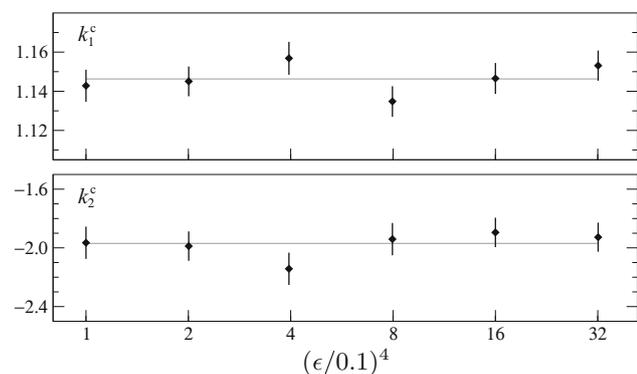
Fig. 2 Dependence of k_1^{c} and k_2^{c} at $L = 24$, $c_{\mathrm{t}} = 1$ and $c = 0.3$ on the simulation step size ϵ (note the log scale on the abscissa). In the range of data shown, ϵ increases from 0.1 to 0.238. Each data point is based on approximately 10^4 statistically independent measurements of the coefficients $\hat{E}_0^{\mathrm{c}}$, $\hat{E}_1^{\mathrm{c}}$ and $\hat{E}_2^{\mathrm{c}}$
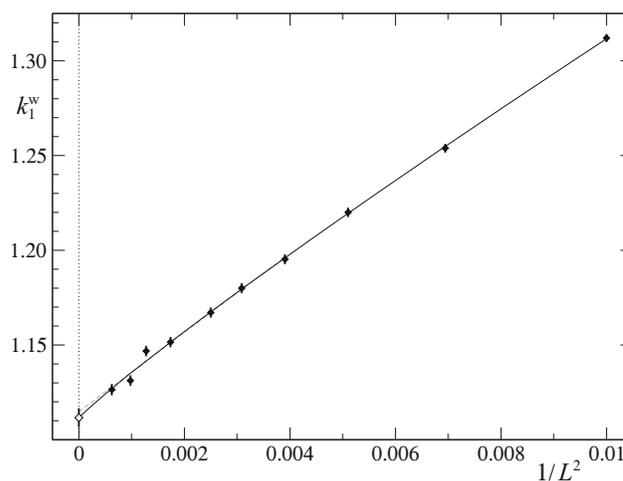
Fig. 3 Extrapolation of the simulation results for k_1^{w} at $c = 0.3$ to the continuum limit (*open point*). The full line is obtained by fitting the data in the range $10 \leq L \leq 40$ with the asymptotic expression (7.3). A linear fit in the range $16 \leq L \leq 40$ (*dashed grey line*) lies practically on top of the full line

$$k_1^{\mathrm{w}} \underset{L \to \infty}{=} a_0 + \{a_1 + b_1 \ln L\}L^{-2} + \mathrm{O}(L^{-3}), \qquad (7.3)$$

where the leading-order term, a_0, coincides with the coefficient k_1 in the continuum theory. The available data are well described by this asymptotic expression down to $L = 10$ if $c \geq 0.3$ and $L = 12$ if $c = 0.2$ (see Fig. 3). Note the fact that the data points in Fig. 3 are uncorrelated and that random fluctuations by more than one standard deviation are consequently not improbable in a sample of 10 points. Plots of the other one-loop coefficients k_1^{c}, k_1^{ws}, $\ldots$ look much the same.

Even though only a short extrapolation is required to reach the continuum limit, there is no way the extrapolation errors can be rigorously assessed. Following standard practice, the errors are estimated by varying the fit procedure within reasonable bounds. In particular, fits are discarded if the quality of the fit is low or if the fit parameters are poorly determined. Another indication of the size of the extrapolation errors is provided by the deviations of the results obtained using the "w" and the "c" data, respectively. The errors of the results quoted in Table 3 include both the statistical and an estimate of the extrapolation errors.

With increasing loop order, the continuum extrapolations tend to become more difficult, partly because one loses statistical precision and partly because the asymptotic formulae describing the leading lattice effects have more and more terms. In particular, with respect to the one-loop coefficients, the coefficients at the next order are obtained about 10 times less precisely and their asymptotic form

$$k_2^{\mathrm{w}} = a_0 + \{a_1 + b_1 \ln L + c_1 (\ln L)^2\}L^{-2} + \mathrm{O}(L^{-3})$$

$$(7.4)$$
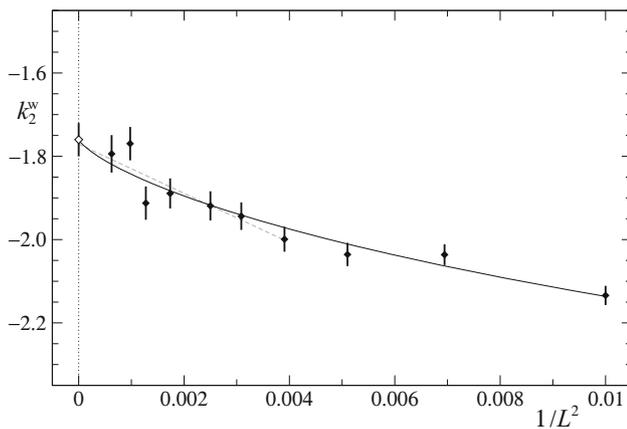
Fig. 4 Extrapolation of the simulation results for k_2^{w} at $c = 0.3$ to the continuum limit (*open point*). The full line is obtained by fitting the data in the range $10 \leq L \leq 40$ according to Eq. (7.4), with the two lowest modes of the design matrix projected away. In the range $16 \leq L \leq 40$ the data can also be fitted by a straight line (*dashed grey line*)

includes an additional logarithm. Combinations of the logarithms in this expression only too easily accurately approximate a constant in the range of the available data, but may blow up closer to the continuum limit and then strongly influence the result of the extrapolation.

In the present context, the goal is not to determine the values of the coefficients a_1, b_1 and c_1, but to perform a short extrapolation of the data to the continuum limit. A sensible way to stabilize fits of the data by the asymptotic expression (7.4) is then to constrain the minimization of the likelihood function to the directions in parameter space orthogonal to its flattest directions (see Fig. 4). The results quoted in Table 3 were obtained in this way and by varying the fit procedure, as in the case of the one-loop coefficients, in order to assess the extrapolation errors.

7.5 Miscellaneous remarks

Lattice effects. Since the smoothing range of the gradient flow decreases with c, it is no surprise that the coefficients k_1, k_2 calculated in NSPT are found to be increasingly sensitive to lattice effects when c is lowered. The continuum extrapolation of the data then becomes more and more difficult and eventually requires larger lattices to be simulated.

Infinite-volume limit. The gradient-flow coupling in infinite volume runs with the flow time t and may be expanded in

Table 4 Ratio $\rho_2 = \beta_2/\beta_0$ of coefficients of the β-function

c	ρ_2	ρ_2^{s}	ρ_2^{t}
0.2	$-2.38(6)$	$-2.51(6)$	$-2.22(7)$
0.3	$-2.99(4)$	$-3.74(5)$	$-2.29(6)$
0.4	$-4.88(6)$	$-8.04(7)$	$-2.21(8)$

powers of α_s at scale $q = (8t)^{-1/2}$, as in Eq. (6.2), the one- and two-loop coefficients in the continuum limit being $k_1 = 1.0978(1)$ [26] and $k_2 = -0.9822(1)$ [27]. In the finite-volume scheme considered in this paper, and after passing to the continuum limit, the infinite-volume limit is reached by sending c to zero. The results listed in Table 3 cannot be reliably extrapolated to $c = 0$, but the values of $k_1, \ldots, k_2^{\mathrm{t}}$ at $c = 0.2$ are actually already quite close to the infinite-volume values quoted above.

Three-loop β-function. The L-dependence of the gradient-flow coupling $\alpha = \bar{g}^2/4\pi$ is governed by the renormalization group equation

$$L\frac{\partial \alpha}{\partial L} = \beta_0\alpha^2 + \beta_1\alpha^3 + \beta_2\alpha^4 + \ldots, \qquad (7.5)$$

where $\beta_0 = 11(2\pi)^{-1}$ and $\beta_1 = 51(2\pi)^{-2}$ are the universal one- and two-loop coefficients of the Callan–Symanzik β-function. Using the results quoted in Table 3, the three-loop coefficient may be calculated in a few lines (see Table 4). The coefficient thus has opposite sign and is significantly larger than in the $\overline{\mathrm{MS}}$ scheme, particularly so at $c = 0.4$ and if the spatial part of the Yang–Mills action density is inserted in the definition (6.1) of the coupling.

8 Conclusions

In stochastic perturbation theory the fields are represented by a series of coefficient fields that solve the underlying stochastic equation order by order in the interactions. Beyond the leading order, the probability distributions of the coefficient fields are however not a priori known and actually depend on the chosen stochastic process. The variances of the observables of interest must consequently be expected to vary with the parameters of the simulation algorithm, an effect that tends to considerably complicate the situation with respect to the one in simulations based on importance sampling.

Table 3 Values of k_1 and k_2 in the continuum limit

c	k_1	k_1^{s}	k_1^{t}	k_2	k_2^{s}	k_2^{t}
0.2	1.089(6)	1.106(7)	1.066(7)	$-1.21(6)$	$-1.29(6)$	$-1.12(7)$
0.3	1.112(5)	1.220(6)	1.005(7)	$-1.76(4)$	$-2.17(5)$	$-1.36(6)$
0.4	1.297(5)	1.685(6)	0.935(6)	$-3.06(6)$	$-4.78(7)$	$-1.47(8)$

The SMD algorithm provides a technically attractive basis for NSPT. Compared to the traditional version of NSPT [1–3], where the starting point is the Langevin equation, a significantly improved efficiency is achieved, particularly so near the continuum limit. Moreover, the available highly accurate integrators for the molecular-dynamics equations allow the integration errors to be easily kept under control. For the reasons mentioned above, some tuning of the friction parameter γ is however required and must take into account the variances of the observables of interest.

The inclusion of the quark fields in the SMD algorithm along the lines of Sect. 5 is straightforward and is not expected to slow down the simulations by a large factor [3]. In general, the cost of NSPT computations very much depends on the observables considered, the order in perturbation theory and the desired precision of the results.

The statistical errors of the expansion coefficients k_l of the gradient-flow coupling, for example, appear to grow rapidly with the loop order l. In practice some loss of precision from one order to the next is tolerable, since the coefficients get multiplied by α_s^{l+1} in the perturbation series (6.2). An extension of the computations to three-loop order would however only make sense if k_1 and k_2 are recalculated with errors about an order of magnitude smaller than the ones quoted in Table 3. Furthermore, the relation between α_s and the bare coupling would need to be worked out to three loops.

Appendix A: OMF molecular-dynamics integrators

Among the popular integrators used in lattice QCD, there are two schemes proposed by Omelyan, Mryglod and Folk (OMF) [10], one of second order in the integration step size and the other of fourth order. Here their efficiency is studied by comparing the distribution (4.8) actually simulated at leading order of stochastic perturbation theory with the exact distribution.

A.1 Second-order OMF integrator

The sequence of update steps (2.8), (2.9) is

$$I_\epsilon = I_{p,\epsilon_1} I_{q,\epsilon_2} I_{p,\epsilon_3} I_{q,\epsilon_2} I_{p,\epsilon_1}, \quad \epsilon_k = r_k \epsilon, \tag{A.1}$$

$$r_1 = 0.1931833275037836, \quad r_2 = 1/2, \quad r_3 = 1 - 2r_1, \tag{A.2}$$

in this case. Using an algebraic manipulation program, the expressions

$$M_{qp} = \epsilon \left(1 - \frac{1}{4} r_3 \epsilon^2 \Delta \right), \tag{A.3}$$

$$\hat{\Delta} = \Delta \left(1 - \frac{1}{2} r_1 \epsilon^2 \Delta \right) \left(1 - \frac{1}{2} r_1 r_3 \epsilon^2 \Delta \right) \left(1 - \frac{1}{4} r_3 \epsilon^2 \Delta \right)^{-1} \tag{A.4}$$

are then easily obtained. The convergence of the SMD algorithm is thus guaranteed if $\epsilon^2 \|\Delta\| < 6.51$. With respect to the leapfrog integrator, and if the step size ϵ is adjusted so that the number of force evaluations per unit time is the same, this scheme achieves a reduction of the relative deviation of $\hat{\Delta}$ from Δ by a factor 25.

A.2 Fourth-order OMF integrator

This integrator is of the form (A.1) too, but with 11 steps instead of 5. The associated step sizes in units of ϵ are

$$r_1 = 0.08398315262876693,$$
$$r_2 = 0.2539785108410595,$$
$$r_3 = 0.6822365335719091,$$
$$r_4 = -0.03230286765269967,$$
$$r_5 = 1/2 - r_1 - r_3, \quad r_6 = 1 - 2(r_2 + r_4), \tag{A.5}$$

and the matrices M_{pq} and M_{qp} are polynomials in $\epsilon^2 \Delta$ of degree 5 and 4 times $\epsilon \Delta$ and ϵ, respectively. In this case, M_{qp} is non-singular and

$$\hat{\Delta} = \Delta \left\{ 1 + a_1 \epsilon^4 \Delta^2 + a_2 \epsilon^6 \Delta^3 + \cdots \right\}, \tag{A.6}$$

$$a_1 = -2.58(1) \times 10^{-5}, \quad a_2 = -1.88(1) \times 10^{-5}, \tag{A.7}$$

is positive if $\epsilon^2 \|\Delta\| < 9.87$.

As is evident from Eqs. (A.6), (A.7), this integrator is impressively accurate. If $\|\Delta\| = 16$ and $\epsilon = 0.2$, for example, the relative deviation of $\hat{\Delta}$ from Δ is at most 1.6×10^{-5} and thus about a factor 12 smaller than the deviation in the case of the second-order OMF integrator (with an adjusted step size of $\epsilon = 0.08$).

Appendix B: Explicit form of d^*

The operator d^* depends on the chosen boundary conditions and the scalar products in the field spaces $\mathcal{H}_0$ and $\mathcal{H}_1$ (cf. Sect. 5.1). For SF boundary conditions

$$(d^*A)(x) = -\frac{T - x_0}{T^3 L^3} \sum_{y_0=0}^{T-1} \sum_y A_0(y) - \sum_{\mu=0}^{3} \partial_\mu^* A_\mu(x), \tag{B.1}$$

where

$$\partial_\mu^* A_\mu(x) = \begin{cases} A_\mu(x) - A_\mu(x - \hat{\mu}) & \text{if } 0 < x_0 < T, \\ 0 & \text{otherwise,} \end{cases} \tag{B.2}$$

while in the case of open-SF boundary conditions,

$$(d^*A)(x) = -\sum_{\mu=0}^{3} \partial_\mu^* A_\mu(x), \tag{B.3}$$

where now

$$\partial_\mu^* A_\mu(x) = \begin{cases} A_\mu(x) - A_\mu(x - \hat{\mu}) & \text{if } \mu > 0 \text{ or } 0 < x_0 < T, \\ 2A_0(x) & \text{if } \mu = 0 \text{ and } x_0 = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.4}$$

Appendix C: Gauge-damped stochastic equations

In the continuous-time limit $\epsilon \to 0$, the gauge-damped SMD algorithm described in Sect. 5 integrates the stochastic equations

$$\partial_t U_t(x, \mu) = g_0 \{\pi_t(x, \mu) - \nabla_\mu \omega_t(x)\} U_t(x, \mu), \tag{C.1}$$

$$\partial_t \pi_t(x, \mu) = -g_0 w_{x,\mu} (\partial_{x,\mu}^a S_G)(U_t) T^a - \gamma \pi_t(x, \mu)$$
$$+ \eta_t(x, \mu) + g_0 [\omega_t(x), \pi_t(x, \mu)], \tag{C.2}$$

$$\partial_t \omega_t(x) = -\gamma \omega_t(x) + \lambda_0 (d^* C_t)(x), \tag{C.3}$$

where π_t, U_t, ω_t denote the momentum, gauge and gauge-damping fields at simulation time t (and the link field C_t is given by Eq. (5.23) with U replaced by U_t). The normalization of the Gaussian random field is such that

$$\langle \eta_t^a(x, \mu) \eta_s^b(y, \nu) \rangle = 2\gamma w_{x,\mu} \delta^{ab} \delta_{\mu\nu} \delta(t - s) \delta_{xy} \tag{C.4}$$

and

$$\nabla_\mu \omega_t(x) = U_t(x, \mu) \omega_t(x + \hat{\mu}) U_t(x, \mu)^{-1} - \omega_t(x) \tag{C.5}$$

stands for the gauge-covariant gradient of the gauge-damping field.

As in the case of the SMD algorithm, the gauge-damping terms can be removed from the continuous-time process by applying a time-dependent gauge transformation to the fields π_t, U_t and η_t. Up to a rescaling of the simulation time and the momentum field by the coupling g_0, Eqs. (C.1), (C.2) then reduce to the standard stochastic molecular-dynamics equations [5–7]. The gauge damping can also easily be shown to coincide with the standard one in the Langevin limit $\gamma \to \infty$ [39].

Appendix D: Tables of simulation results

See Tables 5, 6, 7, 8, 9 and 10

Table 5 Values of the expansion coefficient k_1 at $c = 0.2$

L	k_1^w	k_1^c	k_1^{ws}	k_1^{cs}	k_1^{wt}	k_1^{ct}
10	1.4876(9)	1.3126(12)	1.5044(12)	1.3396(16)	1.4708(13)	1.2857(18)
12	1.3971(10)	1.2908(12)	1.4166(14)	1.3169(16)	1.3777(15)	1.2648(18)
14	1.3199(12)	1.2564(13)	1.3378(15)	1.2784(17)	1.3021(17)	1.2345(19)
16	1.2676(12)	1.2262(13)	1.2867(16)	1.2483(17)	1.2485(18)	1.2042(20)
18	1.2319(13)	1.2026(14)	1.2518(17)	1.2249(18)	1.2121(19)	1.1804(21)
20	1.2063(13)	1.1842(14)	1.2270(18)	1.2069(18)	1.1857(20)	1.1617(21)
24	1.1725(14)	1.1587(14)	1.1910(18)	1.1784(18)	1.1540(20)	1.1390(21)
28	1.1561(14)	1.1466(14)	1.1756(18)	1.1670(18)	1.1367(21)	1.1263(21)
32	1.1369(14)	1.1298(14)	1.1578(18)	1.1514(18)	1.1161(20)	1.1083(21)
40	1.1202(15)	1.1159(15)	1.1401(19)	1.1363(19)	1.1004(22)	1.0957(22)

Table 6 Values of the expansion coefficient k_1 at $c = 0.3$

L	k_1^w	k_1^c	k_1^{ws}	k_1^{cs}	k_1^{wt}	k_1^{ct}
10	1.3120(21)	1.3062(24)	1.4241(28)	1.4409(30)	1.2021(30)	1.1745(34)
12	1.2539(23)	1.2545(24)	1.3710(31)	1.3869(32)	1.1391(32)	1.1250(35)
14	1.2200(25)	1.2225(26)	1.3316(32)	1.3448(33)	1.1105(36)	1.1027(38)
16	1.1952(26)	1.1977(27)	1.3061(33)	1.3164(34)	1.0866(38)	1.0816(39)
18	1.1800(27)	1.1823(28)	1.2937(35)	1.3023(36)	1.0685(39)	1.0647(40)
20	1.1671(27)	1.1691(28)	1.2799(35)	1.2870(36)	1.0566(40)	1.0537(41)
24	1.1515(27)	1.1532(27)	1.2595(35)	1.2646(35)	1.0456(39)	1.0441(40)
28	1.1468(28)	1.1482(28)	1.2593(36)	1.2632(36)	1.0365(41)	1.0354(42)
32	1.1312(28)	1.1322(28)	1.2444(35)	1.2474(36)	1.0203(41)	1.0194(41)
40	1.1264(30)	1.1271(30)	1.2359(39)	1.2378(39)	1.0192(45)	1.0187(45)

Table 7 Values of the expansion coefficient k_1 at $c = 0.4$

L	k_1^w	k_1^c	k_1^{ws}	k_1^{cs}	k_1^{wt}	k_1^{ct}
10	1.3582(35)	1.4408(38)	1.7545(48)	1.8741(50)	0.9928(49)	1.0415(53)
12	1.3373(37)	1.3953(39)	1.7405(52)	1.8244(53)	0.9651(52)	0.9994(55)
14	1.3337(41)	1.3772(43)	1.7303(55)	1.7927(57)	0.9667(57)	0.9930(59)
16	1.3197(43)	1.3527(44)	1.7097(57)	1.7568(58)	0.9592(61)	0.9793(63)
18	1.3183(44)	1.3444(45)	1.7137(59)	1.7513(60)	0.9525(61)	0.9681(63)
20	1.3181(45)	1.3396(46)	1.7111(59)	1.7419(59)	0.9543(64)	0.9673(65)
24	1.3121(44)	1.3271(44)	1.6983(58)	1.7197(58)	0.9544(61)	0.9636(62)
28	1.3116(46)	1.3227(47)	1.7068(59)	1.7226(60)	0.9456(66)	0.9523(67)
32	1.2987(46)	1.3070(47)	1.6918(60)	1.7038(60)	0.9349(65)	0.9399(66)
40	1.3012(50)	1.3066(51)	1.6876(65)	1.6953(65)	0.9433(72)	0.9467(72)

Table 8 Values of the expansion coefficient k_2 at $c = 0.2$

L	k_2^w	k_2^c	k_2^{ws}	k_2^{cs}	k_2^{wt}	k_2^{ct}
10	$-1.863(9)$	$-1.893(12)$	$-1.919(12)$	$-1.985(16)$	$-1.807(12)$	$-1.801(18)$
12	$-1.760(11)$	$-1.742(13)$	$-1.820(14)$	$-1.826(17)$	$-1.700(15)$	$-1.658(19)$
14	$-1.706(12)$	$-1.679(14)$	$-1.755(16)$	$-1.739(17)$	$-1.657(17)$	$-1.620(20)$
16	$-1.655(13)$	$-1.630(14)$	$-1.718(17)$	$-1.702(18)$	$-1.593(19)$	$-1.558(21)$
18	$-1.608(15)$	$-1.586(16)$	$-1.674(19)$	$-1.660(20)$	$-1.542(22)$	$-1.512(24)$
20	$-1.562(16)$	$-1.544(17)$	$-1.611(20)$	$-1.597(21)$	$-1.513(24)$	$-1.490(25)$
24	$-1.486(17)$	$-1.474(18)$	$-1.534(22)$	$-1.526(23)$	$-1.438(25)$	$-1.423(27)$
28	$-1.473(18)$	$-1.465(19)$	$-1.520(24)$	$-1.515(24)$	$-1.426(28)$	$-1.415(29)$
32	$-1.365(18)$	$-1.357(19)$	$-1.438(24)$	$-1.433(24)$	$-1.293(28)$	$-1.282(28)$
40	$-1.317(21)$	$-1.312(21)$	$-1.398(28)$	$-1.395(28)$	$-1.236(32)$	$-1.230(32)$

Table 9 Values of the expansion coefficient k_2 at $c = 0.3$

L	k_2^{w}	k_2^{c}	k_2^{ws}	k_2^{cs}	k_2^{wt}	k_2^{ct}
10	−2.134(23)	−2.229(26)	−2.582(31)	−2.765(33)	−1.695(33)	−1.706(37)
12	−2.036(25)	−2.097(27)	−2.502(34)	−2.618(36)	−1.580(37)	−1.587(40)
14	−2.036(28)	−2.084(30)	−2.469(36)	−2.555(38)	−1.611(41)	−1.624(44)
16	−1.999(30)	−2.034(31)	−2.427(38)	−2.488(40)	−1.580(45)	−1.589(47)
18	−1.944(33)	−1.969(34)	−2.385(42)	−2.433(43)	−1.511(49)	−1.514(50)
20	−1.919(35)	−1.939(36)	−2.349(44)	−2.388(45)	−1.498(52)	−1.500(54)
24	−1.889(36)	−1.905(37)	−2.267(48)	−2.294(48)	−1.518(54)	−1.524(55)
28	−1.912(40)	−1.924(40)	−2.337(51)	−2.357(51)	−1.497(60)	−1.500(60)
32	−1.770(40)	−1.778(40)	−2.222(51)	−2.238(51)	−1.327(60)	−1.328(61)
40	−1.794(45)	−1.800(46)	−2.229(58)	−2.239(58)	−1.368(68)	−1.370(69)

Table 10 Values of the expansion coefficient k_2 at $c = 0.4$

L	k_2^{w}	k_2^{c}	k_2^{ws}	k_2^{cs}	k_2^{wt}	k_2^{ct}
10	−3.118(41)	−3.417(45)	−4.908(56)	−5.384(59)	−1.468(58)	−1.605(63)
12	−3.039(45)	−3.248(47)	−4.858(61)	−5.183(63)	−1.361(64)	−1.463(67)
14	−3.164(50)	−3.321(52)	−4.950(67)	−5.189(70)	−1.511(70)	−1.594(73)
16	−3.124(53)	−3.241(55)	−4.863(71)	−5.041(73)	−1.52(8)	−1.58(8)
18	−3.054(58)	−3.142(59)	−4.82(8)	−4.96(8)	−1.42(8)	−1.46(8)
20	−3.106(61)	−3.179(63)	−4.88(8)	−4.99(8)	−1.47(9)	−1.50(9)
24	−3.080(62)	−3.131(63)	−4.75(8)	−4.83(8)	−1.54(9)	−1.56(9)
28	−3.125(69)	−3.163(70)	−4.89(9)	−4.95(9)	−1.49(10)	−1.51(10)
32	−2.996(70)	−3.025(71)	−4.79(9)	−4.84(9)	−1.33(10)	−1.35(10)
40	−3.06(8)	−3.08(8)	−4.77(10)	−4.79(10)	−1.49(11)	−1.50(12)

References

1. F. Di Renzo, G. Marchesini, P. Marenzoni, E. Onofri, Lattice perturbation theory on the computer. Nucl. Phys. Proc. Suppl. **34**, 795 (1994)
2. F. Di Renzo, E. Onofri, G. Marchesini, P. Marenzoni, Four loop result in SU(3) lattice gauge theory by a stochastic method: lattice correction to the condensate. Nucl. Phys. B **426**, 675 (1994)
3. F. Di Renzo, L. Scorzato, Numerical stochastic perturbation theory for full QCD. JHEP **0410**, 073 (2004)
4. S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Hybrid Monte Carlo. Phys. Lett. B **195**, 216 (1987)
5. A.M. Horowitz, Stochastic quantization in phase space. Phys. Lett. B **156**, 89 (1985)
6. A.M. Horowitz, The second order Langevin equation and numerical simulations. Nucl. Phys. B **280**(FS18), 510 (1987)
7. A.M. Horowitz, A generalized guided Monte Carlo algorithm. Phys. Lett. B **268**, 247 (1991)
8. M. Dalla Brida, M. Garofalo, A. D. Kennedy, Numerical stochastic perturbation theory and gradient flow in ϕ^4 theory. PoS LATTICE2015, 309 (2016)
9. K. Jansen, C. Liu, Kramers equation algorithm for simulations of QCD with two flavors of Wilson fermions and gauge group SU(2). Nucl. Phys. B **453**, 375 (1995). Erratum: Nucl. Phys. B **459**, 437 (1996)
10. I.P. Omelyan, I.M. Mryglod, R. Folk, Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations. Comp. Phys. Commun. **151**, 272 (2003)
11. G. Parisi, Y.-S. Wu, Perturbation theory without gauge fixing. Sci. Sin. **24**, 483 (1981)
12. P.H. Damgaard, H. Hüffel, Stochastic quantization. Phys. Rep. **152**, 227 (1987)
13. M. Lüscher, R. Narayanan, P. Weisz, U. Wolff, The Schrödinger functional—a renormalizable probe for non-Abelian gauge theories. Nucl. Phys. B **384**, 168 (1992)
14. S. Sint, On the Schrödinger functional in QCD. Nucl. Phys. B **421**, 135 (1994)
15. M. Lüscher, Step scaling and the Yang–Mills gradient flow. JHEP **1406**, 105 (2014)
16. K.G. Wilson, Confinement of quarks. Phys. Rev. D **10**, 2445 (1974)
17. P. Weisz, Continuum limit improved lattice action for pure Yang–Mills theory (I). Nucl. Phys. B **212**, 1 (1983)
18. M. Lüscher, P. Weisz, On-shell improved lattice gauge theories. Commun. Math. Phys. **97**, 59 (1985). Erratum: Commun. Math. Phys. **98**, 433 (1985)
19. Y. Iwasaki, Renormalization group analysis of lattice theories and improved lattice action. II—Four-dimensional non-Abelian SU(N) gauge model, preprint UTHEP-118 (1983). arXiv:1111.7054v1
20. M. Lüscher, Computational strategies in lattice QCD, in *Modern perspectives in lattice QCD*, ed. by L. Lellouch et al. (Oxford University Press, New York, 2011)
21. P. Fritzsch, A. Ramos, The gradient flow coupling in the Schrödinger functional. JHEP **1310**, 008 (2013)
22. M. Dalla Brida et al. [ALPHA Collab.], Slow running of the gradient flow coupling from 200 MeV to 4 GeV in $N_f = 3$ QCD. Phys. Rev. D **95**, 014507 (2017)
23. I. Campos et al., Non-perturbative running of quark masses in three-flavour QCD, PoS LATTICE2016, 201 (2016)

24. M. Lüscher, P. Weisz, U. Wolff, A numerical method to compute the running coupling in asymptotically free theories. Nucl. Phys. B **359**, 221 (1991)

25. M. Lüscher, Advanced lattice QCD, in *Probing the Standard Model of Particle Interactions (Les Houches 1997)*, ed. by R. Gupta et al. (Elsevier, Amsterdam, 1999)

26. M. Lüscher, Properties and uses of the Wilson flow in lattice QCD. JHEP **1008**, 071 (2010). Erratum: JHEP **1403**, 092 (2014)

27. R.V. Harlander, T. Neumann, The perturbative QCD gradient flow to three loops. JHEP **1606**, 161 (2016)

28. M. Lüscher, P. Weisz, Perturbative analysis of the gradient flow in non-Abelian gauge theories. JHEP **1102**, 051 (2011)

29. M. Lüscher, P. Weisz, Two-loop relation between the bare lattice coupling and the $\overline{\text{MS}}$ coupling in pure SU(N) gauge theories. Phys. Lett. B **349**, 165 (1995)

30. M. Lüscher, R. Sommer, P. Weisz, U. Wolff, A precise determination of the running coupling in the SU(3) Yang–Mills theory. Nucl. Phys. B **413**, 481 (1994)

31. A. Bode, U. Wolff, P. Weisz, Two-loop computation of the Schrödinger functional in pure SU(3) lattice gauge theory. Nucl. Phys. B **540**, 491 (1999)

32. A. Bode, P. Weisz, U. Wolff, *Two-loop computation of the Schrödinger functional in lattice QCD*, Nucl. Phys. B **576**, 517 (2000). [Erratum: Nucl. Phys. B **600**, 453 (2001); Nucl. Phys. B **608**, 481] (2001)

33. M. Dalla Brida, M. Garofalo, A. D. Kennedy, An investigation of new methods for numerical stochastic perturbation theory in φ^4 theory. arXiv:1703.04406 [hep-lat]

34. J. Zinn-Justin, Renormalization and stochastic quantization. Nucl. Phys. B **275**(FS17), 135 (1986)

35. J. Zinn-Justin, D. Zwanziger, Ward identities for the stochastic quantization of gauge fields. Nucl. Phys. B **295**(FS21), 297 (1988)

36. M. Lüscher, Statistical errors in stochastic perturbation theory, notes (2015). http://luscher.web.cern.ch/luscher/notes/enspt.pdf

37. K. Symanzik, Cutoff dependence in lattice ϕ_4^4 theory, in *Recent Developments in Gauge Theories (Cargèse 1979)*, ed. by G. 't Hooft et al. (Plenum, New York, 1980)

38. K. Symanzik, Some topics in quantum field theory, in *Mathematical Problems in Theoretical Physics*, vol. 153, Lecture Notes in Physics, ed. by R. Schrader et al. (Springer, New York, 1982)

39. D. Zwanziger, Covariant quantization of gauge fields without Gribov ambiguity. Nucl. Phys. B **192**, 259 (1981)