DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Error Bounds of MCMC for Functions with Unbounded Stationary Variance

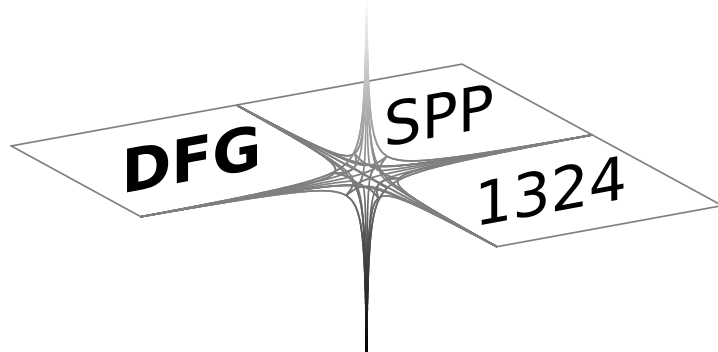D. Rudolf, N. Schweizer
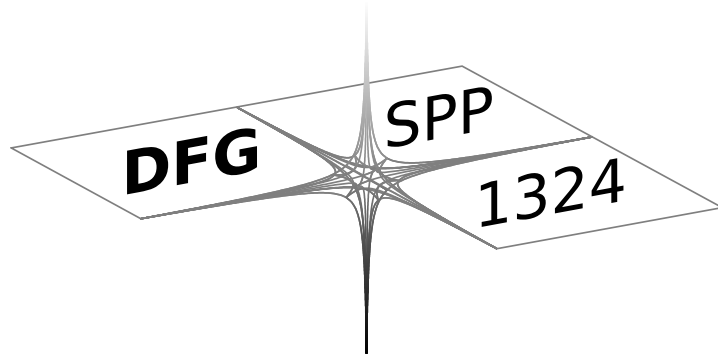
Preprint 163

Edited by

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen”

Error Bounds of MCMC for Functions with Unbounded Stationary Variance

D. Rudolf, N. Schweizer

Preprint 163

The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

ERROR BOUNDS OF MCMC FOR FUNCTIONS WITH UNBOUNDED STATIONARY VARIANCE

DANIEL RUDOLF AND NIKOLAUS SCHWEIZER

ABSTRACT. We prove explicit error bounds for Markov chain Monte Carlo (MCMC) methods to compute expectations of functions with unbounded stationary variance. We assume that there is a $p \in (1, 2)$ so that the functions have finite L_p-norm. For uniformly ergodic Markov chains we obtain error bounds with the optimal order of convergence $n^{1/p-1}$ and if there exists a spectral gap we almost get the optimal order. Further, a burn-in period is taken into account and a recipe for choosing the burn-in is provided.

1. INTRODUCTION

Let $G \subseteq \mathbb{R}^d$ and let $\mathcal{B}(G)$ be the corresponding Borel σ-algebra. We study the problem of computing an expectation of a measurable function, say $f \colon G \to \mathbb{R}$, with respect to a probability measure π. Thus we want to know

$$\mathbb{E}_\pi(f) = \int_G f(x) \, \mathrm{d}\pi(x).$$

We assume that the variance $(\mathbb{E}_\pi(f^2)) - (\mathbb{E}_\pi(f))^2$ is not finite, but that there is a $p \in (1, 2)$ such that

$$\|f\|_p = \left(\int_G |f(x)|^p \, \mathrm{d}\pi(x) \right)^{1/p} < \infty.$$

1

This is the case if f has a singularity, e.g. $f(x) = |x|^{-1/2}$ and π has a bounded strictly positive density over a compact convex set G with $0 \in G$. Here, $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$. Another application is the computation of a second moment of π, say $f(x) = x_d^2$, when the fourth moment of π is infinite due to heavy tails.

Our focus is on situations where π and f are complicated and thus Monte Carlo algorithms are applied. Often one does not have an i.i.d. sequence of random variables with distribution π. For instance, this is the case if π is only known up to a normalizing constant. Such situations naturally arise in Bayesian Statistics and Statistical Physics. Markov Chain Monte Carlo (MCMC) methods are a popular approach for overcoming this problem. The objective of this paper is to provide explicit error bounds of MCMC methods for the computation of $\mathbb{E}_\pi(f)$ when $\mathbb{E}_\pi(f^2) = \|f\|_2^2$ is not finite. So far, the literature has largely focused on cases where $\|f\|_2$ is finite, see [Rud09, BC09, JO10, ŁN11, Rud12, ŁMN13].

The basic idea of MCMC is to approximate π using a Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition kernel K and initial distribution ν. Here π is the stationary and limit distribution. Then, one approximates $\mathbb{E}_\pi(f)$ by

$$(1) \qquad S_{n,n_0}(f) = \frac{1}{n} \sum_{j=1}^{n} f(X_{j+n_0}),$$

where n denotes the number of function evaluations and n_0 the burn-in. The burn-in is the number of steps needed to get sufficiently close to π. For simplicity we assume that the Markov chain is reversible with respect to π. To quantify convergence, we need further assumptions on the Markov chain. We begin with assuming that the Markov chain is uniformly ergodic. Then we relax this assumption to the existence of an (absolute) spectral gap. For reversible Markov chains, existence of such a spectral gap is equivalent to geometric ergodicity, see [RR97]. For $p \in (1, 2]$ we prove bounds on the absolute mean error

$$e_1(S_{n,n_0}, f) = \mathbb{E}\left|S_{n,n_0}(f) - \mathbb{E}_\pi(f)\right|,$$

for functions f, with $\|f\|_p < \infty$. Here, $\mathbb{E}$ denotes the expectation with respect to the distribution of the trajectory of the Markov chain. We consider this absolute error criterion since the root mean square error of S_{n,n_0} is not necessarily finite for $p \in (1, 2)$.

It is known that

$$\inf_{A_n} \sup_{\|f\|_p \leq 1} e_1(A_n, f) \geq c \cdot n^{1/p-1},$$

for some number $c > 0$, where the infimum is taken over all algorithms A_n that use at most n function evaluations of f. For a proof of this fact follow the arguments in [Nov88, Hei94]. Therefore, our bounds cannot decay faster than $n^{1/p-1}$, the optimal order of convergence.

Our absolute mean error bounds satisfy the following properties:

- For uniformly ergodic Markov chains we obtain the optimal order of convergence $n^{1/p-1}$. For Markov chains with spectral gap we come arbitrarily close to the optimal order.
- We quantify the penalty that arises since the initial distribution ν is not the stationary distribution. This penalty appears in our bounds through $\log \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty$, where $\frac{d\nu}{d\pi}$ denotes the density of ν with respect to π. This effect is controlled by the choice of the burn-in n_0. We provide a recipe for choosing n_0.

The main idea of proof is simple and adapted from [Hei94, Proposition 5.4]. The key technique is to apply the interpolation theorem of Riesz-Thorin to the absolute mean error and root mean square error, viewed as operators.

1.1. **Literature overview.** By an ergodic theorem see [MT09, Theorem 17.1.7, p. 427] or [AG11], the MCMC method is well defined, i.e. for a φ-irreducible and Harris recurrent Markov chain

$$\lim_{n \to \infty} S_{n,n_0}(f) = \mathbb{E}_\pi(f)$$

holds almost surely for any f, with finite $\|f\|_1$. For $p \in [2, \infty]$, the ergodic theorem is usually augmented by a central limit theorem, for a survey see [Jon04] and for estimating with confidence see [FJ11]. The corresponding limit theorems for $p \in [0, 2)$ have been less studied, see [CMA13] for recent results and an overview.

Under different convergence assumptions on the Markov chain, various bounds on the mean square error of S_{n,n_0} are known, see for example [Rud09, BC09, JO10, ŁN11, Rud12, ŁMN13]. In particular, for $p \in [2, \infty]$ with finite $\|f\|_p$ in [Rud12, Theorem 3.34 and Theorem 3.41] explicit error bounds of S_{n,n_0} are provided. However, no explicit error bounds are known for $p \in (1, 2)$.

In [Hei94, Proposition 5.4] a simple Monte Carlo method where $(X_n)_{n \in \mathbb{N}}$ is an i.i.d. sequence and X_n has distribution π is studied. A bound of the mean absolute error is provided, namely, for $p \in (1, 2]$ holds

$$(2) \qquad\qquad e_1(S_{n,0}, f) \leq 2^{2/p-1} n^{1/p-1} \|f\|_p.$$

A similar estimate follows by an inequality of v. Bahr and Esseen [vBE65, Theorem 2]. Their bounds work also for some dependent sequences of random variables, but their condition (2) is not satisfied for most Markov chains used for MCMC. Finally let us mention that a slight (nonlinear) modification of the simple Monte Carlo method does also give the optimal rate of the mean square error, for a discussion see [HN02].

1.2. **Outline.** The paper is organized as follows: In the next section we state and discuss our two main results. In Section 3 we provide two applications, the independent Metropolis-Hastings algorithm and the hit-and-run algorithm. Finally, Section 4 contains the proofs. For the reader's convenience the theorem of Riesz-Thorin is stated in the appendix.

2. Main results

By $L_2 = L_2(\pi)$ we denote the set of all square integrable functions with respect to π. For $f \in L_2$ note that the transition kernel K induces the Markov operator

$$(3) \qquad Pf(x) = \int_G f(y)\, K(x, \mathrm{d}y).$$

By

$$\mathrm{gap}(P) = 1 - \|P - \mathbb{E}_\pi\|_{L_2 \to L_2}$$

we denote the spectral gap of the Markov operator. Recall that the transition kernel is reversible with respect to π.

Now we state the error bound for uniformly ergodic Markov chains.

Theorem 1. *Let us assume that we have a Markov chain with transition kernel K and initial distribution ν. Let K be uniformly ergodic, i.e. for an $\alpha \in [0, 1)$ and an $M \in (0, \infty)$ it holds for π-almost all $x \in G$ that*

$$(4) \qquad \|K^n(x, \cdot) - \pi\|_{\mathrm{tv}} \le \alpha^n M,$$

where $\|\cdot\|_{\mathrm{tv}}$ denotes the total variation distance. Further, assume that there exists $\frac{d\nu}{d\pi}$, with finite $\left\|\frac{d\nu}{d\pi}\right\|_\infty$, where $\frac{d\nu}{d\pi}$ denotes the density of ν with respect to π. Let $p \in (1, 2]$ and assume that $n_0 \in \mathbb{N}_0$ satisfies

$$n_0 \ge \frac{\log(2\, M \left\|\frac{d\nu}{d\pi} - 1\right\|_\infty)}{1 - \alpha}.$$

Then

$$e_1(S_{n,n_0}, f) \le \frac{4\,\|f\|_p}{(n \cdot \mathrm{gap}(P))^{1 - 1/p}} + \frac{4\,\|f\|_p}{(n \cdot (1 - \alpha))^{2 - 2/p}}.$$

First, note that uniform ergodicity implies $\mathrm{gap}(P) \ge 1 - \alpha$. The upper bound might be interpreted as follows: The burn-in n_0 is used to decrease the influence of the initial distribution. The number n decreases the error of the averaging procedure. The leading term has the optimal order of convergence $n^{1/p - 1}$ as in (2). The spectral gap appears in the leading term and $1 - \alpha$ appears in the higher

order term. Both quantities describe the price we have to pay for using Markov chains for approximate sampling. If one can sample with respect to π, then $\alpha = 0$, $\text{gap}(P) = 1$ and $\nu = \pi$. Thus $n_0 = 0$ and the error bound is, up to a constant factor, the same as in (2).

Now we state the result for Markov chains which have a spectral gap.

Theorem 2. *Let us assume that we have a Markov chain with transition kernel K and initial distribution ν. Let $\text{gap}(P) > 0$ and further assume that there exists $\frac{d\nu}{d\pi}$, with finite $\left\| \frac{d\nu}{d\pi} \right\|_\infty$, where $\frac{d\nu}{d\pi}$ denotes the density of ν with respect to π. Let $\delta \in (0, 1]$, $p \in (1 + \delta, 2]$ and assume that $n_0 \in \mathbb{N}_0$ satisfies*

$$n_0 \geq \frac{\log(64 \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty) + \log \delta^{-1}}{\delta \cdot \text{gap}(P)}.$$

Then

$$e_1(S_{n,n_0}, f) \leq \frac{4 \left\| f \right\|_p}{(n \cdot \text{gap}(P))^{1 - \frac{1+\delta}{p}}} + \frac{4 \left\| f \right\|_p}{(n \cdot \text{gap}(P))^{2 - \frac{2(1+\delta)}{p}}}.$$

Let us interpret the result. The burn-in n_0 is used to decrease the dependence on the initial distribution and n denotes the sample size of the average procedure. The convergence of the Markov chain is captured by the spectral gap. However, an additional parameter $\delta \in (0, 1]$ appears. This parameter measures a minimal integrability and provides a relation between integrability and convergence of the Markov chain. If a bound on the root mean square error could be derived from $\text{gap}(P) > 0$ for $f \in L_2$, then we would also obtain a bound on the absolute mean error for $\delta = 0$.

For δ close to zero, the rate of convergence in the error bound is arbitrarily close to optimal. But we pay a price. Namely, the burn-in n_0 increases for decreasing δ. There is obviously a trade-off and one might ask for an optimal δ. After some computations by hand one can guess that

$$(5) \qquad \hat{\delta} = \frac{\sqrt{p-1}}{\sqrt{p}} \left(\frac{\log(64 \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty)}{(8\varepsilon^{-1})^{p/(p-1)} \log(8\varepsilon^{-1})} \right)^{1/2},$$

p	δ^*	$N(\delta^*)$	$\hat{\delta}$	$N(\hat{\delta})$
1.1	$6.54 \cdot 10^{-10}$	$5.47 \cdot 10^{19}$	$4.21 \cdot 10^{-11}$	$5.47 \cdot 10^{19}$
1.3	$3.78 \cdot 10^{-4}$	$1.01 \cdot 10^{9}$	$1.48 \cdot 10^{-4}$	$1.03 \cdot 10^{9}$
1.5	$6.58 \cdot 10^{-3}$	$9.17 \cdot 10^{6}$	$3.30 \cdot 10^{-3}$	$9.80 \cdot 10^{6}$

TABLE 1. Required size of the Markov chain sample $N(\delta)$ as prescribed by Theorem 2 for different values of p and $\delta \in \{\delta^*, \hat{\delta}\}$. The parameters are $\|f\|_p = 1$, $\mathrm{gap}(P) = 0.01$, $\left\|\frac{d\nu}{d\pi} - 1\right\|_\infty = 10^{30}$ and $\varepsilon = 0.1$.

ε	δ^*	$N(\delta^*)$	$\hat{\delta}$	$N(\hat{\delta})$
0.01	$1.97 \cdot 10^{-6}$	$1.91 \cdot 10^{13}$	$8.15 \cdot 10^{-7}$	$1.91 \cdot 10^{13}$
0.2	$1.78 \cdot 10^{-3}$	$6.20 \cdot 10^{7}$	$7.23 \cdot 10^{-4}$	$6.59 \cdot 10^{7}$
0.5	$1.11 \cdot 10^{-2}$	$2.57 \cdot 10^{6}$	$6.08 \cdot 10^{-3}$	$2.87 \cdot 10^{6}$

TABLE 2. Required size of the Markov chain sample $N(\delta)$ as prescribed by Theorem 2 for different values of ε and $\delta \in \{\delta^*, \hat{\delta}\}$. The parameters are $\|f\|_p = 1$, $\mathrm{gap}(P) = 0.01$, $\left\|\frac{d\nu}{d\pi} - 1\right\|_\infty = 10^{30}$ and $p = 1.3$.

is a good choice for δ to achieve an error estimate smaller than ε. We justify this heuristic $\hat{\delta}$ as follows: For different values of p and ε we numerically compute δ^* which minimizes the total size of the Markov chain sample $N(\delta) = n(\delta) + n_0(\delta)$ which is needed to obtain an estimate with error ε from Theorem 2. In Table 1 and Table 2 one can see that $\hat{\delta}$ and δ^* have the same behavior for decreasing $\varepsilon \in (0, 1]$ and decreasing $p \in (1, 2)$. Furthermore, the numbers $N(\hat{\delta})$ and $N(\delta^*)$ are quite close. For p close to 1 (Table 1) we also see that the penalty for the lack of integrability leads to a drastic increase in $N(\delta^*)$. This is not surprising, since for p close to 1 also Theorem 1 and (2) exhibit similar behavior. However, for p not too far away from 2, the total size of the Markov chain sample $N(\delta^*)$ and $N(\hat{\delta})$ is reasonable.

3. APPLICATIONS

Now we illustrate our error bounds in two settings. We apply Theorem 1 to the independent Metropolis-Hastings algorithm and Theorem 2 to the hit-and-run algorithm.

3.1. **Independent Metropolis-Hastings algorithm.** Let π be a probability measure with density ρ. We use the independent Metropolis-Hastings algorithm for approximate sampling with respect to π. Let ν be the initial and proposal distribution and denote its density by q. Then, for $x \in G$ and $A \in \mathcal{B}(G)$ the transition kernel is given by

$$M_{\rho,q}(x, A) = \int_A \alpha(x, y) q(y) \, \mathrm{d}y + \mathbf{1}_A(x) \left(1 - \int_G \alpha(x, y) q(y) \, \mathrm{d}y \right),$$

with $\alpha(x, y) = \min \left\{ 1, \frac{\rho(y)q(x)}{\rho(x)q(y)} \right\}$. It is well known that $M_{\rho,q}$ is reversible with respect to π. For $\gamma \in (0, 1]$ and $\kappa \geq \gamma$ let us define

$$\mathcal{R}_{\gamma,\kappa} = \left\{ \rho \colon G \to (0, \infty) \mid \gamma \leq \frac{q(x)}{\rho(x)} \leq \kappa \right\}.$$

By [MT96, Theorem 2.1] we have for $\rho \in \mathcal{R}_{\gamma,\kappa}$ a criterion for uniform ergodicity, i.e.

(6) $$\left\| M_{\rho,q}^n(x, \cdot) - \pi \right\|_{\mathrm{tv}} \leq (1 - \gamma)^n.$$

This leads to the following application of Theorem 1.

Corollary 3. *For $\rho \in \mathcal{R}_{\gamma,\kappa}$ let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain with transition kernel $M_{\rho,q}$ and initial distribution ν. Let $p \in (1, 2]$ and assume that $n_0 \in \mathbb{N}_0$ satisfies*

$$n_0 \geq \gamma^{-1} \log(2\kappa).$$

Then

$$e_1(S_{n,n_0}, f) \leq \frac{4 \, \|f\|_p}{(n\gamma)^{1-1/p}} + \frac{4 \, \|f\|_p}{(n\gamma)^{2-2/p}}.$$

Proof. By (6) the spectral gap is bounded from below by γ. Together with $\frac{d\nu}{d\pi}(x) = \frac{q(x)}{\rho(x)} \leq \kappa$ this implies the error bound. $\qquad\qquad\square$

The result shows that the Metropolis-Hastings algorithm, in a suitable setting, can be used to compute expectations of functions which do not necessarily have finite stationary variance.

3.2. **Hit-and-run algorithm.** Let $G \subset \mathbb{R}^d$ be a measurable set with $0 < \text{vol}_d(G) < \infty$, where vol_d denotes the Lebesgue measure. We assume that G is given by a membership oracle, i.e. we are able to evaluate $\mathbf{1}_G(x)$ for any $x \in \mathbb{R}^d$. The goal is to compute

$$\mathbb{E}_G(f) = \frac{1}{\text{vol}_d(G)} \int_G f(x) \mathrm{d}x$$

for a function $f \colon G \to \mathbb{R}$ with finite $\|f\|_p$ for $p \in (1, 2)$. We impose some additional structure on possible sets G. Namely, let

$$\mathcal{G}_{r,d} = \{G \subset \mathbb{R}^d \mid G \text{ is convex}, B_d \subset G \subset rB_d\},$$

where $r \geq 1$ and $rB_d = \{x \in \mathbb{R}^d \mid |x| \leq r\}$ is the Euclidean ball with radius r. Note that here the input of an approximation scheme is given by a tuple (f, G) with $G \in \mathcal{G}_{r,d}$. The hit-and-run algorithm gives a suitable tool to construct a Markov chain to approximate a sample with uniform distribution in $G \in \mathcal{G}_{r,d}$.

A transition of the hit-and-run algorithm from $x \in G$ works as follows:

(1) Choose a direction, say θ, uniformly distributed on the sphere ∂B_d.

(2) Choose the next state, say $y \in G$, uniformly distributed in $G \cap \{x + \theta r \colon r \in \mathbb{R}\}$.

After randomly choosing a direction θ one samples the next state $y \in G$ uniformly on the line segment determined by the current state x and the direction θ. The resulting hit-and-run algorithm is reversible with respect to the uniform distribution on G.

It is known that the hit-and-run algorithm induces a positive semidefinite Markov operator (on L_2), say H, see [RU13]. Lovaśz and Vempala prove in [LV06, Theorem 4.2, p. 993] a lower bound on the conductance, which implies by the positive semidefiniteness and by the Cheeger inequality, see [LS88], a lower bound on the spectral gap.

Proposition 4. *Let $G \in \mathcal{G}_{r,d}$. Then, the Markov operator H of the hit-and-run algorithm satisfies* $\text{gap}(H) \geq 2^{-51}(dr)^{-2}$.

Now we can apply Theorem 2.

Corollary 5. *Let ν be the uniform distribution on B_d. Let $G \in \mathcal{G}_{r,d}$ and assume that we have a Markov chain with transition kernel, given by the hit-and-run algorithm, and initial distribution ν. Let $\delta \in (0,1]$, $p \in [1+\delta, 2]$ and assume that $n_0 \in \mathbb{N}_0$ satisfies*

$$n_0 \geq 2.3 \cdot 10^{15} (dr)^2 \delta^{-1} (d \log r + \log \delta^{-1} + 4.2).$$

Then

$$e_1(S_{n,n_0}, (f,G)) \leq \frac{4 \left\| f \right\|_p (6.8 \cdot 10^7 dr)^{2(1-\frac{1+\delta}{p})}}{n^{1-\frac{1+\delta}{p}}} + \frac{4 \left\| f \right\|_p (6.8 \cdot 10^7 dr)^{4(1-\frac{1+\delta}{p})}}{n^{2-\frac{2(1+\delta)}{p}}}.$$

Note that the constants in the error bound are rather large. Thus the bound is not useful for simulations. However, it shows that the error behaves only polynomially with respect to the dimension d.

4. Auxiliary results and proofs

Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain with transition kernel K and initial distribution ν. We assume that the transition kernel is reversible with respect to π. Thus π is a stationary distribution.

By (3) the transition kernel induces the Markov operator acting on functions and by

$$\nu P(A) = \int_G K(x, A) \, \nu(\mathrm{d}x), \quad A \in \mathcal{B}(G),$$

it induces the Markov operator acting on signed measures ν on $(G, \mathcal{B}(G))$. If ν is absolutely continuous with respect to π, then also νP is absolutely continuous with respect to π. In particular, $\frac{d(\nu P)}{d\pi} = P(\frac{d\nu}{d\pi})$, for details we refer to [Rud12, Lemma 3.9].

For $p \in [1, \infty]$ we denote by $L_p = L_p(\pi)$ the space of all functions $f \colon G \to \mathbb{R}$ with $\left\| f \right\|_p < \infty$. Note that $\left\| P \right\|_{L_p \to L_p} = 1$ and that $P \colon L_2 \to L_2$ is self-adjoint (because of reversibility).

Next we provide some relations between different convergence properties of reversible Markov chains. If the Markov chain is uniformly ergodic, i.e., if it satisfies (4) for an $\alpha \in [0,1)$ and an $M \in (0,\infty)$, then

$$(7) \qquad \|P^n - \mathbb{E}_\pi\|_{L_p \to L_p} \leq \alpha^n 2M, \quad n \in \mathbb{N},$$

for $p \in \{1, \infty\}$. This implies $\mathrm{gap}(P) \geq 1 - \alpha$. For more details we refer to [Rud12, Proposition 3.24]. Further, by the Riesz-Thorin theorem (see Proposition 11) it follows for $p \in [2, \infty)$ that $\mathrm{gap}(P) > 0$ implies

$$(8)$$
$$\|P^n - \mathbb{E}_\pi\|_{L_p \to L_p} \leq 2^{1-1/p} \|P - \mathbb{E}_\pi\|_{L_2 \to L_2}^{2n/p} = 2^{1-1/p}(1 - \mathrm{gap}(P))^{2n/p}, \quad n \in \mathbb{N}.$$

Now we define a generalized error term of S_{n,n_0} with parameter $p \in [1,2]$ for the computation of $\mathbb{E}_\pi(f)$. Let

$$e_p(S_{n,n_0}, f) := \left(\mathbb{E}\, |S_{n,n_0}(f) - \mathbb{E}_\pi(f)|^p \right)^{1/p}.$$

Note that for $p = 1$ this is the absolute mean error and for $p = 2$ we have the root mean square error. The expectation in the definition of the error is taken with respect to the distribution, say $\mu_{\nu, K}$, of the trajectory $X_1, \dots, X_{n+n_0}$.

4.1. **Proof of Theorem 1.** We prove that under the assumptions of Theorem 1 the inequality

$$(9)$$
$$\sup_{\|f\|_p \leq 1} e_p(S_{n,n_0}, f) \leq 2^{2/p-1} \left(1 + 2\alpha^{n_0} M \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty \right)^{2/p-1}$$
$$\times \left(\frac{2^{1-1/p}}{(n \cdot \mathrm{gap}(P))^{1-1/p}} + \left(\frac{4\,M \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty \alpha^{n_0}}{n^2(1-\alpha)^2} \right)^{1-1/p} \right)$$

holds. From this upper bound the assertion of the theorem follows immediately by $\log \alpha^{-1} \geq 1 - \alpha$ for $\alpha \in [0,1]$ and the choice of the burn-in n_0.

First we state two auxiliary inequalities. By the next proposition we have an upper bound on the root mean square error for $f \in L_2$, see [Rud12, Theorem 3.34]. The next lemma states that the absolute mean error is bounded for $f \in L_1$.

Proposition 6. *Under the assumptions of Theorem 1 we have*

$$\sup_{\|f\|_2 \leq 1} e_2(S_{n,n_0}, f) \leq \frac{\sqrt{2}}{(\,n \cdot \mathrm{gap}(P)\,)^{1/2}} + \frac{2\,M^{1/2}\,\left\|\frac{d\nu}{d\pi} - 1\right\|_\infty^{1/2} \alpha^{n_0/2}}{n(1-\alpha)}.$$

Lemma 7. *Under the assumptions of Theorem 1 we have*

$$\sup_{\|f\|_1 \leq 1} e_1(S_{n,n_0}, f) \leq 2 + 4\alpha^{n_0} M \left\|\frac{d\nu}{d\pi} - 1\right\|_\infty.$$

Proof. By $P^{j+n_0}\left(\frac{d\nu}{d\pi}\right) = 1 + (P^{j+n_0} - \mathbb{E}_\pi)\left(\frac{d\nu}{d\pi} - 1\right)$ and (7) we have

$$e_1(S_{n,n_0}, f)$$

$$\leq \frac{1}{n}\sum_{j=1}^n \mathbb{E}\left|f(X_{j+n_0}) - \mathbb{E}_\pi(f)\right|$$

$$= \frac{1}{n}\sum_{j=1}^n \int_G |f(x) - \mathbb{E}_\pi(f)|\, \frac{d(\nu P^{j+n_0})}{d\pi}(x)\, \mathrm{d}\pi(x)$$

$$= \frac{1}{n}\sum_{j=1}^n \int_G |f(x) - \mathbb{E}_\pi(f)|\, P^{j+n_0}(\frac{d\nu}{d\pi})(x)\, \mathrm{d}\pi(x)$$

$$= \frac{1}{n}\sum_{j=1}^n \left(\|f - \mathbb{E}_\pi(f)\|_1 + \int_G |f(x) - \mathbb{E}_\pi(f)|\,(P^{j+n_0} - \mathbb{E}_\pi)(\frac{d\nu}{d\pi} - 1)(x)\, \mathrm{d}\pi(x)\right)$$

$$\leq 2\,\|f\|_1\, \frac{1}{n}\sum_{j=1}^n \left(1 + 2M\alpha^{j+n_0}\left\|\frac{d\nu}{d\pi} - 1\right\|_\infty\right)$$

$$\leq 2\,\|f\|_1 \left(1 + 2\alpha^{n_0} M \left\|\frac{d\nu}{d\pi} - 1\right\|_\infty\right).$$

$$\square$$

Now we prove (9). We consider the linear operator $T\colon L_p(\pi) \to L_p(\mu_{\nu,K})$ defined by

$$(10) \qquad\qquad T(f) = S_{n,n_0}(f) - \mathbb{E}_\pi(f).$$

Recall that $\mu_{\nu,K}$ denotes the distribution of the sample trajectory. Further, note that

$$\|T\|_{L_p(\pi)\to L_p(\mu_{\nu,K})} = \sup_{\|f\|_p\le 1} e_p(S_{n,n_0}, f).$$

By Proposition 6 and Lemma 7 we obtain

$$\|T\|_{L_1(\pi)\to L_1(\mu_{\nu,K})} \le M_1 \quad \text{and} \quad \|T\|_{L_2(\pi)\to L_2(\mu_{\nu,K})} \le M_2,$$

with

$$M_1 = 2 + 4\alpha^{n_0} M \left\|\frac{d\nu}{d\pi} - 1\right\|_\infty,$$

$$M_2 = \frac{\sqrt{2}}{(n\cdot\text{gap}(P))^{1/2}} + \frac{2\,M^{1/2}\left\|\frac{d\nu}{d\pi}-1\right\|_\infty^{1/2}\alpha^{n_0/2}}{n(1-\alpha)}.$$

The application of the Riesz-Thorin theorem (see Proposition 11) leads to

$$\|T\|_{L_p(\pi)\to L_p(\mu_{\nu,K})} \le M_1^{1-\theta} M_2^{\theta}$$

with $\theta = 2 - 2/p$. The assertion of (9) follows by $(x+y)^r \le x^r + y^r$ for $x, y \ge 0$ and $r \in [0,1]$.

4.2. **Proof of Theorem 2.** We prove that for any $\delta \in (0,1]$ and $p \in [1+\delta, 2]$ by the assumptions of Theorem 2

(11)
$$\sup_{\|f\|_p\le 1} e_1(S_{n,n_0}, f) \le \left(2 + 4(1-\text{gap}(P))^{2\frac{n_0\delta}{1+\delta}}\left\|\frac{d\nu}{d\pi}-1\right\|_\infty\right)^{2\frac{1+\delta}{p}-1}$$

$$\times\left(\frac{2^{1-\frac{1+\delta}{p}}}{(n\cdot\text{gap}(P))^{1-\frac{1+\delta}{p}}} + \frac{\left(64\frac{1+\delta}{\delta}\left\|\frac{d\nu}{d\pi}-1\right\|_\infty(1-\text{gap}(P))^{2\frac{n_0\delta}{1+\delta}}\right)^{1-\frac{1+\delta}{p}}}{(n^2\cdot\text{gap}(P)^2)^{1-\frac{1+\delta}{p}}}\right).$$

From this upper bound and the choice

$$n_0 \ge \frac{1+\delta}{2\delta}\cdot\frac{\log\left(\frac{32(1+\delta)}{\delta}\left\|\frac{d\nu}{d\pi}-1\right\|_\infty\right)}{\log(1-\text{gap}(P))^{-1}}.$$

the assertion of the theorem follows by taking $\log(1 - \text{gap}(P))^{-1} \geq \text{gap}(P)$ and $\delta \in (0, 1]$ into account.

First we state two auxiliary inequalities with parameters $p_1 \in [1, 2]$ and $p_2 \in (2, 4]$. By the next proposition we have an upper bound on the root mean square error for $f \in L_{p_2}$, see [Rud12, Theorem 3.41]. The next lemma states that the absolute mean error is bounded for $f \in L_{p_1}$.

Proposition 8. *Under the assumptions of Theorem 2 we have*

$$\sup_{\|f\|_{p_2} \leq 1} e_2(S_{n, n_0}, f) \leq \frac{\sqrt{2}}{(n \cdot \text{gap}(P))^{1/2}} + \frac{8\sqrt{p_2}}{\sqrt{p_2 - 2}} \cdot \frac{\left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty^{1/2} (1 - \text{gap}(P))^{n_0 \frac{p_2 - 2}{p_2}}}{n \cdot \text{gap}(P)}.$$

Lemma 9. *Under the assumptions of Theorem 2 we have*

$$\sup_{\|f\|_{p_1} \leq 1} e_1(S_{n, n_0}, f) \leq 2 + 4 \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty (1 - \text{gap}(P))^{2 \frac{p_1 - 1}{p_1} n_0}.$$

Proof. By the same steps as in the proof of Lemma 7 we obtain the assertion with Hölder's inequality and (8). $\square$

By the last two inequalities we can apply similar interpolation arguments as in the proof of Theorem 1. We obtain the following:

Lemma 10. *Let $p_1 \in [1, 2]$, $p_2 \in (2, 4]$ and $p \in [p_1, p_2]$. Then, under the assumptions of Theorem 2 we have*

$$\sup_{\|f\|_p \leq 1} e_q(S_{n, n_0}, f) \leq M_1^{\frac{p_1}{(p_2 - p_1)} \left(\frac{p_2}{p} - 1 \right)} \cdot M_2^{\frac{p_2}{p_2 - p_1} \left(1 - \frac{p_1}{p} \right)}$$

with

$$q = 1 + \frac{p_2(p - p_1)}{p_2(p + p_1) - 2pp_1} \in [1, 2],$$

$$M_1 = 2 + 4 \left\| \frac{d\nu}{d\pi} - 1 \right\|_\infty (1 - \text{gap}(P))^{2 \frac{p_1 - 1}{p_1} n_0},$$

$$M_2 = \frac{\sqrt{2}}{(n \cdot \mathrm{gap}(P))^{1/2}} + \frac{8\sqrt{p_2}}{\sqrt{p_2 - 2}} \cdot \frac{\left\|\frac{d\nu}{d\pi} - 1\right\|_\infty^{1/2} (1 - \mathrm{gap}(P))^{n_0 \frac{p_2 - 2}{p_2}}}{n \cdot \mathrm{gap}(P)}.$$

Proof. We consider the linear operator $T \colon L_p(\pi) \to L_q(\mu_{\nu,K})$ defined by (10). Note that

$$\|T\|_{L_p(\pi) \to L_q(\mu_{\nu,K})} = \sup_{\|f\|_p \leq 1} e_q(S_{n,n_0}, f).$$

By Lemma 9 and Proposition 8 we have

$$\|T\|_{L_{p_1}(\pi) \to L_1(\mu_{\nu,K})} \leq M_1 \quad \text{and} \quad \|T\|_{L_{p_2}(\pi) \to L_2(\mu_{\nu,K})} \leq M_2.$$

By the Riesz-Thorin theorem (see Proposition 11) $\|T\|_{L_p(\pi) \to L_q(\mu_{\nu,K})} \leq M_1^{1-\theta} M_2^\theta$ holds for $\theta \in [0,1]$ satisfying $p^{-1} = (1-\theta)p_1^{-1} + \theta p_2^{-1}$ and $q^{-1} = 1 - \frac{\theta}{2}$. $\square$

Note that for $q \in [1, 2]$ we have $e_1(S_{n,n_0}, f) \leq e_q(S_{n,n_0}, f)$. Thus, the proof of (11) follows by an application of Lemma 10 with $p_1 = 1 + \delta$ and $p_2 = 2(1 + \delta)$.

Appendix A. Riesz-Thorin interpolation theorem

Let $(G, \mathcal{G}, \pi)$ and $(\Omega, \mathcal{F}, \mu)$ be probability spaces. Let $p \in [1, \infty]$ and let $L_p(\pi)$ be the space of $\mathcal{G}$-measurable functions $g \colon G \to \mathbb{R}$ with $\|g\|_{p,\pi} = \left(\int_G |g(x)|^p \, d\pi(x)\right)^{1/p} < \infty$ and let $L_p(\mu)$ be the space of $\mathcal{F}$-measurable functions $f \colon \Omega \to \mathbb{R}$ with $\|f\|_{p,\mu} = \left(\int_\Omega |f(x)|^p \, d\mu(x)\right)^{1/p} < \infty$. In the following we formulate a version of the theorem of Riesz-Thorin. For details we refer to [BS88, Chapter 4: Corollary 1.8, Excercise 5, Corollary 2.3].

Proposition 11 (Riesz-Thorin theorem)**.** *Let $1 \leq p_k \leq q_k \leq \infty$ for $k = 1, 2$. We assume that $\theta \in [0, 1]$ and*

$$\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}, \qquad \frac{1}{q} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}.$$

Let T be a linear operator from $L_{p_1}(\pi)$ to $L_{q_1}(\mu)$ and at the same time from $L_{p_2}(\pi)$ to $L_{q_2}(\mu)$ with

$$\|T\|_{L_{p_1}(\pi) \to L_{q_1}(\mu)} \leq M_1, \qquad \|T\|_{L_{p_2}(\pi) \to L_{q_2}(\mu)} \leq M_2.$$

Then

$$(12) \qquad \|T\|_{L_p(\pi) \to L_q(\mu)} \leq M_1^{1-\theta} M_2^{\theta}$$

and if T is a positive operator, i.e. $f \geq 0$ implies $Tf \geq 0$, then (12) holds for all $1 \leq p_k, q_k \leq \infty$ with $k = 1, 2$.

References

[AG11]   S. Asmussen and P. W. Glynn, *A new proof of convergence of mcmc via the ergodic theorem*, Statistics & Probability Letters **81** (2011), no. 10, 1482–1485.

[BC09]   A. Belloni and V. Chernozhukov, *On the computational complexity of MCMC-based estimators in large samples*, Ann. Statist. **37** (2009), no. 4, 2011–2055.

[BS88]   C. Bennett and R. Sharpley, *Interpolation of operators*, Pure and Applied Mathematics, vol. 129, Academic Press Inc., Boston, MA, 1988.

[CMA13] P. Cattiaux and M. Manou-Abi, *Limit theorems for some funtionals with heavy tails of a discrete time Markov chain*, ESAIM: Probability and Statistics **eFirst** (2013).

[FJ11]   M. Flegal and G. Jones, *Implementing MCMC: Estimating with confidence*, in Handbook of Markov chain Monte Carlo, Chapman & Hall (2011), 175–197.

[Hei94] S. Heinrich, *Random approximation in numerical analysis*, Proceedings of the Conference "Functional Analysis" Essen (1994), 123–171.

[HN02]   S. Heinrich and E. Novak, *Optimal summation and integration by deterministic, randomized, and quantum algorithms*, Monte Carlo and Quasi-Monte Carlo Methods 2000, K.-T. Fang, F. J. Hickernell, H. Niederreiter (Eds.), Springer Verlag (2002), 50–62.

[JO10]   A. Joulin and Y. Ollivier, *Curvature, concentration and error estimates for Markov chain Monte Carlo*, Ann. Probab. **38** (2010), no. 6, 2418–2442.

[Jon04] G. Jones, *On the Markov chain central limit theorem*, Probab. Surv. **1** (2004), 299–320.

[ŁMN13] K. Łatuszynski, B. Miasojedow, and W. Niemiro, *Nonasymptotic bounds on the estimation error of mcmc algorithms*, Bernoulli **19** (2013), no. 5A, 20133–2066.

[ŁN11]   K. Łatuszyński and W. Niemiro, *Rigorous confidence bounds for MCMC under a geometric drift condition*, J. Complexity **27** (2011), no. 1, 23–38.

[LS88]   G. Lawler and A. Sokal, *Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality*, Trans. Amer. Math. Soc. **309** (1988), no. 2, 557–580.

[LV06]   L. Lovász and S. Vempala, *Hit-and-run from a corner*, SIAM J. Comput. **35** (2006), no. 4, 985–1005.

[MT96]   K. Mengersen and R. Tweedie, *Rates of convergence of the Hastings and Metropolis algorithms*, Ann. Statist. **24** (1996), no. 1, 101–121.

[MT09]   S. Meyn and R. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, 2009.

[Nov88] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lecture Notes in Mathematics 1349, Springer: Berlin, Heidelberg, New York, London, Paris and Tokyo, 1988.

[RR97]   G. Roberts and J. Rosenthal, *Geometric ergodicity and hybrid Markov chains*, Electron. Comm. Probab. **2** (1997), no. 2, 13–25.

[RU13]   D. Rudolf and M. Ullrich, *Positivity of hit-and-run and related algorithms*, Electron. Commun. Probab. **18** (2013), 1–8.

[Rud09] D. Rudolf, *Explicit error bounds for lazy reversible Markov chain Monte Carlo*, J. Complexity **25** (2009), no. 1, 11–24.

[Rud12] ———, *Explicit error bounds for Markov chain Monte Carlo*, Dissertationes Math. **485** (2012), 93 pp.

[vBE65] B. v. Bahr and C. Esseen, *Inequalities for the rth absolute moment of a sum of random variables,* $1 \leq r \leq 2$, Ann. Math. Statist. **36** (1965), 299–303.

INST. FÜR MATH., UNIVERSITÄT JENA, ERNST-ABBE-PLATZ 2, 07743 JENA, GERMANY

E-mail address: `daniel-rudolf@uni-jena.de`

FR MATHEMATIK, UNIVERSITÄT DES SAARLANDES POSTFACH 151150, 66041 SAARBRÜCKEN, GERMANY

E-mail address: `schweizer@math.uni-sb.de`

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multilevel Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak and H. Woźniakowski. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.

[69] G. Kutyniok and W.-Q Lim. Image Separation Using Shearlets. Preprint 69, DFG-SPP 1324, November 2010.

[70] B. Jin and P. Maass. An Analysis of Electrical Impedance Tomography with Applications to Tikhonov Regularization. Preprint 70, DFG-SPP 1324, December 2010.

[71] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT Format. Preprint 71, DFG-SPP 1324, December 2010.

[72] T. Müller-Gronbach and K. Ritter. A Local Refinement Strategy for Constructive Quantization of Scalar SDEs. Preprint 72, DFG-SPP 1324, December 2010.

[73] T. Rohwedder and R. Schneider. An Analysis for the DIIS Acceleration Method used in Quantum Chemistry Calculations. Preprint 73, DFG-SPP 1324, December 2010.

[74] C. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. Preprint 74, DFG-SPP 1324, December 2010.

[75] C. Bender. Primal and Dual Pricing of Multiple Exercise Options in Continuous Time. Preprint 75, DFG-SPP 1324, December 2010.

[76] H. Harbrecht, M. Peters, and R. Schneider. On the Low-rank Approximation by the Pivoted Cholesky Decomposition. Preprint 76, DFG-SPP 1324, December 2010.

[77] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Adaptive Wavelet Methods for Elliptic Stochastic Partial Differential Equations. Preprint 77, DFG-SPP 1324, January 2011.

[78] G. Plonka, S. Tenorth, and A. Iske. Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform. Preprint 78, DFG-SPP 1324, January 2011.

[79] A. Mugler and H.-J. Starkloff. On Elliptic Partial Differential Equations with Random Coefficients. Preprint 79, DFG-SPP 1324, January 2011.

[80] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. A Derandomization of the Euler Scheme for Scalar Stochastic Differential Equations. Preprint 80, DFG-SPP 1324, January 2011.

[81] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. Preprint 81, DFG-SPP 1324, January 2011.

[82] K. Grella and C. Schwab. Sparse Tensor Spherical Harmonics Approximation in Radiative Transfer. Preprint 82, DFG-SPP 1324, January 2011.

[83] D.A. Lorenz, S. Schiffler, and D. Trede. Beyond Convergence Rates: Exact Inversion With Tikhonov Regularization With Sparsity Constraints. Preprint 83, DFG-SPP 1324, January 2011.

[84] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. Preprint 84, DFG-SPP 1324, January 2011.

[85] S. Dahlke and W. Sickel. On Besov Regularity of Solutions to Nonlinear Elliptic Partial Differential Equations. Preprint 85, DFG-SPP 1324, January 2011.

[86] S. Dahlke, U. Friedrich, P. Maass, T. Raasch, and R.A. Ressel. An adaptive wavelet method for parameter identification problems in parabolic partial differential equations. Preprint 86, DFG-SPP 1324, January 2011.

[87] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and Variational Stabilization for Convection-Diffusion Equations. Preprint 87, DFG-SPP 1324, January 2011.

[88] T. Jahnke. On Reduced Models for the Chemical Master Equation. Preprint 88, DFG-SPP 1324, January 2011.

[89] P. Binev, W. Dahmen, R. DeVore, P. Lamby, D. Savu, and R. Sharpley. Compressed Sensing and Electron Microscopy. Preprint 89, DFG-SPP 1324, March 2011.

[90] P. Binev, F. Blanco-Silva, D. Blom, W. Dahmen, P. Lamby, R. Sharpley, and T. Vogt. High Quality Image Formation by Nonlocal Means Applied to High-Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). Preprint 90, DFG-SPP 1324, March 2011.

[91] R. A. Ressel. A Parameter Identification Problem for a Nonlinear Parabolic Differential Equation. Preprint 91, DFG-SPP 1324, May 2011.

[92] G. Kutyniok. Data Separation by Sparse Representations. Preprint 92, DFG-SPP 1324, May 2011.

[93] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to Compressed Sensing. Preprint 93, DFG-SPP 1324, May 2011.

[94] H.-C. Kreusler and H. Yserentant. The Mixed Regularity of Electronic Wave Functions in Fractional Order and Weighted Sobolev Spaces. Preprint 94, DFG-SPP 1324, June 2011.

[95] E. Ullmann, H. C. Elman, and O. G. Ernst. Efficient Iterative Solvers for Stochastic Galerkin Discretizations of Log-Transformed Random Diffusion Problems. Preprint 95, DFG-SPP 1324, June 2011.

[96] S. Kunis and I. Melzer. On the Butterfly Sparse Fourier Transform. Preprint 96, DFG-SPP 1324, June 2011.

[97] T. Rohwedder. The Continuous Coupled Cluster Formulation for the Electronic Schrödinger Equation. Preprint 97, DFG-SPP 1324, June 2011.

[98] T. Rohwedder and R. Schneider. Error Estimates for the Coupled Cluster Method. Preprint 98, DFG-SPP 1324, June 2011.

[99] P. A. Cioica and S. Dahlke. Spatial Besov Regularity for Semilinear Stochastic Partial Differential Equations on Bounded Lipschitz Domains. Preprint 99, DFG-SPP 1324, July 2011.

[100] L. Grasedyck and W. Hackbusch. An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples. Preprint 100, DFG-SPP 1324, August 2011.

[101] N. Chegini, S. Dahlke, U. Friedrich, and R. Stevenson. Piecewise Tensor Product Wavelet Bases by Extensions and Approximation Rates. Preprint 101, DFG-SPP 1324, September 2011.

[102] S. Dahlke, P. Oswald, and T. Raasch. A Note on Quarkonial Systems and Multilevel Partition of Unity Methods. Preprint 102, DFG-SPP 1324, September 2011.

[103] A. Uschmajew. Local Convergence of the Alternating Least Squares Algorithm For Canonical Tensor Approximation. Preprint 103, DFG-SPP 1324, September 2011.

[104] S. Kvaal. Multiconfigurational time-dependent Hartree method for describing particle loss due to absorbing boundary conditions. Preprint 104, DFG-SPP 1324, September 2011.

[105] M. Guillemard and A. Iske. On Groupoid C*-Algebras, Persistent Homology and Time-Frequency Analysis. Preprint 105, DFG-SPP 1324, September 2011.

[106] A. Hinrichs, E. Novak, and H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Preprint 106, DFG-SPP 1324, September 2011.

[107] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander. Efficient Analysis of High Dimensional Data in Tensor Formats. Preprint 107, DFG-SPP 1324, September 2011.

[108] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider. Optimization Problems in Contracted Tensor Networks. Preprint 108, DFG-SPP 1324, October 2011.

[109] S. Dereich, T. Müller-Gronbach, and K. Ritter. On the Complexity of Computing Quadrature Formulas for SDEs. Preprint 109, DFG-SPP 1324, October 2011.

[110] D. Belomestny. Solving optimal stopping problems by empirical dual optimization and penalization. Preprint 110, DFG-SPP 1324, November 2011.

[111] D. Belomestny and J. Schoenmakers. Multilevel dual approach for pricing American style derivatives. Preprint 111, DFG-SPP 1324, November 2011.

[112] T. Rohwedder and A. Uschmajew. Local convergence of alternating schemes for optimization of convex problems in the TT format. Preprint 112, DFG-SPP 1324, December 2011.

[113] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. Preprint 113, DFG-SPP 1324, December 2011.

[114] Y. Dong, T. Görner, and S. Kunis. An iterative reconstruction scheme for photoacoustic imaging. Preprint 114, DFG-SPP 1324, December 2011.

[115] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along generated sets. Preprint 115, DFG-SPP 1324, February 2012.

[116] H. Chen and R. Schneider. Numerical analysis of augmented plane waves methods for full-potential electronic structure calculations. Preprint 116, DFG-SPP 1324, February 2012.

[117] J. Ma, G. Plonka, and M.Y. Hussaini. Compressive Video Sampling with Approximate Message Passing Decoding. Preprint 117, DFG-SPP 1324, February 2012.

[118] D. Heinen and G. Plonka. Wavelet shrinkage on paths for scattered data denoising. Preprint 118, DFG-SPP 1324, February 2012.

[119] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. Preprint 119, DFG-SPP 1324, March 2012.

[120] C. Bender and J. Steiner. A-posteriori estimates for backward SDEs. Preprint 120, DFG-SPP 1324, April 2012.

[121] M. Espig, W. Hackbusch, A. Litvinenkoy, H.G. Matthiesy, and P. Wähnert. Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Preprint 121, DFG-SPP 1324, May 2012.

[122] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. Preprint 122, DFG-SPP 1324, June 2012.

[123] A. Mugler and H.-J. Starkloff. On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. Preprint 123, DFG-SPP 1324, June 2012.

[124] P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R.L. Schilling. On the convergence analysis of Rothe's method. Preprint 124, DFG-SPP 1324, July 2012.

[125] P. Binev, A. Cohen, W. Dahmen, and R. DeVore. Classification Algorithms using Adaptive Partitioning. Preprint 125, DFG-SPP 1324, July 2012.

[126] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation of hierarchical Tucker and Tensor-Train tensors. Preprint 126, DFG-SPP 1324, July 2012.

[127] M. Kovács, S. Larsson, and K. Urban. On Wavelet-Galerkin methods for semilinear parabolic equations with additive noise. Preprint 127, DFG-SPP 1324, August 2012.

[128] M. Bachmayr, H. Chen, and R. Schneider. Numerical analysis of Gaussian approximations in quantum chemistry. Preprint 128, DFG-SPP 1324, August 2012.

[129] D. Rudolf. Explicit error bounds for Markov chain Monte Carlo. Preprint 129, DFG-SPP 1324, August 2012.

[130] P.A. Cioica, K.-H. Kim, K. Lee, and F. Lindner. On the $L_q(L_p)$-regularity and Besov smoothness of stochastic parabolic equations on bounded Lipschitz domains. Preprint 130, DFG-SPP 1324, December 2012.

[131] M. Hansen. $n-$term Approximation Rates and Besov Regularity for Elliptic PDEs on Polyhedral Domains. Preprint 131, DFG-SPP 1324, December 2012.

[132] R. E. Bank and H. Yserentant. On the H^1 -stability of the L_2 -projection onto finite element spaces. Preprint 132, DFG-SPP 1324, December 2012.

[133] M. Gnewuch, S. Mayer, and K. Ritter. On Weighted Hilbert Spaces and Integration of Functions of Infinitely Many Variables. Preprint 133, DFG-SPP 1324, December 2012.

[134] D. Crisan, J. Diehl, P.K. Friz, and H. Oberhauser. Robust Filtering: Correlated Noise and Multidimensional Observation. Preprint 134, DFG-SPP 1324, January 2013.

[135] Wolfgang Dahmen, Christian Plesken, and Gerrit Welper. Double Greedy Algorithms: Reduced Basis Methods for Transport Dominated Problems. Preprint 135, DFG-SPP 1324, February 2013.

[136] Aicke Hinrichs, Erich Novak, Mario Ullrich, and Henryk Wozniakowski. The Curse of Dimensionality for Numerical Integration of Smooth Functions. Preprint 136, DFG-SPP 1324, February 2013.

[137] Markus Bachmayr, Wolfgang Dahmen, Ronald DeVore, and Lars Grasedyck. Approximation of High-Dimensional Rank One Tensors. Preprint 137, DFG-SPP 1324, March 2013.

[138] Markus Bachmayr and Wolfgang Dahmen. Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations. Preprint 138, DFG-SPP 1324, April 2013.

[139] Felix Lindner. Singular Behavior of the Solution to the Stochastic Heat Equation on a Polygonal Domain. Preprint 139, DFG-SPP 1324, May 2013.

[140] Stephan Dahlke, Dominik Lellek, Shiu Hong Lui, and Rob Stevenson. Adaptive Wavelet Schwarz Methods for the Navier-Stokes Equation. Preprint 140, DFG-SPP 1324, May 2013.

[141] Jonas Ballani and Lars Grasedyck. Tree Adaptive Approximation in the Hierarchical Tensor Format. Preprint 141, DFG-SPP 1324, June 2013.

[142] Harry Yserentant. A short theory of the Rayleigh-Ritz method. Preprint 142, DFG-SPP 1324, July 2013.

[143] M. Hefter and K. Ritter. On Embeddings of Weighted Tensor Product Hilbert Spaces. Preprint 143, DFG-SPP 1324, August 2013.

[144] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts. Preprint 144, DFG-SPP 1324, August 2013.

[145] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate functions by trigonometric polynomials based on rank-1 lattice sampling. Preprint 145, DFG-SPP 1324, September 2013.

[146] C. Bender, N. Schweizer, and J. Zhuo. A primal-dual algorithm for BSDEs. Preprint 146, DFG-SPP 1324, October 2013.

[147] D. Rudolf. Hit-and-run for numerical integration. Preprint 147, DFG-SPP 1324, October 2013.

[148] D. Rudolf and M. Ullrich. Positivity of hit-and-run and related algorithms. Preprint 148, DFG-SPP 1324, October 2013.

[149] L. Grasedyck, M. Kluge, and S. Krämer. Alternating Directions Fitting (ADF) of Hierarchical Low Rank Tensors. Preprint 149, DFG-SPP 1324, October 2013.

[150] F. Filbir, S. Kunis, and R. Seyfried. Effective discretization of direct reconstruction schemes for photoacoustic imaging in spherical geometries. Preprint 150, DFG-SPP 1324, November 2013.

[151] E. Novak, M. Ullrich, and H. Woźniakowski. Complexity of Oscillatory Integration for Univariate Sobolev Spaces. Preprint 151, DFG-SPP 1324, November 2013.

[152] A. Hinrichs, E. Novak, and M. Ullrich. A Tractability Result for the Clenshaw Curtis Smolyak Algorithm. Preprint 152, DFG-SPP 1324, November 2013.

[153] M. Hein, S. Setzer, L. Jost, and S. Rangapuram. The Total Variation on Hypergraphs - Learning on Hypergraphs Revisited. Preprint 153, DFG-SPP 1324, November 2013.

[154] M. Kovács, S. Larsson, and F. Lindgren. On the Backward Euler Approximation of the Stochastic Allen-Chan Equation. Preprint 154, DFG-SPP 1324, November 2013.

[155] S. Dahlke, M. Fornasier, U. Friedrich, and T. Raasch. Multilevel preconditioning for sparse optimization of functionals with nonconvex fidelity terms. Preprint 155, DFG-SPP 1324, December 2013.

[156] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. On the complexity of computing quadrature formulas for marginal distributions of SDEs. Preprint 156, DFG-SPP 1324, January 2014.

[157] M. Giles, T. Nagapetyan, and K. Ritter. Multi-Level Monte Carlo Approximation of Distribution Functions and Densities. Preprint 157, DFG-SPP 1324, February 2014.

[158] F. Dickmann and N. Schweizer. Faster Comparison of Stopping Times by Nested Conditional Monte Carlo . Preprint 158, DFG-SPP 1324, February 2014.

[159] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on sampling along rank-1 lattice with generating vector of Korobov form . Preprint 159, DFG-SPP 1324, February 2014.

[160] S. Dereich and T. Müller-Gronbach. Quadrature for Self-Affine Distributions on $\mathbb{R}^d$. Preprint 160, DFG-SPP 1324, March 2014.

[161] S. Dereich and S. Li. Multilevel Monte Carlo for Lévy-driven SDEs: central limit theorems for adaptive Euler schemes. Preprint 161, DFG-SPP 1324, March 2014.

[162] M. Kluge. Sampling Rules for Tensor Reconstruction in Hierarchical Tucker Format. Preprint 162, DFG-SPP 1324, April 2014.

[163] D. Rudolf and N. Schweizer. Error Bounds of MCMC for Functions with Unbounded Stationary Variance. Preprint 163, DFG-SPP 1324, April 2014.