# Performance Bounds for Dispersion Analysis: A Comparison Between Monte Carlo and Perron-Frobenius Operator Approach

Abhishek Halder*, Raktim Bhattacharya*

April 25, 2011

**Abstract**

We compare computational cost and accuracy between two different approaches of dispersion analysis. One is the conventional Monte Carlo method and the other being the Perron-Frobenius operator approach, that directly propagates the joint probability density function using Liouville equation. It is shown that with same computational budget, Perron-Frobenius operator approach rewards better accuracy than Monte-Carlo based dispersion analysis. In particular, we show that propagation of uncertainty through the Perron-Frobenius operator is exact in the sense that the joint density computation incurs no more than the path integration error. On contrary, the rate of approximation from Monte Carlo simulation, has fractional decay. We also establish performance guarantees for approximating marginals from the joint density and obtain the optimal approximation algorithm for piecewise constant approximating class. These results justify why Perron-Frobenius operator outperforms Monte Carlo, as observed numerically in uncertainty propagation setting [Halder and Bhattacharya, 2011] and in nonlinear filtering setting [Dutta and Bhattacharya, 2011].

## 1 Introduction

Dispersion analysis refers to nonparametric uncertainty quantification in dynamical systems. With the increasing complexity of dynamical systems being studied in science and engineering, methods for quantification and systematic propagation of uncertainty, have gained considerable research interest for guaranteeing robust performance in safety critical operations. Some representative applications include risk assessment in the atmospheric transport of chemical, biological, radiological and nuclear (CBRN) hazards [Terejanu *et. al.*, 2007], weather forecasting [Poroseva *et. al.*, 2006], planetary landing of rovers [Knocke *et. al.*, 2004], uncertainty evolution in fluids [Galbally *et. al.*, 2010; Pettersson *et. al.*, 2009] and materials [Ghanem and Spanos, 2003; Greene *et. al.*, 2011]. The most popular way of doing this is Monte Carlo (MC) simulation, where one simulates the dynamics with randomly sampled initial conditions and constructs a histogram at each snapshot of interest. One may then look at the constructed static histogram as an approximation of the underlying probability density function (PDF). However, the difficulty with this approach is that it's computationally too expensive for simulating large-scale systems, typically those with nonlinear dynamics involving uncertainties in many states and parameters. Another potential concern is that the method of constructing histogram as a post-processing step, is a piecewise constant approximation that suffers from 'curse of dimensionality' [Bellman, 1957]. Hence in addition to runtime complexity, the accuracy of histogram approximation exacerbates with the increasing dimensionality of the system.

With the emergence of uncertainty quantification as a significant thread of research in scientific computing, many alternative methods have been proposed for dispersion analysis. One prominent candidate is the polynomial chaos (PC) method [Soize and Ghanem, 2004] where one derives a set of deterministic ordinary differential equations (ODEs) using either Galerkin projection [Spanos and Ghanem, 1989] or stochastic collocation [Xiu and Hesthaven, 2005]. Although PC have been implemented successfully in many engineering applications [Pettit, 2004; Mathelin *et. al.*, 2005, Prabhakar *et. al.*, 2010; Dutta and Bhattacharya, 2010], one bottleneck is that deriving the deterministic ODEs from PC expansions is difficult for generic nonlinear dynamical systems [Debusschere *et. al.*, 2004]. Furthermore, the resulting set of deterministic coupled nonlinear ODEs need to be solved in a higher dimensional state space. For example, if the original state space dimension is $d$ (i.e. the dynamics is given by $d$ first order ODEs) and each state is approximated through $N$-term approximation via PC expansion, then we must solve total $Nd$ ODEs to obtain the PC coefficients. To ensure good convergence, $N$ need to be large. Hence

*Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843-3141 (ahalder@tamu.edu, raktim@tamu.edu)

for large $d$, the problem becomes intractable for simulation. Moreover, the finite dimensional approximation of the probability space degrades the computational performance if long term statistics is desired.

Other than PC, frameworks for uncertainty quantification include concentration-of-measure inequalities [Lucas *et. al.*, 2008], statistical linearization [Roberts and Spanos, 2003] and moment closure methods [Bergman *et. al.*, 1994]. The first only provides a margin of uncertainty, the applicability of the second depends on the nonlinearity of the dynamics and parametric methods like moment closure lack accuracy compared to nonparametric methods which compute evolution of the entire PDF instead of first few moments of it. Due to the lack of availability of nonparametric methods which can guarantee accuracy in high-dimensional nonlinear setting, coupled with its ease of implementation, MC is still preferred in many industry scale and mission specific simulations [Powell *et. al.*, 2000; Balaram *et. al.*, 2002].

Borrowing ideas from statistical physics, in [Halder and Bhattacharya, 2011] the authors have shown that the nonparametric propagation of uncertainty can be computed *exactly* by solving Liouville partial differential equation associated with the Perron-Frobenius (PF) operator [Lasota and Mackey, 1994]. The proposed technique is easy to implement through method-of-characteristics (MOC) and hence from a practitioner's perspective, existing MC codes can quickly transition to the Perron-Frobenius operator formalism. Further, since each characteristic can be computed independent of the other, the MOC implementation is suitable for parallelization. The formulation was incorporated in NASA Jet Propulsion Laboratory's Dynamics Simulator for Entry, Descent and Surface Landing (DESENDS) platform [Balaram *et. al.*, 2010] for dispersion analysis in planetary exploration.

Although it's apparent that an *exact* computation of the spatio-temporal evolution of the PDF based on Perron-Frobenius operator formulation outperforms an *approximate* computation of the same based on MC histogram, as evidenced by the numerical results in [Halder and Bhattacharya, 2011], one may still ask the margin of accuracy leveraged by the proposed method compared to MC. Further, no rigorous results were known on whether the MOC computation overhead affects the runtime complexity for calculating exact PDF compared to MC based histogram construction. The purpose of this article is to show that with same computational budget, the PF operator formalism enables exact computation while MC compromises with the accuracy.

### 1.1 Main results

In this section, we outline the structure of this paper and clearly identify its contributions. To compare computational performance between MC and the proposed method, we analyze each step of these frameworks. In the first step, a prescribed initial joint PDF supported over the space of initial conditions and parameters, is sampled. This is described in section 3. Our treatment here is mostly expository to demonstrate how ideas from different fields can be brought together in the context of this problem. Section 3.1 shows how the sampling strategy depends on the target density. For uniform sampling, accuracy of quasi-Monte Carlo samplers can be characterized in terms of number of samples and number of dimensions. For others, in high dimensions, Markov Chain Monte Carlo samplers guarantee asymptotic convergence. To characterize runtime complexity (section 3.2), an analysis of 'burn-in period' is required that till date can be done for few restrictive cases. For uniform samplers, however, we provide a runtime summary in Table 2. Section 4 concentrates on the second step, i.e. propagation of joint PDF. Section 4.1 compares accuracy and section 4.2 deals with runtime complexity. Finally, in section 5, we analyze the performance of computing marginal PDF of interest that can be seen as a post-processing. Accuracy and runtime are compared in section 5.1 and 5.2 respectively. It can be noted that unlike propagation of joint PDF, marginal PDF computation is an approximation algorithm for both MC and Perron-Frobenius operator formalism. As a result, part of section 4.1 and all of section 5.1, draw heavily on the theory of constructive approximation to guarantee necessary and sufficient conditions on accuracy. This pursuit connects the rate of convergence of the approximation algorithms with interpolation between function spaces and can be juxtaposed with more traditional analysis in nonparametric density estimation exercised by the statistics community. In the latter, only sufficiency is guaranteed under conservative assumptions on the density. We also improve the marginal approximation algorithm of [Halder and Bhattacharya, 2011] by deriving the optimal piecewise constant marginal approximation in the Perron-Frobenius operator framework. To help the readers, section 5.1.1 and the two appendices supplement the necessary background. The comparative performance guarantees established in this work are summarized in Table 1.

## 2 Notations

Most notations are standard. The set of whole numbers is denoted by $\mathbb{N}_0 = \mathbb{N} \bigcup \{0\}$, where $\mathbb{N}$ stands for the set of natural numbers. $\mathbb{R}^d$ is the real Euclidean space in $d$ dimensions. The symbol $|.|$ denotes magnitude that returns

Table 1: Summary of results

| | Accuracy[a] | | Runtime[b] | |
|---|---|---|---|---|
| | MC | PF | MC | PF |
| Sampling initial PDF | $\begin{cases}\text{For QMC samplers, Theorem 1}\\ \text{For MCMC samplers, asymptotic convergence}\end{cases}$ | | $\begin{cases}T_{\text{gen}} = O\left(\xi\left(\nu, d\right)\right), \text{ in general}\\ \text{For QMC samplers, Table 2}\end{cases}$ | |
| Uncertainty propagation | $O\left(N^{-\alpha/d}\right),\ 0 < \alpha < 1$ | $\begin{cases}\text{zero}\\ O\left((\Delta t)^q\right)\end{cases}$ | $\begin{cases}T_{\text{dyn}} + O\left(\nu d\right)\\ = O\left(\nu\,\zeta\left(d\right)\right) + O\left(\nu d\right)\end{cases}$ | $\begin{cases}T_{\text{dyn}} + O\left(\nu d\right)\\ = O\left(\nu\,\zeta\left(d\right)\right) + O\left(\nu d\right)\end{cases}$ |
| Marginal approximation | $O\left(N^{-\alpha/m}\right),\ 0 < \alpha < 1$ | $O\left(N^{-\alpha/m}\right),\ 0 < \alpha < 1$ | $O\left(\nu^{d-m+1}\right)$ | $O\left(\nu\right)$ |

[a] $N$ = number of tiling partitions of the domain, $d$ = dimension of the support of joint PDF, $m$ = dimension of the support of marginal PDF, $\Delta t \coloneqq \max_{k=1,\cdots,K}\{\Delta t_k\}$ = largest step-size for a variable step integrator, $q$ = order of the integrator

[b] $\nu$ = number of samples, $\zeta\left(d\right)$ = time complexity for propagating one sample in $d$ dimensions

absolute value for scalar argument and Euclidean length for vector argument. $\|\,.\,\|_{X(\Omega)}$ is the norm w.r.t. the Banach space $X$ over domain $\Omega \subset \mathbb{R}^d$ while $|\,.\,|_{Y(\Omega)}$ stands for the seminorm. By $A \hookrightarrow B$, we mean the function space $A$ is embedded in the function space $B$. The symbols $\lesssim$, $\gtrsim$ denote weak inequality up to a constant. $\lceil\alpha\rceil$ returns the smallest integer bigger or equal to $\alpha$; $\lfloor\alpha\rfloor$ returns the largest integer smaller or equal to $\alpha$. The big $O$ notation is often used to describe the order of numerical accuracy or that of computational complexity of an algorithm. By the notation $f\left(n\right) \in O\left(g\left(n\right)\right)$, we mean $\exists\ k > 0$ and $n_0$, such that $|f\left(n\right)| \leqslant kg\left(n\right),\ \forall\ n > n_0$. Asymptotically, $|f\left(n\right)| \leqslant kg\left(n\right)$ for some positive $k$. $\#\left(S\right)$ denotes cardinality of the set $S$. The divergence of a vector field is denoted by $\operatorname{div}\left(.\right)$. Also, the indicator function is denoted by $\chi$ and the symbol $\operatorname{Vol}\left(.\right)$ returns the volume or associated Lebesgue measure.

# 3 Sampling of initial distribution

To describe how uncertainties in initial conditions and parameters evolve in a dynamical system, we need to quantify the uncertainty in the combined space of initial conditions and parameters. In this paper, such a quantification is done in weak distributional sense. Assuming the underlying probability measure to be absolutely continuous, the initial distribution is specified as a PDF supported over the initial conditions and parameters. Then the question arises, how to generate a user-specified number of samples from such prescribed initial joint PDF?

An important observation here is that the users' freedom to specify any number of multidimensional samples precludes grid-based approach. Algorithmically, keeping number of samples independent of dimensions, saves the sampling step from 'curse of dimensionality' [Bellman, 1957]. However, to ensure the quality of samples (i.e. how accurately do the samples represent the original specified distribution), dimensions do come into play. Some explicit guarantees can be given depending on the algorithm and the underlying distribution to sample, as expatiated below.

## 3.1 Accuracy

If the initial density is jointly uniform, then multidimensional *Halton sequence* [Halton, 1960] provides a deterministic method to generate a prescribed number of samples. Since the construction is based on a deterministic algorithm (as opposed to randomized algorithm), it's often called a quasi Monte Carlo (QMC) method of sample generation. The Halton sequence is a multidimensional extension of *van der Corput sequence with respect to base* $b$, denoted by $\mathcal{V}_b$, whose $n^{\text{th}}$ term is given by

$$\mathcal{V}_b\left(n\right) \coloneqq \sum_{r=0}^{\infty} \frac{a_r\left(n\right)}{b^{r+1}}, \qquad n \in \mathbb{N}, \tag{1}$$

where $a_r(n)$ is the $r^{\text{th}}$ digit of the $b$-adic expansion of

$$n-1=\sum_{r=0}^{\infty} a_r(n)\, b^r. \tag{2}$$

The $n^{\text{th}}$ term of the Halton sequence in unit hypercube $[0,1]^d$ is defined as the $d$-tuple

$$\mathcal{H}_n := \left(\mathcal{V}_{b_1}(n), \cdots, \mathcal{V}_{b_d}(n)\right), \tag{3}$$

where the $b_j$s are pairwise coprime, for $j=1,\cdots,d$. Accuracy of the generated Halton points is quantified through the notion of *discrepancy*, denoted by $D_\nu(\mathcal{S}_d) \in (0,1)$, that intuitively measures the amount of irregularity in the generated $d$-dimensional sample set of cardinality $\nu$ i.e. $\mathcal{S}_d := \{\mathcal{H}_1, \cdots, \mathcal{H}_\nu\}$. Since the samples of uniform distribution are desired to be scattered evenly throughout the domain, lower discrepancy is better.

**Definition 1.** ***(Discrepancy and star discrepancy)*** *Let $\mathscr{F}$ denote the nonempty family of subsets of $[0,1]^d$. The discrepancy $D_\nu(\mathcal{S}_d)$ of a sample set $\mathcal{S}_d$ of cardinality $\nu$, is defined as*

$$D_\nu(\mathcal{S}_d) := \sup_{\mathcal{F}\in\mathscr{F}} \left| \frac{1}{\nu}\sum_{i=1}^{\nu} \chi(\mathcal{H}_i \in \mathcal{F}) - \text{Vol}(\mathcal{F}) \right|. \tag{4}$$

*If instead of $\mathscr{F}$, we use $\mathscr{G}$, the nonempty family of subsets of semi-open unit hypercube $[0,1)^d$, then the above defines star discrepancy $D^\star_\nu(\mathcal{S}_d)$ as*

$$D^\star_\nu(\mathcal{S}_d) := \sup_{\mathcal{G}\in\mathscr{G}} \left| \frac{1}{\nu}\sum_{i=1}^{\nu} \chi(\mathcal{H}_i \in \mathcal{G}) - \text{Vol}(\mathcal{G}) \right|. \tag{5}$$

The following theorem ensures that the multidimensional Halton sequence, as defined in (3), is of low discrepancy.

**Theorem 1** (Niederreiter, 1992)**.** $\forall\ \nu \geqslant 1, \quad D^\star_\nu(\mathcal{S}_d) < \dfrac{d}{\nu} + \dfrac{1}{\nu}\displaystyle\prod_{j=1}^{d}\left(\frac{b_j-1}{2\log b_j}\log\nu + \frac{b_j+1}{2}\right)$.

There are other low discrepancy sequences like Sobol, Faure and Niederreiter sequences [Niederreiter, 1992], which suitably permute Halton sequences to lower the discrepancy for large $d$. If the initial joint density is other than uniform, then the availability of a low discrepancy uniform random number generator still comes handy for inverse transform sampling or other specialized methods [Devroye, 1986].

Unfortunately, sampling non-uniform density through transform techniques suffers from computational inefficiency in high dimensions. Hence, to sample arbitrary initial PDF in high dimensions, one must resort to a Markov Chain Monte Carlo (MCMC) sampler. The core idea behind MCMC is to create a Markov chain whose stationary distribution is the one from where we want to sample. Two popular ways of doing this are *Metropolis-Hastings algorithm* [Chib and Greenberg, 1995; Diaconis and Saloff-Coste, 1998; Liu, 2001] and *Gibbs sampler* (also known as *Glauber dynamics* and *heat-bath algorithm*) [Glauber, 1963; Turčin, 1971; Diaconis *et. al.*, 2008]. Since MCMC guarantees that the constructed Markov chain *asymptotically* converges to the desired density, to adjudge the performance of MCMC, one can ask: how long must the chain be run to get sufficiently close to the target density? This rate of convergence, often called 'burn-in period' in MCMC literature, quantifies the performance of a sample generation algorithm. We describe it next.

## 3.2 Computational time

Although numerically implementing Metropolis-Hastings and Gibbs sampler are straightforward, deriving rigorous bounds for burn-in period is a daunting task. In particular, sharp computable bounds are not available in general setting. Hence two approaches of tackling this problem have evolved in the statistics community. One is the practitioners' approach, where instead of computing an upper bound for burn-in period, a *convergence diagnostics* is performed to test whether the realized samples "seem to be stable" after "large enough" number of iterations. Some widely used convergence diagnostics can be found in [Cowles and Carlin, 1996]. The advantage of these diagnostics is that they often work in practice for any general setting. On the flip side, lack of theoretical rigor may lead to a premature claim for convergence [Matthews, 1993] or introduce bias in the resulting estimates [Cowles *et. al.*, 1999].

The second approach strives to find a *rigorous upper bound for the burn-in period.* Here the distance between the $n^{\text{th}}$ iterate of the Markov chain and the stationary target density is estimated in total variation norm. The

idea is to upper bound this distance as a function of $n$, to answer how many steps are necessary to be $\epsilon$ close to the target density. This thread of research is an ongoing pursuit. In our context of continuous state space, most results [see section 3, Roberts and Rosenthal, 2004; Meyn and Tweedie, 2009] impose some restrictions on the Markov chain (e.g. uniform ergodicity, geometric ergodicity etc.). Others rely on symmetry [see section 2, Diaconis and Saloff-Coste, 1998] or geometric techniques [section 4, Diaconis and Saloff-Coste, 1998] for finite state space.

Due to lack of practical and computable MCMC bounds in continuous state space, we simply denote the time required to sample as $T_{\text{gen}} = O(\xi(\nu, d))$, a function of $\nu$ and $d$. To make the comparison meaningful, we propagate the uncertainty with same set of MCMC samples, for both MC and PF method. This precludes $T_{\text{gen}}$ from exponential dependence on $d$, i.e. grid-based curse of dimensionality. However, if the initial density is jointly uniform, then the QMC samplers of section 3.1 are in force and runtime complexity results are known for them (see Table 2).

Table 2: Computational speed of QMC samplers

| Runtime complexity $T_{\text{gen}} = O(\xi(\nu, d))$ for QMC sequences | | | |
|---|---|---|---|
| Halton sequence | $O(d \log \nu)$ | [Halton and Smith, 1964] | |
| Faure sequence | $O\left(d\left(\frac{\log \nu}{\log \eta_d}\right)^2\right)$ | [Fox, 1986] | $\eta_d$ is the first prime number $\geqslant d$ |
| Sobol sequence | $O\left(d(\log \nu)^2\right)$ | [Sobol, 1979] | Sobol's original implementation |
| Sobol sequence with XOR and assembly-like shift-and-count | $O\left(d\left\lceil \frac{\log \vartheta}{w} \right\rceil\right)$ | [Bratley and Fox, 1988] | $\vartheta > \nu$ is a user-specified upper bound, $w$ is the computer word-length |
| Sobol sequence with XOR but without shift-and-count | $O\left(d\left\lceil \frac{\log \vartheta}{w} \right\rceil + \log \nu\right)$ | [Fox, 1986] | |

# 4 Propagation of uncertainty

To determine the evolution of parametric and initial condition uncertainty, the joint density needs to be transported in the extended state space. Given the vector field in nonlinear state space form

$$\dot{z} = f(z, p), \qquad z \in \mathbb{R}^{n_s}, \quad p \in \mathbb{R}^{n_p}, \tag{6}$$

we define the augmented state space $x := \begin{bmatrix} z & p \end{bmatrix}^T$. Here $n_s$ and $n_p$ denote the number of states and number of parameters, respectively. The dimension of the extended state space is $n := n_s + n_p$. Let the uncertainty in the initial conditions ($z_0 := z(t = 0)$) and parameters ($p$), be quantified through the initial joint density $\rho_0(x) := \rho(x, t = 0)$. In dispersion analysis, the question of interest is to study the spatio-temporal propagation of uncertainties in $\Omega \times (0, \infty)$, $\Omega \in \mathbb{R}^n$, i.e. to compute $\rho(x, t) \geqslant 0$, where $\int_\Omega \rho(x, t)\, dx = 1$ for almost every $t \in (0, \infty)$. This evolution of probability density subject to the dynamics (6) is governed by the first order partial differential equation (PDE)

$$\frac{\partial}{\partial t}\rho(x, t) + \sum_{i=1}^{n_s} \frac{\partial}{\partial x_i}\left(\rho(x, t) F_i(x)\right) = 0, \tag{7}$$

known as the *Liouville equation*, where $F(x) : \mathbb{R}^n \to \mathbb{R}^n$, is the augmented dynamics $\begin{bmatrix} f_{n_s \times 1} & 0_{n_p \times 1} \end{bmatrix}^T$. Equation (7) can be seen as the transport equation associated with the PF operator [Lasota and Mackey, 1994], that advects probability mass in the extended state space while keeping the total mass conserved.

It can be shown that (7) can be solved using MOC. In this approach, the trajectories of the augmented dynamics $\dot{x} = F(x)$, become the characteristic curves. Along these characteristics, the joint PDF value can be

updated by solving the ODE initial value problem

$$\frac{d}{dt}\rho\left(x,t\right) = -\rho\left(x,t\right)\operatorname{div}\left(F\left(x\right)\right), \tag{8}$$

with $\rho_0\left(x\right) := \rho\left(x, t=0\right)$, given. For deriving (8) from (7), refer [section II, Halder and Bhattacharya, 2011]. Some illustrative examples can be found in the same reference showing (8) enables exact solution for $\rho\left(x,t\right)$ through direct integration.

This approach of exactly computing $\rho\left(x,t\right)$ is in sharp contrast with the traditional MC where only samples are propagated using the dynamics $\dot{x} = F\left(x\right)$ and then joint PDF values are approximately constructed through histogram. It's obvious that with the same number of samples, MOC based PF solution for joint PDF will reward better accuracy than MC. In section 4.1, we provide a quantitative estimate for the same. Section 4.2 assures that runtime complexity is not compromised in this pursuit of numerical accuracy.

## 4.1 Accuracy

The main idea here is that constructing joint PDF through MC histogram is an *approximate method* (piecewise constant approximation) while computing the same via PF operator based Liouville equation, is an *exact method*. Except the truncation error of the integrator, for any generic nonlinear dynamics, the joint PDF weights updated through MOC computation of the Liouville PDE, is exact at the sample sites. The finite sample computation of MOC does not incur any loss of generality. To elicit this, consider a case where the user queries the instantaneous joint PDF value at a location of the extended state space where none of the finite samples have landed. To determine this value exactly, by back integrating the dynamics till $t = 0$, we can determine the initial condition this sample would have come from. If the corresponding initial condition is found to lie outside the support of the prescribed initial joint density, the joint PDF value at the query site is zero. Otherwise, the instantaneous joint PDF value can be computed exactly by forward integrating the Liouville equation along the characteristic curve for that single initial condition.

### 4.1.1 Numerical accuracy for computing joint PDF using MC histogram

The joint PDF from MC simulation is constructed by making a multivariate histogram. If the number of samples is of the order of inverse of integration error for propagating each sample, then truncation error affects the accuracy of MC histogram construction. For a reasonably higher order integrator with small step-size, the major source of error is the *piecewise constant approximation* of the joint PDF via frequentist bin counting. This occurs as a post-processing step. To keep things simple, we only consider the case for *tiling histograms* i.e. the partitioning of the domain consists of disjoint union of hyper-rectangles. Even in this simple setting, the approximation may take place in a linear (uniform tiling) or nonlinear (nonuniform tiling, e.g. adaptive) fashion. It section 5.1.2, we will show that under mild assumptions, the true joint density lies in the smoothness space $B_1^\alpha\left(L_1\left(\Omega\right)\right) \hookrightarrow B_\infty^\alpha\left(L_1\left(\Omega\right)\right)$, $0 < \alpha < 1$. Here $\Omega \subset \mathbb{R}^n$, denotes the instantaneous domain of the MC samples, $n := n_s + n_p$ being the dimension of the extended state space. Such a knowledge about the underlying function that we are trying to approximate, enables us to guarantee an optimal rate of approximation $N^{-\alpha/n}$, for uniform tiling of $N$ buckets, using linear approximation theory to be detailed in the sequel (see section 5.1.3). Since $0 < \alpha < 1$ and $n \in \mathbb{N}$, for a fixed dimension $n$, no grid refinement can make the error decay faster than $O\left(N^{-1}\right)$. For a fixed smoothness $\alpha \in (0,1)$, such fractional decay of error aggravates further with the increase in dimension $n$. This result can not be improved using nonlinear approximation (Appendix B).

### 4.1.2 Numerical accuracy for computing joint PDF using Liouville equation

MOC implementation of Liouville equation is an *exact method* to update the joint PDF weights along each trajectory. Thus, for a given dynamics, if the ODE (8) can be solved analytically, then the total error to compute joint PDF is zero. However, for most nonlinear dynamics, (8) would require numerical integration and the joint PDF computation would be as accurate as the truncation error of the integrator (modulo round off error). A popular choice is the fourth order Runge-Kutta integrator with *fifth order truncation error*. For example, the `ode45` command [Shampine and Reichelt, 1997] in MATLAB® implements a Dormand-Prince (4,5) pair [Dormand and Prince, 1980] for variable step-size fourth order Runge-Kutta method. Then integrating (8) using `ode45` with $\Delta t = O\left(10^{-2}\right)$ would induce a (local) truncation error of $O\left(10^{-10}\right)$. One may resort to even higher order integrators [Hairer *et. al.*, 1993; Hairer and Wanner, 1996] if further accuracy is desired. Using fixed step-size fourth order Runge-Kutta discretization, the rate of convergence for $N$ steps, is $O\left(N^{-4}\right)$.

## 4.2 Computational time

Since both MC and PF operator based method propagate the samples through dynamics, the only difference in computational speed comes from the construction of the respective joint PDF vectors. In MC method, this is done by constructing a multidimensional histogram over the extended state space while in PF method, one additional equation is propagated for each sample, to update the joint PDF values at each time, at each sample site.

### 4.2.1 Time for computing joint PDF using MC histogram

Computational performance for univariate MC histogram construction is well understood and a considerable body of literature exists [see for example, Haas *et. al.*, 1995; Ioannidis, 2003] in the database and networking community. Unlike the one-dimensional case [Jagadish *et. al.*, 1998], constructing optimal multivariate histogram is NP-hard [Muthukrishnan, 1998]. Hence several approximate algorithms have been proposed for the multi-dimensional case based on greedy heuristics [Khanna et. al., 1997], wavelets [Vitter and Wang, 1999], kernel methods [Gunopulos *et. al.*, 2000] and discrete cosine transform [Greenwald and Khanna, 2001].

The optimality of a histogram can be cast as a primal-dual formulation [Jagadish *et. al.*, 1998]. Let the sample data-set be given by $X := \{X_1, X_2, \cdots, X_\nu\}$, where each sample point $X_i$ is a $d$-tuple of the form $(x_{i_1}, x_{i_2}, \cdots, x_{i_d})$, $i = 1, 2, \cdots, \nu$. Given $\#(X) =: \nu$ number of samples in $d$ dimensions, the primal problem strives to find the *minimum error histogram* $\underset{\mathcal{H}(X)\in\mathscr{H}(X)}{\operatorname{argmin}} \| \rho(X) - \mathcal{H}(X) \|_{L_p(\Omega)}$ with a fixed computational budget of $B$ buckets in $d$ dimensions. The minimization takes place over $\mathscr{H}(X)$, the set of all admissible histograms supported over the sample data-set $X$. The error norm $\| \cdot \|_{L_p(\Omega)}$ makes sense since $\rho(X)$ stays in $L_p$ at all times, whenever the initial density $\rho_0 \in L_p$, $1 \leqslant p \leqslant \infty$. This will be detailed in section 5.1.2. Using dynamic programming, it can be shown that the one dimensional primal problem is solvable in $O\left(B\nu^2\right)$ time for $p = 2$ and in $O\left(B\nu^3\right)$ time otherwise. On the other hand, the dual problem, with $\#(X) =: \nu$ number of samples, seeks a histogram such that the dual error has at most $\mathcal{E}$ deviation from the primal error $\mathcal{E}^\star_{\text{primal}} := \min_{\mathcal{H}(X)\in\mathscr{H}(X)} \| \rho(X) - \mathcal{H}(X) \|_{L_p(\Omega)}$. Notice that the primal problem has $B$ fixed but error is free. In the dual problem, error is specified as a constraint but $B$ is free. Hence the former is known as *space bounded histogram problem* while the latter is called *error bounded histogram problem*. The best known runtime for one dimensional dual problem is due to [Guha and Koudas, 2002] that provides an algorithm to ensure dual error $\leqslant (1+\mathcal{E})\,\mathcal{E}^\star_{\text{primal}}$ using $B^\star_{\text{dual}} = B$ buckets, in $O\left(\left(\frac{B}{\mathcal{E}}\log\nu\right)^3 + \nu\right)$ time. In one dimension, both primal and dual problems require $O(B\nu)$ storage.

In multiple dimensions, there are very few results [see for example, Muthukrishnan and Suel, 2005] that attempt the static primal-dual problem in *approximately* optimal sense (due to NP-hardness) to get a bound in terms of dimension $d$, sample size $\nu$ and approximation parameter. In [Muthukrishnan and Suel, 2005], the treatment was for a class of metrics for tiling histogram. In this paper, instead of analyzing the approximate optimization problem, we estimate the time complexity for constructing static histogram from $\nu$ MC samples in $d$ dimensions with a specified mesh. Without loss of generality, we assume the data to lie in the first quadrant and that the histogram grid is uniform. If the domain is bounded by a $d$-dimensional hypercube of length $L$ and along each dimension, the step-size of our uniform grid is $b$, then the objective is to estimate the time to construct a histogram supported over $B := \left(\frac{L}{b}\right)^d$ buckets. A simple way to do this is shown in Algorithm 1.

From the algorithm, we can see that the inner loop checks whether a sample point falls inside a $d$-dimensional bucket or not. Clearly, this has worst case complexity $O(d)$. Hence Algorithm 1 takes $O(\nu d)$ time and is linear in both sample size and dimensions. It can be noticed that the complexity result holds even when the specified grid is non-uniform, i.e. $B = \prod_{i=1}^{d} \frac{L_i}{b_i}$. If propagating nonlinear dynamics for each $d$-dimensional sample takes $O(\zeta(d))$ time, then propagation using dynamics (till the snapshot of interest) takes total $T_{\text{dyn}} = O(\nu\zeta(d))$ time. Thus overall runtime for MC joint PDF construction is $T_{\text{dyn}} + O(\nu d)$.

### 4.2.2 Time for computing joint PDF using Liouville equation

The Liouville equation is solved along each sample trajectory through MOC. The associated ODE has joint PDF $\rho$ as the dependent variable with time $t$ being the independent variable. The integrand contains the divergence of the $d$-dimensional vector field and hence can be computed in $O(d)$ time, provided the individual gradients have been computed offline using numerical interpolation when the dynamics is available as data instead of analytical expressions. This is the situation often encountered in practice and allows us to compute a bound that does not

**Algorithm 1** Calculate joint PDF using MC histogram

**Require:** $X$, $L$, $b$

$B \leftarrow \left(\frac{L}{b}\right)^d$

bucket = $\mathbf{0}_{B\times 1}$ ▷ Initialize the vector of buckets

**for** $i = 1$ to $\nu$ **do** ▷ The sample index

  bucketID = 1 ▷ Initialize bucket index

  **for** $j = 1$ to $d$ **do** ▷ The dimension index

    $k \leftarrow \left\lfloor \frac{X(i,j)}{b} \right\rfloor + 1$

    bucketID $\leftarrow$ bucketID $+ (k-1)(B)^{\frac{j-1}{d}}$

  **end for**

  bucket (bucketID) $\leftarrow$ bucket (bucketID) + 1 ▷ Count the number of samples falling inside each bucket

**end for**

depend on the exact functional form of the dynamics. Thus the cost for MOC solution of the Liouville equation with $\nu$ samples is $O(\nu d)$. The overall cost for computing joint PDF through Liouville equation, thus becomes $T_{\text{dyn}} + O(\nu d)$. When compared to MC construction, the complexity is identical. Notice however that in MC method, constructing joint PDF via histogram is a post-processing step. In PF operator based construction of joint PDF via Liouville equation, the joint PDF values are computed "on the fly" along each trajectory.

# 5 Computation of marginals

Before discussing the marginal approximation algorithm and its computational performance, in section 5.1, we first review some preliminaries for the same. The main purpose is to ensure notational consistency as the subsequent sections depend on the tools described here. Then section 5.2 proceeds for obtaining a rate-of-convergence of the approximation algorithm followed by assessing the computational speed in section 5.3.

## 5.1 Mathematical preliminaries

Let $\Omega \subset \mathbb{R}^d$ be an open, connected domain. Let's denote $\overline{\Omega}$ as its closure and $\partial\Omega := \overline{\Omega} \setminus \Omega$ as the boundary. Further, let $\Omega_e := \mathbb{R}^d \setminus \overline{\Omega}$ denote the associated exterior domain.

**Definition 2.** *(**Lipschitz domain**) A bounded domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$, is said to be a Lipschitz domain, if there exist constants $\lambda, \mu > 0$, and a finite number of local coordinate systems $\left(x_1^k, x_2^k, \cdots, x_d^k\right)$, $1 \leqslant k \leqslant K \in \mathbb{N}$, and associated local Lipschitz continuous mappings*

$$\phi_k : \{\widehat{x}^k = \left(x_2^k, \cdots, x_d^k\right) \in \mathbb{R}^{d-1} |\ \ |x_i^k| \leqslant \lambda,\ 2 \leqslant i \leqslant d\} \longmapsto \mathbb{R}$$

*such that*

$$\partial\Omega = \bigcup_{k=1}^{K} \{\left(x_1^k, \widehat{x}^k\right) |\ \ x_1^k = \phi_k\left(\widehat{x}^k\right),\ \ |\widehat{x}^k| < \lambda\},$$

$$\{\left(x_1^k, \widehat{x}^k\right) |\ \phi_k\left(\widehat{x}^k\right) < x_1^k < \phi_k\left(\widehat{x}^k\right) + \mu,\ \ |\widehat{x}^k| < \lambda\} \subset \Omega,\ \ 1 \leqslant k \leqslant K,$$

$$\{\left(x_1^k, \widehat{x}^k\right) |\ \phi_k\left(\widehat{x}^k\right) - \mu < x_1^k < \phi_k\left(\widehat{x}^k\right),\ \ |\widehat{x}^k| < \lambda\} \subset \Omega_e,\ \ 1 \leqslant k \leqslant K.$$

Intuitively, such a domain simply guarantees the existence of finite number of balls to cover the boundary $\partial\Omega$ and that the intersection of $\partial\Omega$ with any such ball results a Lipschitz 1 (see definition 8) graph. Geometrically, it means that one can fit a cone with finite angle whose vertex is on $\partial\Omega$, such that the cone remains in $\Omega$. In other words, both $\Omega$ and $\Omega_e$ are *locally* situated on exactly one side of the boundary. In our context of PDF evolution, we will assume that the support of the joint PDF remains a Lipschitz domain at all times. All subsequent results for the quality of marginal approximation algorithm, will invoke this assumption.

We recall the function space $L_p(\Omega)$ as the normed linear space of $p^{\text{th}}$ order integrable functions on $\Omega$ for $0 < p < \infty$ and as the normed linear space of essentially bounded functions for $p = \infty$.

**Definition 3.** *($L_p$ **space**) A function $f(.)$ is said to be in $L_p(\Omega)$ if the corresponding norm is finite.*

$$\| f \|_{L_p(\Omega)} := \begin{cases} \left( \int_\Omega |f(x)|^p \, dx \right)^{1/p} & \text{for } 0 < p < \infty, \\ \operatorname*{ess\,sup}_{x \in \Omega} |f(x)| & \text{for } p = \infty. \end{cases}$$

For $1 \leqslant p \leqslant \infty$, they define Banach spaces except for $p = 2$, which becomes a Hilbert space equipped with the inner product $\langle f, g \rangle_{L_2(\Omega)} := \int_\Omega f(x) \overline{g(x)} \, dx$. For $0 < p < 1$, they are only quasi-Banach spaces. For $1 \leqslant p < q \leqslant \infty$, we have the embedding: $L_q(\Omega) \hookrightarrow L_p(\Omega)$.

To work with the functional smoothness spaces for multivariate case, the concept of (mixed) weak derivatives is useful since it extends the idea of classical derivatives to functions which are integrable, but may not be differentiable.

**Definition 4.** *(**Weak derivatives**) Let $\nu = (\nu_1, \nu_2, \cdots, \nu_d)$ be a multi-index of magnitude $|\nu| := \sum_{i=1}^{d} \nu_i$, where $\nu_i \in \mathbb{N}_0$. The function $g \in L_1(\Omega)$ is said to be the $\nu^{th}$ weak derivative of a function $f \in L_1(\Omega)$, denoted by $D^\nu f$, if*

$$\int_\Omega f D^\nu \varphi \, dx = (-1)^{|\nu|} \int_\Omega \varphi g \, dx, \qquad \forall \ \varphi \in C_0^\infty(\Omega).$$

Of particular interest is when the weak derivative of a function lands up in $L_p$. We say, $D^\nu f \in L_p(\Omega)$ if $\exists \, g \in L_p(\Omega)$ such that $\int_\Omega f D^\nu \varphi \, dx = (-1)^{|\nu|} \int_\Omega \varphi g \, dx, \quad \forall \ \varphi \in C_0^\infty(\Omega)$. With this notion of weak derivatives, we can introduce Sobolev spaces $W^r(L_p(\Omega))$ as more general smoothness spaces of functions than $L_p(\Omega)$.

**Definition 5.** *(**Sobolev space**) The Sobolev space $W^r(L_p(\Omega)) := \{ f \in L_p(\Omega) \mid D^\nu f \in L_p(\Omega), |\nu| \leqslant r \}$, consists of all functions in $L_p(\Omega)$, whose weak derivatives up to $r^{th}$ order, are also in $L_p(\Omega)$.*

The above defines Banach spaces with norm $\| f \|_{W^r_{L_p(\Omega)}} := \| f \|_{L_p(\Omega)} + |f|_{W^r(L_p(\Omega))}$, where the seminorm $|f|_{W^r(L_p(\Omega))} := \sum_{|\nu|=1}^{r} \| D^\nu f \|_{L_p(\Omega)}$. Clearly, the space $W^0(L_p(\Omega))$ is identical to $L_p(\Omega)$. Further, $W^r(L_p(\Omega)) \hookrightarrow L_p(\Omega)$ and $W^{r+1}(L_p(\Omega)) \hookrightarrow W^r(L_p(\Omega))$ (reduction theorem). A natural question is: when does $W^r(L_p(\Omega))$ embed in $L_q(\Omega)$? *Sobolev embedding theorem* answers this question.

**Theorem 2.** *(**Sobolev Embedding Theorem**) Consider a Lipschitz domain $\Omega \subset \mathbb{R}^d$, $r \in \mathbb{N}_0$ and $1 \leqslant p \leqslant \infty$. Let the critical Sobolev exponent $\tau$ be given by $\frac{1}{\tau} = \frac{1}{p} - \frac{r}{d}$. Then, we have the following compact embeddings for Sobolev spaces $W^r(L_p(\Omega))$.*

$$\begin{aligned} &\text{(a)} \quad W^r(L_p(\Omega)) \hookrightarrow L_q(\Omega), \quad 1 \leqslant q \leqslant \tau, \quad \text{if } r - \frac{d}{p} < 0, \\ &\text{(b)} \quad W^r(L_p(\Omega)) \hookrightarrow L_q(\Omega), \ 1 \leqslant q < \infty, \quad \text{if } r - \frac{d}{p} = 0, \\ &\text{(c)} \quad W^r(L_p(\Omega)) \hookrightarrow C(\overline{\Omega}), \qquad\qquad\quad \text{if } r - \frac{d}{p} > 0. \end{aligned}$$

The original proof was due to [Sobolev, 1938] and the modified versions can be found in [Gilbarg and Trudinger, 1977].

Measuring the smoothness of a function using (weak) derivatives is too crude for many practical purposes. For any function $f(.)$, *modulus of smoothness* provides a generic characterization.

**Definition 6.** *(**Modulus of smoothness**) The $r^{th}$ modulus of smoothness of $f \in L_p(\Omega)$, $0 < p \leqslant \infty$, is defined as*

$$\omega_r(f,t)_{L_p(\Omega)} := \sup_{0 \leqslant |h| \leqslant t} \| \Delta_h^r(f, x) \|_{L_p(\Omega_{r,h})}$$

*where the $r^{th}$ order difference operator $\Delta_h^r(f, x) := \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} f(x + kh)$, and the vectors $x := (x_1, \cdots, x_d)$, $h := (h_1, \cdots, h_d) \in \mathbb{R}^d$. Further, $\Omega_{r,h} := \{ x \mid x + kh \in \Omega, \ k = 0, 1, \cdots, r \} \subset \Omega$.*

It can be seen that $\omega_r(f,t)_{L_p(\Omega)}$ is finite for any $t$. Moreover, it is a continuous, monotone increasing function of $t$ that vanishes at $t=0$. Intuitively, how fast $\omega_r(f,t)_{L_p(\Omega)}$ decays as $t\to 0$, is a measure of how smooth the function $f$ is. Besov spaces capture this generic smoothness characterization of a function through $\omega_r(f,t)_{L_p(\Omega)}$.

**Definition 7.** *(**Besov space**) The Besov space $B_q^\alpha(L_p(\Omega))$ consists of functions $f\in L_p(\Omega)$ which have smoothness $\alpha$.*

We say $f\in B_q^\alpha(L_p(\Omega))$, $0<\alpha,p,q\leqslant\infty$, if the associated norm $\|f\|_{B_q^\alpha(L_p(\Omega))}:=\|f\|_{L_p(\Omega)}+|f|_{B_q^\alpha(L_p(\Omega))}$, is finite; where the Besov seminorm is defined as

$$|f|_{B_q^\alpha(L_p(\Omega))}:=\begin{cases}\left(\int_0^\infty\left(t^{-\alpha}\,\omega_r(f,t)_{L_p(\Omega)}\right)^q\frac{dt}{t}\right)^{1/q} & \text{for } 0<q<\infty,\\ \sup\limits_{t>0}\left|t^{-\alpha}\,\omega_r(f,t)_{L_p(\Omega)}\right| & \text{for } q=\infty,\end{cases}$$

where $r:=\lfloor\alpha\rfloor+1$. The parameter $q$ provides a finer gradation of functional smoothness. We have $B_q^\alpha(L_p(\Omega))\hookrightarrow L_p(\Omega)$ and $B_{q_1}^\alpha(L_p(\Omega))\hookrightarrow B_{q_2}^\beta(L_p(\Omega))$ for $0<\beta<\alpha$, $0<q_1,q_2\leqslant\infty$, $1\leqslant p\leqslant\infty$ (reduction theorem). Also, we have the following embeddings: $B_{q_1}^\alpha(L_p(\Omega))\hookrightarrow B_{q_2}^\alpha(L_p(\Omega))$ for $0<q_1<q_2\leqslant\infty$, $1\leqslant p\leqslant\infty$; $B_q^\alpha(L_{p_1}(\Omega))\hookrightarrow B_q^\alpha(L_{p_2}(\Omega))$ for $1<p_2\leqslant p_1$. Since $\alpha$ need not be a natural number, Besov spaces are more general smoothness spaces than Sobolev spaces. However, the spaces $W^r(L_p(\Omega))$ and $B_q^r(L_p(\Omega))$, $r\in\mathbb{N}$, are not the same, unless $p=2$, $q=2$, when they enjoy a Hilbert space structure.[1] Various equivalent norms and interpolation properties of Besov spaces can be found in [DeVore and Popov, 1988].

**Definition 8.** *(**Lipschitz and generalized Lipschitz space**) The Lipschitz space $\mathrm{Lip}(\alpha,r,L_p)$ is defined as*

$$\mathrm{Lip}(\alpha,r,L_p):=\{f:\omega_r(f,t)_{L_p}\leqslant Mt^\alpha\ \ \forall\, t>0\}. \tag{9}$$

*For $r=1$, $\mathrm{Lip}(\alpha,1,L_p)$ is often simply written as $\mathrm{Lip}(\alpha,L_p):=\{f:\|f(x+h)-f(x)\|_{L_p(\Omega_{1,h})}\leqslant Mh^\alpha\ \forall\, h>0\}$. In general, it's cumbersome to track the definition for different values of $r$. To get around this, generalized Lipschitz spaces $\mathrm{Lip}^\star(\alpha,L_p)$ are introduced as*

$$\mathrm{Lip}^\star(\alpha,L_p):=\mathrm{Lip}(\alpha,r,L_p),\qquad \alpha>0,\quad r:=\lfloor\alpha\rfloor+1. \tag{10}$$

*Notice that for $0<\alpha<1$, $\mathrm{Lip}^\star(\alpha,L_p)=\mathrm{Lip}(\alpha,L_p)$. But in general, $\mathrm{Lip}^\star(\alpha,L_p)$ is weaker than $\mathrm{Lip}(\alpha,L_p)$.*

**Definition 9.** *(**Bounded variation space**) The space of $p^{th}$ bounded variation, $BV_p$ is given by*

$$BV_p:=\left\{f:|f|_{BV_p}:=\sup_{0\leqslant x_0\leqslant\cdots\leqslant x_N\leqslant 1}\left(\sum_{i=1}^N|f(x_i)-f(x_{i-1})|^p\right)^{1/p}<\infty\right\},\qquad 1\leqslant p<\infty, \tag{11}$$

*where $N$ is arbitrary. For $p=1$, often it is simply denoted as $BV$ space.*

**Definition 10.** *(**DeVore diagram**) DeVore diagram is a tool to visualize functional smoothness spaces and various embeddings among them through a plot of $\frac{1}{p}$ (as the abscissa) versus smoothness $\alpha$ (as the ordinate), $1\leqslant p\leqslant\infty$, $0\leqslant\alpha$.*

To illustrate its utility, consider the DeVore diagram in Fig.1(a). On the $x$-axis, we have $L_p$ spaces, $1\leqslant p\leqslant\infty$. The embedding $L_q(\Omega)\hookrightarrow L_p(\Omega)$ for $1\leqslant p<q\leqslant\infty$, mentioned earlier, implies that any space on the $x$-axis of DeVore diagram, is embedded to another on its right. Similarly, if we consider the discrete natural number levels $r\in\mathbb{N}$ on the $y$-axis, then $C^\infty(\Omega)\hookrightarrow\cdots\hookrightarrow C^r(\Omega)\hookrightarrow C^{r-1}(\Omega)\hookrightarrow\cdots\hookrightarrow C(\Omega)\hookrightarrow L_\infty(\Omega)$, as shown. In general, a point $\left(\frac{1}{q},r\right)$ in Fig.1(a) represents the Sobolev space $W^r(L_q(\Omega))$ and as we climb vertically up from $L_q(\Omega)$, the spaces get smaller, smoother and nested, i.e., $W^r(L_q(\Omega))\hookrightarrow W^{r-1}(L_q(\Omega))\hookrightarrow\cdots\hookrightarrow L_q(\Omega)$. In particular, we can now provide an illustration of Sobolev embedding theorem. To answer whether $W^r(L_p(\Omega))\hookrightarrow L_q(\Omega)$, we simply draw a line (solid, red) from $L_q(\Omega)$ with slope $d$, called *Sobolev embedding line*. The theorem says that any point

[1] Sobolev spaces have been generalized to have fractional order smoothness (see [Stein, 1970]; [Adams, 1975], for example). However, it can be shown that (see Theorem 6.7 in [DeVore and Sharpley, 1993]) the fractional order Sobolev space $W^r(L_p(\Omega))$, where $r>0$ is a fraction, is *equivalent* to the special family of Besov space $B_p^r(L_p(\Omega))$, $1\leqslant p<\infty$. Hence, in this paper, without loss of generality, the Sobolev spaces are dealt with integer smoothness only.

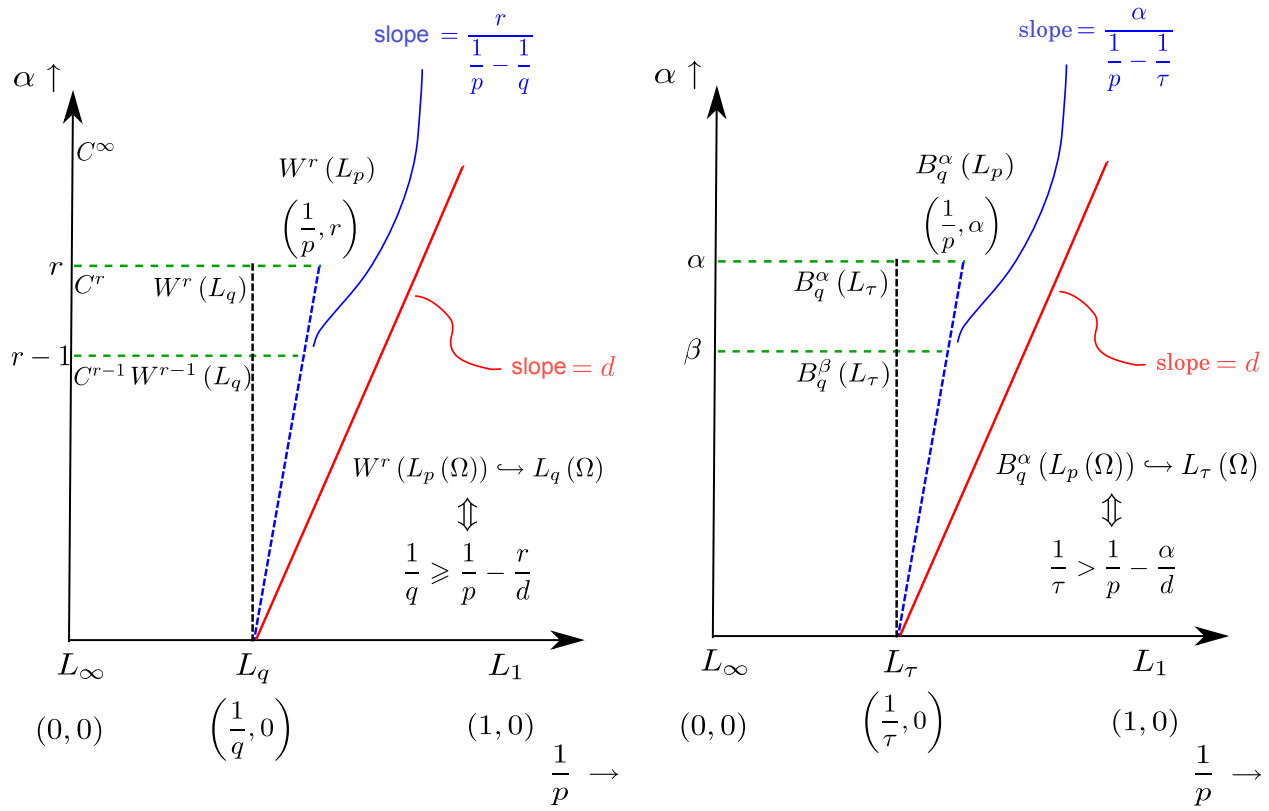


Figure 1: DeVore diagrams illustrating (a) (left) Sobolev spaces, (b) (right) Besov spaces and their embeddings.

$\left(\frac{1}{p}, r\right)$ lying to the left or on the Sobolev embedding line, is compactly embedded in $L_q(\Omega)$. Sobolev himself showed [Sobolev, 1938] such embeddings are strict in the sense that points to the right of the Sobolev embedding line, can not be embedded in $L_q(\Omega)$. We can intuitively see from the diagram that as we go to higher dimensions ($d$ large), it becomes increasingly difficult to have Sobolev embeddings. Notice that although Theorem 2 holds for any $1 \leqslant q \leqslant \infty$, there is a minor technical nuisance when $\frac{1}{q}$ is pushed left to the origin ($L_\infty$ space). In such a situation, if $\left(\frac{1}{p}, r\right)$ is on the Sobolev embedding line, then Sobolev embedding fails (part (b) of Theorem 1). This is sometimes referred as *endpoint Sobolev anomaly*.

Fig.1(a) shows that only the quantized integer levels of smoothness are populated by the Sobolev spaces. In Fig.1(b), we illustrate that Besov spaces 'fill in the gaps' by allowing any smoothness $\alpha \geqslant 0$. Fig.1(b) allows any point $\left(\frac{1}{p}, \alpha\right)$ in the DeVore diagram to be defined as the Besov space $B^\alpha_q(L_p(\Omega))$. As before, we have the vertical embeddings $B^\alpha_{q_1}(L_\tau(\Omega)) \hookrightarrow B^\beta_{q_2}(L_\tau(\Omega)) \hookrightarrow \cdots \hookrightarrow L_\tau(\Omega)$ for $0 < \beta < \alpha$, $0 < q_1, q_2 \leqslant \infty$, $1 \leqslant \tau \leqslant \infty$. Furthermore, we have Sobolev-type embedding theorem for Besov spaces that says any point $\left(\frac{1}{p}, \alpha\right)$ lying strictly left to the Sobolev embedding line (solid, red) through $L_\tau(\Omega)$, is compactly embedded in $L_\tau(\Omega)$. As before, such an embedding is strict in the sense that points to the right of the Sobolev embedding line, can not be embedded in $L_\tau(\Omega)$. However, if the point $\left(\frac{1}{p}, \alpha\right)$ is on the Sobolev embedding line, then whether it can be embedded in $L_\tau(\Omega)$ depends on $q$. In such a situation, $B^\alpha_q(L_p(\Omega)) \hookrightarrow L_\tau(\Omega)$ iff $q \leqslant \max(p, 1)$. Further structural properties of Besov spaces can be found in [Triebel, 1992; Peetre, 1976; Besov and Kalyabin, 2003; DeVore and Yu, 1991].

## 5.2 Accuracy

To analyze the accuracy of the approximation algorithms for computing marginals, we will proceed in two steps. First, we will discuss the regularity of the true marginal densities to be approximated (section 5.1.1). Second,

the approximation set up and rate of approximation will be described (section 5.1.2).

### 5.2.1 Regularity of the marginals

To decide the regularity of the marginal PDFs, one must begin with the regularity of the joint PDFs computed through Liouville equation. Since we are interested to approximate the marginal snapshots, we will restrict ourselves to spatial regularity only. Then, the question of interest is: at fixed $t = t^\star$, what kind of functional smoothness class does the joint density $\rho(x, t^\star)$ belong to?

It's obvious that being a density function, at least $\rho(x, t^\star) \in L_1(\mathbb{R}^n)$, where $n := n_s + n_p$ is the dimension of the extended state space (i.e. support of the joint density). To characterize further, one must know the regularity of the nonlinear dynamics. To see this, one can expand the gradient term in Liouville equation using the product rule of differentiation and then notice that, even though the initial density $\rho_0(x) \in L_1(\mathbb{R}^n)$, depending on the regularity of $f(x)$, the terms $\dfrac{\partial \rho}{\partial x_i}$ may or may not lie in $L_1(\mathbb{R}^n)$. Hence, we can not always guarantee that $\rho(x, t^\star) \in W^1(L_1(\mathbb{R}^n))$.

In the literature, some attempts [DiPerna and Lions, 1989; Emamirad and Protopopescu, 1996; Le Bris and Lions, 2004; Le Bris and Lions, 2008] have been made to characterize the smoothness of the evolving joint PDF, subject to a dynamics whose smoothness is specified. Since such PDF evolution is governed by a transport PDE (Liouville equation for ODE dynamics, Fokker-Planck equation for SDE dynamics), characterizing smoothness of the joint PDF calls for a regularity analysis of the associated PDE. The analysis proceeds in two steps. First, an *a priori* estimate for the solution of the PDE is obtained, followed by a regularization procedure. For this two steps to be possible, one needs to impose some restrictions on the vector field. An important result in this direction is due to [DiPerna and Lions, 1989], stated below.

**Theorem 3.** ***(Regularity of the joint PDF under Liouville evolution with locally Sobolev dynamics)*** *[DiPerna and Lions, 1989] When the dynamics $f := f_1 + f_2$ is such that $f \in W^1_{loc}(L_1(\mathbb{R}^n))$, $\nabla.f \in L_\infty(\mathbb{R}^n)$, $\dfrac{|f_1|}{1+|x|} \in L_1(\mathbb{R}^n)$ and $\dfrac{|f_2|}{1+|x|} \in L_\infty(\mathbb{R}^n)$, then the solution of eqn. (7) at time $t = t^\star$ satisfies $\rho(x, t^\star) \in L_p(\mathbb{R}^n)$, provided the initial density $\rho_0(x) \in L_p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$.*

**Remark 1.** *In [Emamirad and Protopopescu, 1996], it was demonstrated that if we remove the asymptotic growth condition $\dfrac{|f|}{1+|x|} \in (L_1 + L_\infty)(\mathbb{R}^n)$, then the above statement holds for $p = 1$.*

**Remark 2.** *Further, it has been shown [Ambrosio, 2004] that the requirement $f \in W^1_{loc}(L_1(\mathbb{R}^n))$ can be relaxed by $f \in BV_{loc}(\mathbb{R}^n)$. Since $W^1(L_1(\mathbb{R}^n)) \hookrightarrow BV(\mathbb{R}^n) \hookrightarrow B^1_\infty(L_1(\mathbb{R}^n))$, it's natural to ask whether one can provide PDF regularity with Besov smoothness $\left(i.e.\ f \in B^1_\infty(L_1(\mathbb{R}^n))\right)$ for the dynamics. To the best of our knowledge, this remains yet to be addressed.*

Now we are going to claim that under mild assumption, $\rho(x, t^\star) \in B^\alpha_1(L_p(\mathbb{R}^n))$, $\quad 0 < \alpha < 1$, $p = 1$. To prove such a claim, it would have been great if we could show $\rho(x, t^\star) \in W^1(L_p)$. Then we could use the embedding $W^1(L_p) \hookrightarrow B^1_1(L_p) \hookrightarrow B^\alpha_1(L_p)$ and conclude the proof. However, as argued in the second paragraph of this section, we can not, in general, guarantee $\rho(x, t^\star) \in W^1(L_1)$ for a joint PDF under Liouville evolution. Theorem 3 guarantees the evolved PDF snapshot to be in $L_p(\mathbb{R}^n)$ whenever the initial density is in $L_p(\mathbb{R}^n)$. In the most general case, $p = 1$ and we seek a subspace of $L_1(\mathbb{R}^n)$ where our PDF may evolve to.

**Assumption 1.** $\rho(x, t^\star) \in BV(\mathbb{R}^n)$.

The assumption assures that for $p = 1$, we can slightly relax our demand for Sobolev regularity and enables us to show $\rho(x, t^\star) \in B^\alpha_1(L_1(\mathbb{R}^n))$, $\quad 0 < \alpha < 1$. This assumption can be justified by thinking the ODE dynamics as a weak noise limit of SDE dynamics. In the latter case, the PDF evolution occurs with bounded Wasserstein distance [Jordan *et. al*, 1998] and hence with bounded total variation [Gibbs and Su, 2002]. In practice, such an assumption is commonplace in nonparametric density estimation [Obereder *et. al.*, 2007]. Now we relax the space where the evolved joint PDF may lie, by interpolating between $L_1(\mathbb{R}^n)$ and $\mathrm{BV}(\mathbb{R}^n)$.

**Proposition 1.** *Assumption 1 implies $\rho(x, t^\star) \in B^\alpha_1(L_1(\mathbb{R}^n))$, $0 < \alpha < 1$.*

*Proof.* Since $\mathrm{BV}(\mathbb{R}^n) \hookrightarrow B^1_1(L_1(\mathbb{R}^n)) \hookrightarrow B^\alpha_1(L_1(\mathbb{R}^n))$, $0 < \alpha < 1$, the result follows from Assumption 1. $\square$

**Remark 3.** *It is not clear to us if the assumption 1 can be lifted while establishing Besov regularity of the joint PDF. One way to do this could be arguing that under some regularity of the nonlinear dynamics, Liouville operator is pseudo-differential [Taylor, 1981] and then proving that the operator preserves Besov regularity, similar to [Nguyen, 2010]. Pursuing this direction is beyond the scope of this paper.*

So far we have argued the Besov regularity of the $n$-variate joint PDF $\rho_{x_{n\times 1}}(x_1, x_2, \cdots, x_n)$. To arrive at the Besov regularity of the $m$-variate marginal density $\rho_{x_{m\times 1}}(x_1, x_2, \cdots, x_m)$, $1 \leqslant m < n$, $m \in \mathbb{N}$, we utilize the Fubini property which, roughly speaking, says that a function is in a Besov space on $X_1 \times X_2$ with certain smoothness iff it lies separately in each Besov space on $X_1$ and on $X_2$ respectively, with the same smoothness.

**Theorem 4.** ***(Fubini property for Besov spaces)*** *For any smoothness $s > d\left(\frac{1}{\min(p,1)} - 1\right)$, $B_p^s\left(L_p\left(\mathbb{R}^d\right)\right)$ has the Fubini property. More generally, if $X_1$ and $X_2$ are compact manifolds, then*

$$B_p^s\left(L_p\left(X_1 \times X_2\right)\right) = L_p\left(X_1, B_p^s\left(L_p\left(X_2\right)\right)\right) \bigcap L_p\left(X_2, B_p^s\left(L_p\left(X_1\right)\right)\right).$$

*Proof.* See [Triebel, 1983, Theorem 2.5.13]. □

**Corollary 5.**

$$\rho_{x_{n\times 1}}(x_1, x_2, \cdots, x_n) \in B_1^\alpha\left(L_1\left(\mathbb{R}^n\right)\right) \Rightarrow \rho_{x_{m\times 1}}(x_1, x_2, \cdots, x_m) \in B_1^\alpha\left(L_1\left(\mathbb{R}^m\right)\right), \quad 0 < \alpha < 1, \quad 1 \leqslant m < n, \quad m \in \mathbb{N}.$$

*Proof.* Since $p = 1$ and the smoothness $\alpha > n(1-1) = 0$, the result follows from the above theorem. □

Corollary 5 shows that marginalization preserves Besov smoothness. This smoothness characterization of marginal density will be instrumental for deriving rate of convergence for approximating the same (see Theorem 6).

### 5.2.2 Rate of approximation

Having obtained an estimate for the true marginal PDF, we now concentrate on the performance of the piecewise constant approximation for the same. First, let's outline the approach of our analysis. Consider the generic approximation set up as described in Fig. 2(a), where one approximates a function $f \in \mathcal{X}$ using simpler functions drawn from $\Sigma_N^r$. Here $N$ denotes the number of partitions of the support and $r$ quantifies smoothness. For example, $\Sigma_N^r$ could be piecewise algebraic polynomials of order $r$. Our aim is twofold. First, given a fixed computational budget $N$, if someone picks up a function from a subspace of $\mathcal{X}$, we want to know the best rate of approximation for $f$, achievable in this set up, as a function of $N$. Second, we want to answer the inverse problem. Namely, if the user specifies a desired rate of approximation, can we characterize the subspace of $\mathcal{X}$, to which the function $f$ must belong to? These two answers, together establish the one-to-one correspondence between a function space and a rate of approximation.

To keep things simple, we will restrict ourselves to the framework of linear approximation, where $\Sigma_M^r + \Sigma_N^r = \Sigma_{M+N}^r$. Using the theory of $\mathcal{K}$ functionals and modulus of smoothness (definition 6), one can derive [DeVore and Lorentz, 1993] a 'black box theorem' for Lipschitz domain $\Omega \subset \mathbb{R}^m$ (definition 2) that says, whenever one can prove a *Jackson-type inequality* for $f \in \mathcal{Y}_r$, where $\mathcal{Y}_r$ is a subspace of (hence more regular than) $\mathcal{X}$ (see Fig. 2(a))

$$E_N(f)_{\mathcal{X}} \coloneqq \inf_{S \in \Sigma_N^r} \|f - S\|_{\mathcal{X}(\Omega)} \lesssim N^{-r} |f|_{\mathcal{Y}_r}, \tag{12}$$

and a companion *Bernstein-type inequality* for $g \in \Sigma_N^r$

$$|g|_{\mathcal{Y}_r} \lesssim N^r \parallel g \parallel_{\mathcal{X}}, \tag{13}$$

then

$$E_N(f)_{\mathcal{X}} \lesssim N^{-\alpha/m} \iff f \in \mathcal{A}_q^\alpha\left(\mathcal{X}, \mathcal{Y}_r\right)_{\theta, q=\infty} \tag{14}$$

where $\theta \coloneqq \dfrac{\alpha}{r} \in (0,1)$ measures how different is the characterizing space of functions $\mathcal{A}_q^\alpha$ (obtained as an interpolation between $\mathcal{X}$ and $\mathcal{Y}_r$ via $\theta$) compared to $\mathcal{Y}_r$. A small value of $\theta$ means $A_q^\alpha$ is slightly different than $\mathcal{Y}_r$, while a large $\theta$ implies $A_q^\alpha$ is slightly different than $\mathcal{X}$, i.e. a large subspace of $\mathcal{X}$. Clearly, as $\alpha$ increases, the quality of approximation improves.

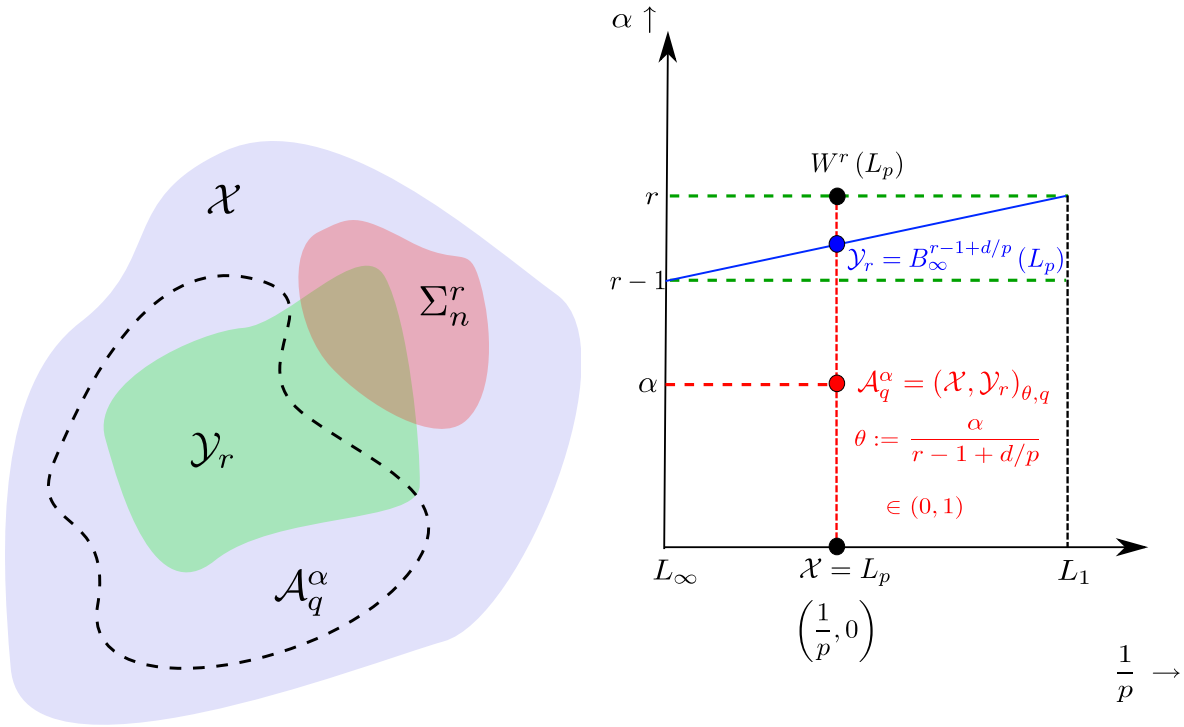


Figure 2: (a) The approximation set up: functions from $\mathcal{X}$ (*blue*) are to be approximated by those from $\Sigma_n^r$ (*red*). By choosing a space $\mathcal{Y}_r$ (*green*) to satisfy (12) and (13), we can characterize the approximation space $\mathcal{A}_q^\alpha$ (*with dotted boundary*) from 'black box theorem'; (b) DeVore diagram showing the characterization of the approximation space $\mathcal{A}_q^\alpha$ via interpolation between $\mathcal{X}$ and $\mathcal{Y}_r$.

To illustrate how the 'black box theorem' characterizes the approximation space via interpolation, consider approximating univariate functions in $\mathcal{X} = L_p$ using functions from $\Sigma_N^r = \Sigma_N =$ the space of trigonometric polynomials of degree $N$ over $[-\pi, \pi]$. In this situation, one can prove [DeVore and Lorentz, 1993] a Jackson-type and a Bernstein-type inequality with $\mathcal{Y}_r = W^r(L_p)$ and characterize $\mathcal{A}_q^\alpha = (L_p, W^r(L_p))_{\frac{\alpha}{r},q} = B_q^\alpha(L_p)$. In particular, $E_N(f)_{L_p[-\pi,\pi]} \lesssim N^{-\alpha} \iff f \in B_\infty^\alpha(L_p)$. However, if the approximating space $\Sigma_N^r$ consists of piecewise algebraic polynomials of order $r$, then $\mathcal{Y}_r = W^r(L_p(\Omega))$ does not work since the LHS of the corresponding Bernstein-type inequality (eqn. (13)) becomes undefined. It can be shown [DeVore and Lorentz, 1993] that $\mathcal{Y}_r = B_\infty^{r-1+d/p}(L_p(\Omega))$ works instead. This is pictorially illustrated in Fig. 2(b) where the approximation space is characterized by interpolating between $B_\infty^{r-1+d/p}(L_p(\Omega))$ and $L_p(\Omega)$, in lieu of interpolating between $W^r(L_p(\Omega))$ and $L_p(\Omega)$. The reason why a piecewise polynomial of order $r$ has smoothness $B_\infty^{r-1+d/p}(L_p(\Omega))$ is explained in Appendix A.

The marginal computation algorithm through level sets (or slices) is shown as a schematic in Fig. 3(b). Let the $j^{\text{th}}$ level set of the joint density be denoted as $\Lambda_j,\ j = 1, 2, \cdots, (N+1)$. Further, let's assume that the level $\Lambda_j$ contains $n_j$ points and the total number of samples representing the joint PDF be $\nu := \sum_{j=1}^{N+1} n_j$. Consider the following piecewise constant approximations of the marginals from MC and PF-based joint PDFs

$$S_{\text{MC}} = \sum_{j=1}^{N} \alpha_j \chi_{I_j}, \qquad \alpha_j := \frac{n_j}{\nu} \tag{15}$$

$$S_{\text{PF}} = \sum_{j=1}^{N} \beta_j \chi_{I_j}, \qquad \beta_j := \frac{1}{n_j} \sum_{i=1}^{n_j} \varphi_i, \tag{16}$$

where the value of the marginal PDF at level $\Lambda_j$ is held fixed through the interval $I_j := [\Lambda_j, \Lambda_{j+1})$ and $\varphi_i$ denotes the joint PDF value at a sample point $i$. Notice that in eqn.(15), one has to compute $n_j = \sum_{i=1}^{\nu} \chi_{ij}$ where $\chi_{ij} := \chi_{z_i \in \Lambda_j}$, $z_i$ being the $i^{\text{th}}$ multi-sample, $i = 1, \cdots, \nu$.

The true marginal PDF $\rho_m := \rho_{x_{m\times 1}}(x_1, x_2, \cdots, x_m)$, being a measurable function, has at least $L_1(\Omega)$, $\Omega \in \mathbb{R}^m$, regularity. In what follows, we prove that approximating $\rho_m$ by piecewise constant function $S := \sum_{j=1}^{N} c_j \chi_{I_j}$ results $O\left(N^{-\alpha/m}\right)$ approximation error with $\alpha \in (0, 1)$.

**Theorem 6.** *Let $\rho_m \in L_1(\Omega)$ and $\Sigma_N^1$ be the class of all piecewise constant functions supported over the disjoint union of $N$ equipartitions of the Lipschitz domain $\Omega \in \mathbb{R}^m$. Then the approximation error*

$$E_N(\rho_m)_{L_1(\Omega)} := \inf_{S \in \Sigma_N^1} \|\rho_m - S\|_{L_1(\Omega)} \lesssim N^{-\alpha/m}, \qquad 0 < \alpha < 1.$$

*Proof.* Here $\mathcal{X} = L_1(\Omega)$. From proposition 3 in Appendix A, we can take $\mathcal{Y}_r = B_\infty^1(L_1(\Omega))$ since

$$S \in B_\infty^n(L_1(\Omega)) \hookrightarrow B_\infty^1(L_1(\Omega)), \qquad 1 < n := n_s + n_p. \tag{17}$$

So we are in the setting depicted in Fig.2 with $p = 1$. The Jackson-type and Bernstein-type inequalities for this situation were proved in [DeVore and Lorentz, 1993] for $1 \leqslant p < \infty$. In particular, for $p = 1$, using 'black box theorem' we get

$$E_N(\rho_m)_{L_1(\Omega)} \lesssim N^{-\alpha/m}, \quad \iff \quad \rho_m \in \left(L_1(\Omega), B_\infty^1(L_1(\Omega))\right)_{\theta,q=\infty} = B_\infty^\alpha(L_1(\Omega)), \;\; 0 < \alpha < 1. \tag{18}$$

From Corollary 5, we know $\rho_m \in B_1^\alpha(L_1(\Omega)) \hookrightarrow B_\infty^\alpha(L_1(\Omega))$, $0 < \alpha < 1$, where the last relation follows from the embedding of Besov spaces discussed in Section 5.1.1. Hence the $\Leftarrow$ direction of (18) guarantees the best rate of approximation as $O\left(N^{-\alpha/m}\right)$. This concludes the proof. Additionally, the $\Rightarrow$ direction of (18) guarantees the sufficiency, i.e. with fixed computational budget $N$, if the desired rate of approximation is $O\left(N^{-\alpha/m}\right)$, then the function must belong to the specified interpolation space. □

**Remark 4.** *Since $0 < \alpha < 1$, quality of approximation improves as $\alpha$ gets closer to unity and vice versa. In particular, the approximation error curve always lies above $O\left(N^{-1}\right)$ curve. For a fixed $\alpha$, the quality of approximation exacerbates with increasing dimension $(m)$. The trend is illustrated in Fig.3 (a)–(c). As $m$ increases, one must increase $\alpha$ to guarantee a fixed accuracy, meaning the performance guarantee in high dimensions can not be met unless inordinate smoothness is ensured, a fact referred as the 'curse of dimensionality' [Bellman, 1957].*

**Remark 5.** *Notice that the approximation space $\Sigma_N^1$ is linear since $S_1 \in \Sigma_M^1$, $S_2 \in \Sigma_N^1 \;\Rightarrow\; S_1 + S_2 \in \Sigma_{M+N}^1$. Using the Sobolev-type embedding theorem for Besov spaces described in section 5.1.1, a corresponding 'black box theorem' can be found [DeVore, 2009] for nonlinear approximation, provided the approximating spaces are nested $(\Sigma_N \subset \Sigma_{N+1})$ and satisfy certain growth condition $\left(\Sigma_M + \Sigma_N \subset \Sigma_{a(M+N)},\, a \geqslant 1\right)$. In Appendix B, we summarize different 'black box type theorems' for nonlinear approximation and argue that in our context, nonlinear approximation methods do not improve the rate of convergence of linear approximation obtained in this section.*

Our next agenda is to prove that neither $S_{MC}$ nor $S_{PF}$ is the optimal approximation. We show that the optimal piecewise constant approximation is attained by holding the median of the joint PDF values corresponding to the samples lying at level $\Lambda_j$, over the interval $I_j$.

**Proposition 2.** $E_N(\rho_m)_{L_1(\Omega)} := \underset{S \in \Sigma_N^1}{\operatorname{argmin}} \|\rho_m - S\|_{L_1(\Omega)} = S_{\mathrm{PF}}^\star := \sum_{j=1}^N \gamma_j \chi_{I_j}$, *where* $\gamma_j := \operatorname{median}\left(\{\varphi_i\}_{i=1}^{n_j}\right)$.

*Proof.* The proof is by contradiction. Let's assume, if possible, for $\epsilon_1 > 0$, $\exists\, \delta_1 > 0$ such that $E_N(\rho_m)_{L_1(\Omega)} = \|\rho_m - S_1\|_{L_1(\Omega)} = \|\rho_m - S_{PF}^\star\|_{L_1(\Omega)} - \delta_1$, where $S_1 := \sum_{j=1}^N (\gamma_j \pm \epsilon_1)\chi_{I_j}$. In words, an $\epsilon_1$ perturbation from the median value, lowers the approximation error by an amount $\delta_1$. Now consider $0 < \epsilon_2 < \epsilon_1$ and let $S_2 := \sum_{j=1}^N (\gamma_j \pm \epsilon_2)\chi_{I_j}$ such that $\|\rho_m - S_2\|_{L_1(\Omega)} = \|\rho_m - S_{PF}^\star\|_{L_1(\Omega)} - \delta_2$. Since $\Omega = \bigcup_{j=1}^N I_j$ with $I_j \bigcap I_k = \varnothing$, $\forall\, j \neq k$, for any generic $S = \sum_{j=1}^N c_j \chi_{I_j}$, we can write $\int_\Omega |\rho_m - S|\, dx = \sum_{j=1}^N \int_{I_j} |\rho_m - c_j|\, dx$. Now we can look at the inner integral for each interval and see that if $c_j$ is taken as $\gamma_j := \operatorname{median}(\varphi_i)_{i=1}^{n_j}$ then by definition of median, there are as many joint PDF samples less than it than greater than it. More formally, if we respectively substitute $(\gamma_j \pm \epsilon_1)$ and $(\gamma_j \pm \epsilon_2)$, then $\delta_2 > \delta_1$, i.e. as we move closer toward the median, the approximation error decreases. This contradicts our assumption. Hence the statement. □

**Remark 6.** *The above is similar to a result in order statistics [David and Nagaraja, 2003] that guarantees median as the minimizer of absolute deviation for random samples in $L_1$ norm. The connection can be seen, for example, by replacing the inner integral in the proof above by its quasi-Monte-Carlo (QMC) approximation. Interestingly, the same theory, also guarantees that in the $L_2$ norm, sample mean is the minimizer. This is true in our case too resulting $S_{PF}$ in eqn.(16) as the optimal approximation provided $\rho_m \in L_2(\Omega)$. In general, the MC and PF marginal approximation errors $\|\rho_m - S_{MC}\|_{L_1}$ and $\|\rho_m - S_{PF}\|_{L_1}$ are suboptimal.*

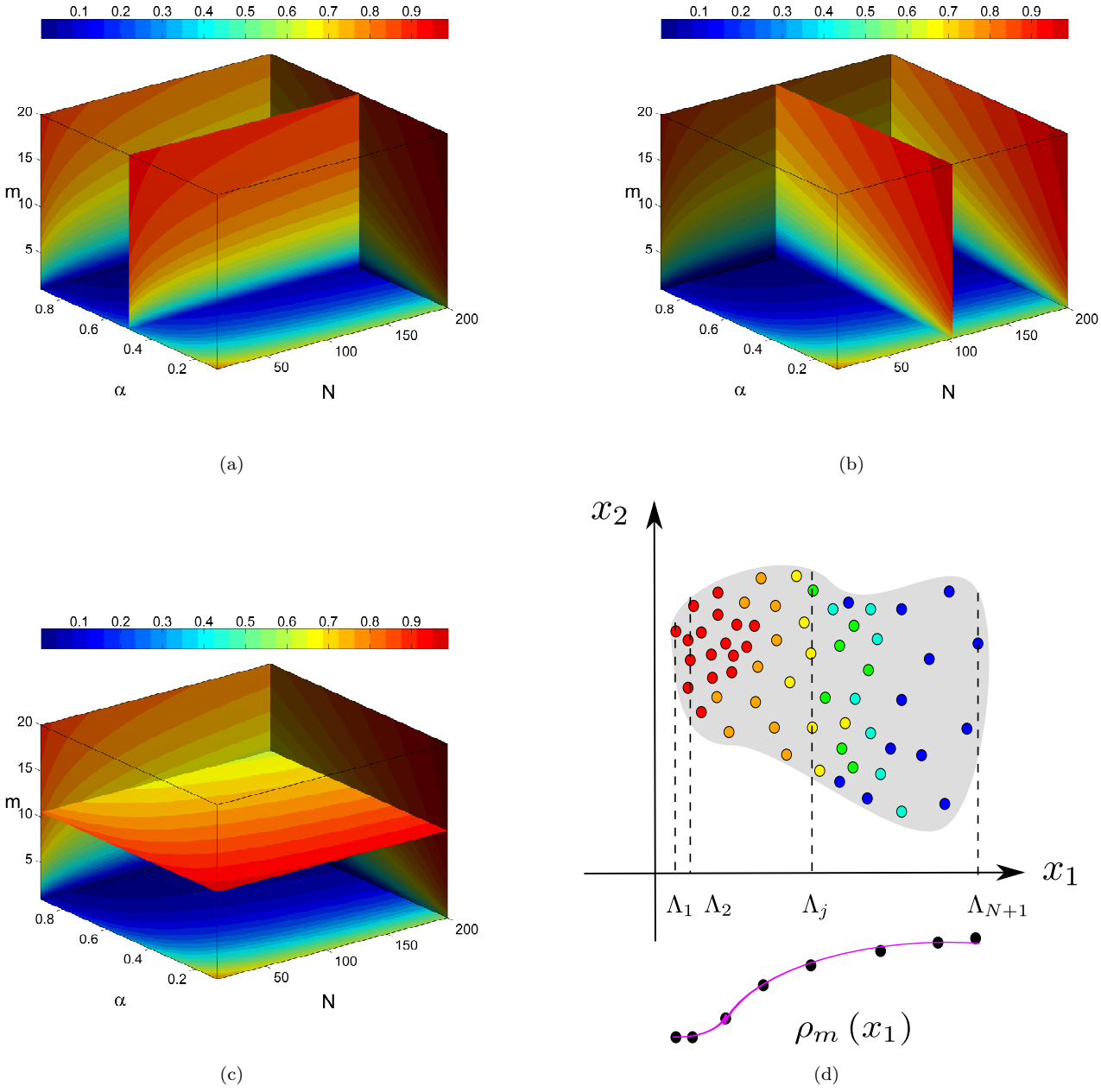


Figure 3: (a)–(c) The color value indicates the $m$ dimensional marginal approximation error $E_N(\rho_m)$ as a function of smoothness $\alpha$, support equipartition of cardinality $N$, and dimension $m$. The accuracy improves as $\alpha$ decreases, $N$ increases and $m$ decreases and vice versa. (d) The schematic shows the marginal computation algorithm. Here computing univariate marginal $\rho_m(x_1)$ from the bivariate joint PDF is shown. The gray region denotes the instantaneous support of the joint PDF. The sample colors indicate joint PDF values at that point (red = high, blue = low etc.).

**Remark 7.** *In Fig. 4, the approximated univariate marginal snapshots are plotted for the three-state hypersonic entry dynamics described in [Halder and Bhattacharya, 2011], at non-dimensional time* $\tilde{t} = 0.3$. *The states are altitude* $(h)$ *in Km, total velocity* $(V)$ *in Km/sec and flight path angle (FPA) in degrees, respectively. The initial joint PDF was assumed to be* $\mathcal{N}(\mu, \Sigma)$, *with* $\mu = [80\,\text{Km}, 3.5\,\text{Km/sec}, -2^\circ]^T$ *and* $\Sigma$ *is constructed by assuming 10% variance along each dimension about the respective means.*

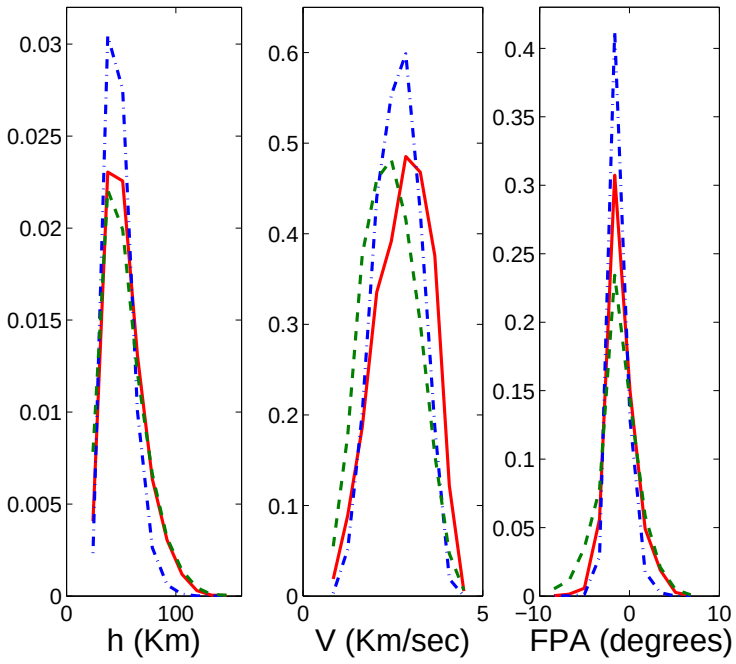


Figure 4: Univariate marginal snapshots at non-dimensional time $\tilde{t} = 0.3$ for the three-state hypersonic entry dynamics described in [Halder and Bhattacharya, 2011]. In each of the above subplots, $S_{\text{MC}}$ (*red, solid*) is the MC approximation (eqn. (15)), $S_{\text{PF}}$ (*blue, dash-dot*) is the suboptimal PF approximation (eqn. (16)), and $S^\star_{\text{PF}}$ (*green, dashed*) is the optimal PF approximation (Proposition 2).

## 5.3 Computational time

We recall that total number of samples $\nu := \sum_{j=1}^{N+1} n_j$. Both deterministic [Blum *et. al.*, 1973; Dor and Zwick, 1999] and randomized algorithms [Floyd and Rivest, 1975; Cunto and Munro, 1989] are available to compute the median $(\gamma_j)$ in linear $(O(n_j))$ time. Thus constructing $S^\star_{PF}$ requires $\sum_{j=1}^{N+1} O(n_j) = O(\nu)$ time. On the other hand, constructing $\beta_j$ requires exactly $(n_j - 1) + 1 = n_j$ operations. Hence, computing $S_{PF}$ requires exact $\nu$ operations. However, for computing $S_{MC}$, one needs to evaluate $n_j = \sum_{i=1}^{\nu} \chi_{ij}$ where $\chi_{ij} := \chi_{z_i \in \Lambda_j}$, $z_i$ being the $i^{\text{th}}$ multi-sample, $i = 1, \cdots, \nu$. The key observation here is that for PF operator based approximation algorithms, the availability of joint PDF values allow the approximation to be constructed by looking at the joint PDF vector on each level set. In MC method, however, we have to count samples in multiple dimensions of the level sets. Hence the dimensionality of the level set $(n - m)$ affects the time complexity of approximate MC marginals. For example, computing univariate marginals $(m = 1)$ from a tri-variate $(n = 3)$ joint PDF requires counting samples over two-dimensional level sets [see Fig. 6, Halder and Bhattacharya, 2011].

At a given $\ell$-dimensional level set $\Lambda_j$, MC implementation executes $\chi_{ij} := \chi_{z_i \in \Lambda_j}$, $z_i$, $i = 1, \cdots, \nu$, total $\nu^\ell$ times and is the rate determining step since constructing $\alpha_j$ requires additional $O(\nu) + 1$ operations. Hence constructing

MC marginal needs total $\sum_{j=1}^{N+1} O\left(\nu^{\ell}\right)$ number of operations. In the worst case, $N+1=\nu$; thus worst case complexity is $O\left(\nu^{\ell+1}\right)$. For us, $\ell=n-m$ and hence computing $S_{\mathrm{MC}}$ needs $O\left(\nu^{n-m+1}\right)$ time. For a fixed computational budget of $B$ buckets, the so called $B$-histogram [Thaper *et. al.*, 2002] marginal requires $O\left(B\nu^{n-m}\right)$ time. Notice that none of these are optimal.

# 6 Conclusion

Rigorous performance bounds for dispersion analysis were obtained with Monte Carlo (MC) simulation and a Perron-Frobenius (PF) operator based formulation, proposed recently in [Halder and Bhattacharya, 2011]. Analysis presented in this paper demonstrate that with same number of samples, the propagation of samples and *approximate* PDF construction via histogram have same runtime complexity with that of 'on the fly' *exact* joint PDF propagation via Liouville equation. However, with a fixed computational budget, the accuracy of the MC constructed PDF is always worse than $O\left(N^{-1}\right)$, $N$ being the number of tiling histogram partitions. However, the MOC implementation of the Perron-Frobenius operator formalism has zero total error, in case the ODE initial value problem (IVP) corresponding to the MOC formulation of Liouville PDE is analytically integrable. In case of numerical integration, the error can be no more than the small integration error, depending on the order of the ODE IVP solver. From the statistical analysis perspective, often the low dimensional marginal density is the object of interest. We analyzed the performance bounds for marginal approximations in MC and PF framework. We derived the optimal marginal approximation in PF framework and showed that in regard with marginal approximation, PF operator formulation improves both accuracy and runtime complexity, compared to MC simulation.

# Acknowledgement

This research was supported by National Science Foundation award # 1016299, with D. Helen Gill as the program manager. The first author was benefitted from communications with Ronald DeVore and Piotr Indyk. Their help is gratefully acknowledged.

# References

[1] Halder, A., and Bhattacharya, R., Dispersion Analysis in Hypersonic Flight During Planetary Entry Using Stochastic Liouville Equation. *Journal of Guidance, Control and Dynamics*, Vol. 34, No. 2, pp. 459–474, 2011.

[2] Dutta, P., and Bhattacharya, R., Hypersonic State Estimation using the Frobenius-Perron Operator. *Journal of Guidance, Control and Dynamics*, Vol. 34, No. 2, pp. 325–344, 2011.

[3] Terejanu, G. A., Singh, T., and Scott, P. D., Unscented Kalman Filter/Smoother for a CBRN Puff-Based Dispersion Model. *10th International Conference on Information Fusion*, Quebec City, Canada, 2007.

[4] Poroseva, S. V., Letschert, J., and Hussaini, M. Y., Application of Evidence Theory to Quantify Uncertainty in Forecast of Hurricane Path. *18th Conference on Probability and Statistics*, 86th Annual Meeting of the American Meteorologial Society, Atlanta, GA, 2006.

[5] Knocke, P. C., Wawrzyniak, G. G., Kennedy, B. M., Desai, P. N., Parker, T. J., Golombek, M. P., Duxbury, T. C., and Kass, D. M., Mars Exploration Rovers Landing Dispersion Analysis. *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, AIAA Paper 2004-5093, 2004.

[6] Galbally, D., Fidkowski, K., Willcox, K. and Ghattas, O., Nonlinear Model Reduction for Uncertainty Quantification in Large-Scale Inverse Problems. *International Journal for Numerical Methods in Engineering*, Vol. 81, No. 12, pp. 1581–1608, 2010.

[7] Pettersson, P., Iaccarino, G., and Nordström, J., Numerical Analysis of the Burgers' Equation in the Presence of Uncertainty. *Journal of Computational Physics*, Vol. 228, pp. 8394–8412, 2009.

[8] Ghanem, R. G., and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, Revised Ed., Dover Publications, 2003.

[9] Steven Greene, M., Liu, Y., Chen, W., and Liu, W. K., Computational Uncertainty Analysis in Multiresolution Materials via Stochastic Constitutive Theory. *Computer Methods in Applied Mechanics and Engineering*, Vol. 200, No. 1, pp. 309–325, 2011.

[10] Bellman, R. A., *Dynamic Programming*, Princeton University Press, 1957.

[11] Soize, C., and Ghanem, R., Physical Systems with Random Uncertainties: Chaos Representations with Arbitrary Probability Measure. *SIAM Journal on Scientific Computing*, Vol. 26, No. 2, pp. 395–410, 2004.

[12] Spanos, P., and Ghanem, R., Stochastic Finite Element Expansion for Random Media. *Journal of Engineering Mechanics*, Vol. 115, No. 5, pp. 1035–1053, 1989.

[13] Xiu, D., and Hesthaven, J., High-Order Collocation Methods for Differential Equations with Random Inputs. *SIAM Journal on Scientific Computing*, Vol. 27, No. 3, pp. 1118–1139, 2005.

[14] Pettit, C., Uncertainty Quantification in Aeroelasticity: Recent Results and Research Challenges. *Journal of Aircraft*, Vol. 41, No. 5, pp. 1217–1229, 2004.

[15] Mathelin, L., Hussaini, M. Y., and Zang, T. A., Stochastic Approaches to Uncertainty Quantification in CFD Simulations. *Numerical Algorithms*, Vol. 38, No. 1, pp. 209–236, 2005.

[16] Prabhakar, A., Fisher, J., and Bhattacharya, R., Polynomial Chaos Based Analysis of Probabilistic Uncertainty in Hypersonic Flight Dynamics. *Journal of Guidance, Control and Dynamics*, Vol. 33, No. 1, pp. 222–234, 2010.

[17] Dutta, P., and Bhattacharya, R., Nonlinear Estimation of Hypersonic State Trajectories in Bayesian Framework with Polynomial Chaos. *Journal of Guidance, Control and Dynamics*, Vol. 33, No. 6, pp. 1765–1778, 2010.

[18] Debusschere, B. J., Najm, H. N., Pébay, P. P., Knio, O. M., Ghanem, R., and Le Maître, O. P., Numerical Challenges in the Use of Polynomial Chaos Representations for Stochastic Processes. *SIAM Journal on Scientific Computing*, Vol. 26, No. 2, pp. 698–719, 2004.

[19] Lucas, L. J., Owhadi, H., and Oritz, M., Rigorous Verification, Validation, Uncertainty Quantification and Certification through Concentration-of-Measure Inequalitites. *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, pp. 4591–4609, 2008.

[20] Roberts, J. B., and Spanos, P. D., *Random Vibration and Statistical Linearization*, Dover Publications, 2003.

[21] Bergman, L. A., Wojtkiewicz, S. F., Johnson, E. A., and Spencer Jr., B. F., Some Reflections On the Efficacy of Moment Closure Methods. *Second International Conference on Computational Stochastic Mechanics*, Athens, Greece, June 13–15, 1994.

[22] Powell, R. W., Striepe, S. A., Desai, P. N., Queen, E. M., Tartabini, P. V., Brauer, G. L., Cornick, D. E., Olson, D. W., Petersen, F. M., Stevenson, R., Engel, M. C., and Marsh, S. M., Program to Optimize Simulated Trajectories (POST II), Vol. II Utilization Manual. *ver. 1.1.1.G*, NASA Langley Research Center, Hampton, VA, May 2000.

[23] Balaram, J., Austin, R., Banerjee, P., Bentley, T., Henriquez, D., Martin, B., McMahon, E., and Sohl, G., DESENDS – A High Fidelity Dynamics and Spacecraft Simulator for Entry, Descent and Surface Landing. *2002 IEEE Aerospace Conference*, Piscataway, NJ, 2002.

[24] Lasota, A., and Mackey, M., *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics*, Applied Mathematical Sciences, Vol. 97, Springer-Verlag, New York, 1994.

[25] Balaram, J., Bhattacharya, R., and Halder, A., Beyond Monte Carlo: Statistical Verification and Validation of Space Systems. *NASA Jet Propulsion Laboratory DRDF Final Report 1328*, 2010.

[26] Halton, J. H., On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-dimensional Integrals. *Numerische Mathematik*, Vol. 2, pp. 84–90, 1960.

[27] Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, 1st Ed., SIAM, Philadelphia, 1992.

[28] Devroye, L., *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, 1986.

[29] Chib, S., and Greenberg, E., Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, Vol. 49, No. 4, pp. 327–335, 1995.

[30] Diaconis, P., and Saloff-Coste, L., What Do We Know About the Metropolis Algorithm? *Journal of Computer and System Sciences*, Vol. 57, No. 1, pp. 20–36, 1998.

[31] Liu, J. S., *Monte Carlo Strategies for Scientific Computing*, Springer Series in Statistics, Springer-Verlag, New York, 2001.

[32] Glauber, R., Time Dependent Statistics of the Ising Model. *Journal of Mathematical Physics*, Vol. 4, pp. 294–307, 1963.

[33] Turčin, V., On the Computation of Multidimensional Integrals by the Monte Carlo Method. *Theory of Probability and Its Applications*, Vol. 16, No. 4, pp. 720–724, 1971.

[34] Diaconis, P., Khare, K., and Saloff-Coste, L., Gibbs Sampling, Exponential Families and Orthogonal Polynomials. *Statistical Science*, Vol. 23, pp. 151–178, 2008.

[35] Cowles, M. K., and Carlin, B. P., Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, Vol. 91, No. 434, pp. 883–904, 1996.

[36] Matthews, P., A Slowly Mixing Markov Chain with Implications for Gibbs Sampling. *Statistics & Probability Letters*, Vol. 17, No. 3, pp. 231–236, 1993.

[37] Cowles, M. K., Roberts, G. O., and Rosenthal, J. S., Possible Biases Induced by MCMC Convergence Diagnostics. *Journal of Statistical Computation and Simulation*, Vol. 64, pp. 87–104, 1999.

[38] Roberts, G. O., and Rosenthal, J. S., General State Space Markov Chains and MCMC Algorithms. *Probability Surveys*, Vol. 1, pp. 20–71, 2004.

[39] Meyn, S., and Tweedie, R. S., *Markov Chains and Stochastic Stability*, Second Ed., Cambridge University Press, 2009.

[40] Halton, J. H., and Smith, G. B., Algorithm 247: Radical-Inverse Quasirandom Point Sequence. *Communications of the ACM*, Vol. 7, pp. 701–702, 1964.

[41] Fox, B. L., Algorithm 647: Implementation and Relative Efficiency of Quasirandom Sequence Generators. *ACM Transactions on Mathematical Software*, Vol. 12, No. 4, pp. 362–376, 1986.

[42] Sobol, I. M., On the Systematic Search in a Hypercube. *SIAM Journal of Numerical Analysis*, Vol. 16, pp. 790–793, 1979.

[43] Bratley, P., and Fox, B. L., Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator. *ACM Transactions on Mathematical Software*, Vol. 14, No. 1, pp. 88–100, 1988.

[44] Shampine, L. F., and Reichelt, M. W., The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, Vol. 18, No. 1, pp. 1–22, 1997.

[45] Dormand, J. R. and Prince, P. J., A Family of Embedded Runge-Kutta Formulae. *Journal of Computational and Applied Mathematics*, Vol. 6, No. 1, pp. 19–26, 1980.

[46] Hairer, E., Nrsett, S. P., and Wanner, G., *Solving Ordinary Differential Equations I: Nonstiff Problems*, 2nd Ed., Springer-Verlag, 1993.

[47] Hairer, E., and Wanner, G., *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, 2nd Ed., Springer-Verlag, 1996.

[48] Haas, P., Naughton, J., Seshadri, S., and Stokes, L., Sampling Based Estimation of the Number of Distinct Values of an Attribute. *Proceedings of the International Conference on Very Large Databases*, pp. 311–322, 1995.

[49] Ioannidis, Y., The History of Histograms (abridged). *Proceedings of the International Conference on Very Large Databases*, Vol. 29, pp. 19–30, 2003.

[50] Jagadish, H. V., Koudas, N., Muthukrishnan, S., Poosala, V., Sevcik, K. C., and Suel, T., Optimal Histograms with Quality Guarantees. *$24^{th}$ International Conference on Very Large Databases*, San Francisco, CA, 1998.

[51] Muthukrishnan, S., Poosala, V., and Suel, T., Partitioning Two-dimensional Arrays: Algorithms, Complexity and Applications. *International Conference on Database Theory*, Jerusalem, Israel, 1999.

[52] Khanna, S., Muthukrishnan, S., and Skiena, S., Efficient Array Partitioning. *Proccedings of the $24^{th}$ International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, Springer, Issue: 1256, pp. 616–626, 1997.

[53] Vitter, J., and Wang, M., Approximate Computation of Multidimensional Aggregates on Sparse Data Using Wavelets. *Proccedings of the ACM SIGMOD International Conference on Management of Data*, pp. 193–204, 1999.

[54] Gunopulos, D., Kollios, G., Tsptras, V., and Domeniconi, C., Approximating Multi-dimensional Aggregate Range Queries Over Real Attributes. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 463–474, 2000.

[55] Greenwald, M., and Khanna, S., Space Efficient Online Computation of Quantile Summaries. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 58–66, 2001.

[56] Guha, S., and Koudas, N., Approximating A Data Stream for Querying and Estimation: Algorithms and Performance Evaluation. *Proceedings of the $18^{th}$ International Conference Data Engineering*, pp. 567–576, 2002.

[57] Muthukrishnan, S., and Suel, T., Approximation Algorithms for Array Partitioning Problems. *Journal of Algorithms*, Vol. 54, No. 1, pp. 85–104, 2005.

[58] Sobolev, S. L., On a Theorem of Functional Analysis. *Mat. Sb.*, Vol. 46, No. 4, pp. 471–497, 1938. (in Russian) English translation in *Transactions of the American Mathematical Society*, Vol. 34, No. 2, pp. 39–68, 1963.

[59] Gilbarg, D., and Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order.* Springer Verlag, First ed., 1977.

[60] Stein, E. M., *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, First ed., 1970.

[61] Adams, R. A., *Sobolev Spaces.* Academic Press, First ed., 1975.

[62] DeVore, R. A., and Sharpley, R. C., Besov Spaces on Domains in $\mathbb{R}^d$. *Transactions of the American Mathematical Society*, Vol. 335, No. 2, pp. 843–864, 1993.

[63] DeVore, R. A., and Popov, V. A., Interpolation of Besov Spaces. *Transactions of the American Mathematical Society*, Vol. 305, No. 1, pp. 397–414, 1988.

[64] Triebel, H., *Theory of Function Spaces.* Vol. II, Birkhäuser Verlag, Basel, 1992.

[65] Peetre, J., *New Thoughts on Besov Spaces.* Duke University Mathematics Series, Durham, 1976.

[66] Besov, O., and Kalyabin, G., Spaces of Differentiable Functions. *Function Spaces, Differential Operators, and Nonlinear Analysis*, The Hans Triebel Anniversary Volume, D. Haroske, T. Runst, H.-J. Schmeisser (eds.), pp. 3–21, Birkhäuser Verlag, Basel, 2003.

[67] DeVore, R. A., and Yu, X. M., $\mathcal{K}$ functionals for Besov Spaces. *Journal of Approximation Theory*, Vol. 67, pp. 38–50, 1991.

[68] DiPerna, R. J., and Lions, P. L., Ordinary Differential Equations, Transport Theory and Sobolev Spaces. *Inventiones Mathematicae*, Vol. 98, pp. 511–547, 1989.

[69] Emamirad, H., and Protopoescu, V., The Liouville Equation in $L^1$ Spaces. *Applied Mathematics Letters*, Vol. 9, No. 1, pp. 49–53, 1996.

[70] Le Bris, C., and Lions, P. L., Renormalized Solutions of Some Transport Equations with Partially $W^{1,1}$ Velocities and Applications. *Annali di Matematica pura ed applicata*, Vol. 183, pp. 97–130, 2004.

[71] Le Bris, C., and Lions, P. L., Existence and Uniqueness of Solutions to Fokker-Planck Type Equations with Irregular Coefficients. *Communications in Partial Differential Equations*, Vol. 33, No. 7, pp. 1272–1317, 2008.

[72] Ambrosio, L., Transport Equation and Cauchy Problem for BV Vector Fields. *Inventiones Mathematicae*, Vol. 158, pp. 227–260, 2004.

[73] Jordan, R., Kinderlehrer, D., and Otto, F., The Variational Formulation of the Fokker-Planck Equation. *SIAM Journal of Mathematical Analysis*, Vol. 29, No. 1, pp. 1–17, 1998.

[74] Gibbs, A. L., and Su, F. E., On Choosing and Bounding Probability Metrics. *International Statistical Review*, Vol. 70, No. 3, pp. 419–435, 2002.

[75] Obereder, A., Scherzer, O., and Kovac, A., Bivariate Density Estimation using BV Regularisation. *Computational Statistics & Data Analysis*, Vol. 51, No. 12, pp. 5622–5634, 2007.

[76] Taylor, M. E., *Pseudodifferential Operators.* Princeton University Press, 1981.

[77] Nguyen, T., The Seiberg-Witten Equations on Manifolds with Boundary II: Lagrangian Boundary Conditions for a Floer Theory. *Arxiv Preprint*, available at `http://arxiv.org/abs/1008.2017`, 2010.

[78] Triebel, H., *Theory of Function Spaces.* Vol. I, Birkhäuser Verlag, Basel, 1983.

[79] DeVore, R. A. and Lorentz, G. G., *Constructive Approximation.* First ed., Springer-Verlag, 1993.

[80] DeVore, R. A., Nonlinear Approximation and Its Applications. *Multiscale, Nonlinear and Adaptive Approximation*, R. A. DeVore and A. Kunoth (eds.), Springer, pp. 169–201, 2009.

[81] David, H. A., and Nagaraja, H. N., *Order Statistics*, Wiley-Interscience, 2003.

[82] Blum, M., Floyd, R. W., Pratt, V., Rivest, R., and Tarjan, R., Time Bounds for Selection. *Journal of Computer and System Sciences*, Vol. 7, pp. 448–461, 1973.

[83] Dor, D., and Zwick, U., Selecting the Median. *SIAM Journal on Computing*, Vol. 28, No. 5, pp. 1722–1758, 1999.

[84] Floyd, R. W., and Rivest, R. L., Expected Time Bounds for Selection. *Communications of the ACM*, Vol. 18, No. 3, pp. 165–172, 1975.

[85] Cunto, W., and Munro, J. I., Average Case Selection. *Journal of the ACM*, Vol. 36, No. 2, pp. 270–279, 1989.

[86] Thaper, N., Guha, S., Indyk, P. and Koudas, N., Dynamic Multidimensional Histograms. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 428–439, 2002.

[87] Zygmund, A., Smooth Functions. *Duke Mathematical Journal*, Vol. 12, No. 1, pp. 47–76, 1945.

[88] DeVore, R. A., Nonlinear Approximation. *Acta Numerica*, Vol. 7, pp. 51–150, 1998.

[89] DeVore, R. A., and Popov, V., Free Multivariate Splines. *Constructive Approximation*, Vol. 3, pp. 239–248, 1987.

# A Besov smoothness of a spline

For the analysis below, we take $1 \leqslant p < \infty$. Unless otherwise mentioned, we consider functions in domain $\Omega \subset \mathbb{R}^d$.

**Lemma 1.** *A piecewise constant polynomial* $S \in \mathrm{Lip}\left(\frac{d}{p}, L_p\right)$.

*Proof.* Without loss of generality, we consider $S(x) \coloneqq \chi_{[0,1]^d}$, where $x = (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d$. Following definition 6 in section 5.1.1, we have

$$\begin{aligned} \| \Delta_h^1 (S,x) \|_{L_p([0,1]^d)}^p &= \| S(x+h) - S(x) \|_{L_p([0,1]^d)}^p \\ &= \int_{[0,1]^d} |S(x+h) - S(x)|^p \, dx \\ &= \int_{-h_1}^0 \int_{-h_2}^0 \cdots \int_{-h_d}^0 1^p \, dx + \int_{1-h_1}^1 \int_{1-h_2}^1 \cdots \int_{1-h_d}^1 1^p \, dx \\ \Rightarrow \omega_r (S,t)_{L_p([0,1]^d)} = \sup_{|h| \leqslant t} \| \Delta_h^1 (S,x) \|_{L_p([0,1]^d)} &= \sup_{|h| \leqslant t} \left( 2 \prod_{i=1}^d h_i \right)^{1/p} = \begin{cases} 2^{1/p} \, t^{1/p} & \text{for } d = 1, \\ 2^{\frac{d+1}{p}} \, d^{-d/2p} \, t^{d/p} & \text{for } d \geqslant 2, \end{cases} \end{aligned} \tag{19}$$

where the constrained optimization has the well-known geometric interpretation that the hyperrectangle of maximum volume inscribed in a given $d$-dimensional sphere of radius $t$, is the hypercube with volume $\left( \frac{2\,t}{\sqrt{d}} \right)^d$. From (19) and (9), we have

$$\omega_r (S,t)_{L_p([0,1]^d)} \lesssim t^{d/p} \Rightarrow S \in \text{Lip} \left( \frac{d}{p}, L_p \right). \tag{20}$$

This completes the proof. □

**Lemma 2** (Zygmund, 1945)**.** $\text{Lip}^\star (\alpha, L_p) = \text{Lip} (\alpha, L_p)$, *for* $0 < \alpha < 1$.

**Remark 8.** *Here is a counterexample illustrating that the restriction over* $\alpha$ *is necessary in the above lemma. Consider the univariate function* $f(x) = x \log x$ *supported over* $[0,1]$. *We want to test whether this function lands up in Lipschitz and/or generalized Lipschitz space with* $\alpha = 1$. *Thus for the weak Lipschitz space,* $r = 2$. *Using series expansion, it can be easily verified that* $|f(x+h) - f(x)| \nleqslant Mh$ *implying* $f(x) \notin \text{Lip}\, 1$. *However,* $|f(x+2h) - 2f(x+h) + f(x)| \leqslant Mh$ *for some constant* $M$, *implying* $f(x) \in \text{Lip}^\star 1$. *Thus the spaces in the above lemma are not same for* $\alpha = 1$. $\text{Lip}^\star 1$ *is also known as Zygmund space after [Zygmund, 1945], who first pointed out this distinction.*

**Lemma 3.** $\text{Lip}^\star (\alpha, L_p) = B_\infty^\alpha (L_p)$.

*Proof.* From definition 7 in section 5.1.1, we know that

$$B_\infty^\alpha (L_p) = \left\{ f : |f|_{B_\infty^\alpha (L_p(\Omega))} \coloneqq \sup_{t>0} \left| t^{-\alpha} \, \omega_r (f,t)_{L_p(\Omega)} \right| = M < \infty \right\}, \qquad r \coloneqq \lfloor \alpha \rfloor + 1. \tag{21}$$

Thus $f \in B_\infty^\alpha (L_p) \Rightarrow t^{-\alpha} \, \omega_r (f,t)_{L_p(\Omega)} \leqslant M$, the value of the Besov supremum quasi-norm in eqn. (21). Rearranging, $\omega_r (f,t)_{L_p(\Omega)} \leqslant Mt^\alpha$ and notice that the condition $r = \lfloor \alpha \rfloor + 1$ appears in both eqn. (7) and (10). Hence from (10), $f \in B_\infty^\alpha (L_p) \Rightarrow f \in \text{Lip}^\star (\alpha, L_p)$. This completes one direction of the proof. For proving the opposite direction $f \in \text{Lip}^\star (\alpha, L_p) \Rightarrow f \in B_\infty^\alpha (L_p)$, one can proceed likewise by starting from (10) to arrive at (7). We omit this for brevity. □

**Proposition 3.** *A piecewise constant polynomial* $S \in B_\infty^{d/p} (L_p)$.

*Proof.* The proof follows by combining lemma 1, 2 and 3. □

Although this proposition suffices for our purpose, more generally, the following holds.

**Theorem 7.** *For an* $r^{th}$ *order spline* $S \in \mathscr{S}^r$, *we have* $S \in B_\infty^{r-1+d/p} (L_p)$.

*Proof.* We sketch the outline of the proof. First, we want to prove by induction that if $S \in \mathscr{S}^r$, then $S^{(r-1)} \in \text{Lip} \left( \frac{d}{p}, L_p \right)$ $\forall\, r \in \mathbb{N}$. For this, lemma 1 need to be used as an induction step for $r = 1$ (piecewise constant spline). For the other induction step (piecewise linear spline), an explicit calculation akin to eqn. (19) confirms the claim for $r = 2$. Together with an inductive hypothesis, we can then establish $S \in \mathscr{S}^r \Rightarrow S^{(r-1)} \in \text{Lip} \left( \frac{d}{p}, L_p \right)$. But then, via lemma 2 and 3, proposition 3 assures $S \in \mathscr{S}^r \Rightarrow S^{(r-1)} \in B_\infty^{d/p} (L_p)$. Finally, invoking the reduction theorem[2] for Besov spaces, we finish the proof. □

[2] In section 5.1.1, we touched upon the reduction theorem for Besov spaces, which says $g^{(r)} \in B_\infty^\alpha \Rightarrow g \in B_\infty^{r+\alpha}$, $\forall r \in \mathbb{N}$. The result is evident from Fig. 1(b) and a formal proof is available in [DeVore and Lorentz, 1993].

# B Performance guarantees for nonlinear approximation

In nonlinear approximation, due to nonuniform partitioning of the support, the approximating space does not satisfy superposition principle. However, when $\Sigma_N \subset \Sigma_{N+1}$ and $\Sigma_M + \Sigma_N \subset \Sigma_{a(M+N)}$, $a \geqslant 1$ fixed, then we can state 'black box theorems' similar to the linear approximation case, for functions in Lipschitz domain $\Omega \subset \mathbb{R}^d$. The exact statements of such theorems, depend on the type of nonlinear approximation being performed, e.g. $N$-term approximation (in wavelet basis or splines), adaptive approximation, tree approximation, greedy algorithms etc. Without detailing these frameworks, we provide the readers a flavor of 'black box type theorems' for two specific kinds of nonlinear approximation: $N$-term approximation and adaptive approximation. As before, our job is to characterize the best rate of approximation. For the proofs, the modus operandi is to establish a Jackson type and a companion Bernstein type inequality and then to leverage the Sobolev type embedding theorem for Besov spaces, as described in section 5.1.1. See [DeVore, 1998, 2009] for nice exposition with necessary details.

## B.1 $N$-term approximation

### B.1.1 Free partition splines

Consider approximating $f \in L_p(\Omega)$ using piecewise polynomials of order $r$, taken from the approximating space $\Sigma_N^r$. The global approximation is $S \coloneqq \sum_{j=1}^{N} p_j(x)\,\chi_{I_j}$, $j = 1, \cdots, N$, $x \in \mathbb{R}^d$, and $p_j(x) \in \mathscr{P}_r\left(\mathbb{R}^d\right)$, the space of all $d$-variate algebraic (real) polynomials of order $r$. We want to estimate the best rate of (nonlinear) approximation $\sigma_N(f)_{L_p(\Omega)} \coloneqq \inf_{S \in \Sigma_N^r} \| f - S \|_{L_p(\Omega)}$.

Unfortunately, no such characterization is known for $d > 1$. Part of the difficulty stems from the fact that the growth condition for the approximating space can not be guaranteed in multivariate setting. Also, characterizing the smoothness space via classical interpolation no longer applies since some evidence suggest that the approximation space $\mathcal{A}_q^\alpha$ becomes a nonlinear function space. Refer [section 6.5, DeVore, 1998] for discussion along these lines.

In univariate set up ($\Omega \subset \mathbb{R}$), $\mathcal{X} = L_p(\Omega)$ and taking $\mathcal{Y}_r = B_\tau^\alpha(L_\tau(\Omega))$, for $0 < \alpha < r$, one can prove [Ch.12, DeVore and Lorentz, 1993] a *Jackson type inequality* for $f \in \mathcal{Y}_r$ of the form

$$\sigma_N(f)_{L_p(\Omega)} \lesssim N^{-\alpha} |f|_{B_\tau^\alpha(L_\tau(\Omega))}, \tag{22}$$

and a companion *Bernstein type inequality* for $S \in \Sigma_N^r$, of the form

$$|S|_{B_\tau^\alpha(L_\tau(\Omega))} \lesssim N^\alpha \| S \|_{L_p(\Omega)}, \tag{23}$$

where $\dfrac{1}{\tau} = \alpha + \dfrac{1}{p} \geqslant 1$. Then the 'black box theorem' for free knot spline guarantees

$$\sigma_N(f)_{L_p(\Omega)} \lesssim N^{-s} \quad \Longleftrightarrow \quad f \in \mathcal{A}_q^s(\mathcal{X}, \mathcal{Y}_r)_{\theta,q} = B_q^s(L_q(\Omega)), \ \frac{1}{q} = s + \frac{1}{p}, \tag{24}$$

where $\theta \coloneqq \dfrac{s}{\alpha}$ and $0 < q \leqslant \infty$, $0 < p < \infty$. The case $p = \infty$ requires a different treatment as in [DeVore and Popov, 1987]. Since $q < p$, the nonlinear approximation space $B_q^s(L_q(\Omega))$ is much larger than the corresponding linear approximation space $B_\infty^s(L_p(\Omega))$, as shown in Fig. 5 for $d = 1$. This reveals the prowess of nonlinear approximation in guaranteeing a desired rate $N^{-s}$ for less regular functions compared to linear approximation.

### B.1.2 Wavelet basis

To leverage the power of nonlinear approximation for $d > 1$, the $N$-term approximation was attempted using box splines [DeVore and Popov, 1987] but was later replaced by computationally more efficient $d$-dimensional wavelets of smoothness $r > 0$. For details, we refer the readers to [section 7.4–7.8, DeVore, 1998]. We state here the crux of the result.

Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. Consider $\mathcal{X} = L_p(\Omega)$ and $\mathcal{Y}_r = B_\tau^\alpha(L_\tau(\Omega))$. Then, for $0 < \alpha < r$, the *Jackson type inequality* for $f \in \mathcal{Y}_r$ is of the form

$$\sigma_N^{\text{wavelet}}(f)_{L_p(\Omega)} \lesssim N^{-\alpha} |f|_{B_\tau^\alpha(L_\tau(\Omega))}, \tag{25}$$

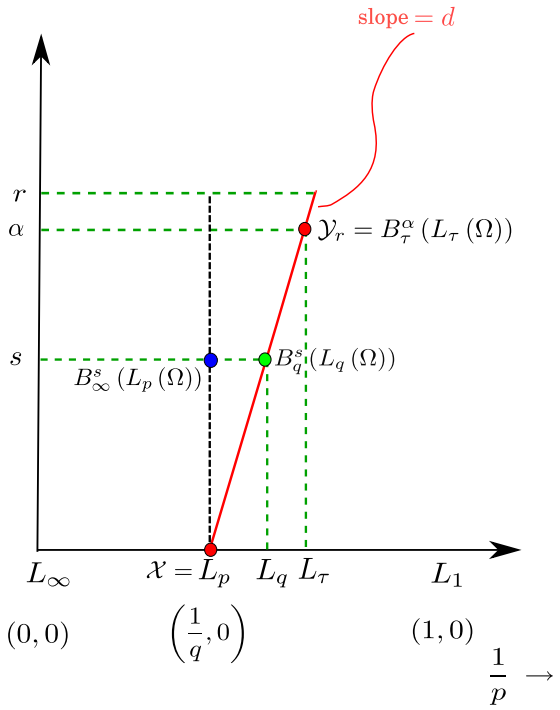


Figure 5: DeVore diagram showing the spaces associated with nonlinear approximation of functions in $\Omega \subset \mathbb{R}^d$. The linear approximation space $B^s_\infty(L_p(\Omega))$ (*blue circle*) is characterized along the vertical line (*black, dotted*) while the nonlinear approximation space $B^s_q(L_q(\Omega))$ (*green circle*) is characterized along (for $N$-term approximation) or to the left (for adaptive approximation) of the Sobolev embedding line (*red, solid*) with slope $d$.

and a companion *Bernstein type inequality* for $S$, an $N$-term wavelet approximation of smoothness $r$, can be written as

$$|S|_{B^\alpha_\tau(L_\tau(\Omega))} \lesssim N^\alpha \parallel S \parallel_{L_p(\Omega)}, \tag{26}$$

where $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p} \geqslant 1$. Then the 'black box theorem' for multivariate wavelet approximation yields

$$\sigma_N^{\text{wavelet}}(f)_{L_p(\Omega)} \lesssim N^{-s/d} \iff f \in \mathcal{A}^s_q(\mathcal{X},\mathcal{Y}_r)_{\theta,q} = B^s_q(L_q(\Omega)), \ \frac{1}{q} = \frac{s}{d} + \frac{1}{p}, \tag{27}$$

where $\theta := \frac{s}{\alpha}$ and $0 < q \leqslant \infty$, $1 < p < \infty$. From Fig. 5, it's evident that given a desired rate, $N$-term wavelets can approximate functions which linear multivariate approximation can not. Conversely, if we desire a best rate (denoted by $\star$) of approximation, then $s^\star_{\text{nonlinear}} \geqslant s^\star_{\text{linear}}$. Depending on how large $s^\star_{\text{nonlinear}}$ is compared to $s^\star_{\text{linear}}$, the best rate of approximation $O\left(N^{-s^\star_{\text{nonlinear}}/d}\right)$ outperforms $O\left(N^{-s^\star_{\text{linear}}/d}\right)$. Finally, note that $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}$ is the equation for Sobolev embedding line (*solid red* in Fig. 5) as described in section 5.1.1. Hence for best $N$-term approximation, the characterization of the approximating space occurs *along the Sobolev embedding line*.

## B.2 Adaptive approximation

If the nonlinear approximation of $f \in L_p(\Omega)$, $\Omega \subset \mathbb{R}^d$, is done by adaptively partitioning the domain in dyadic hypercubes, and we decide to approximate each cell by an $r^{\text{th}}$ order polynomial, then one can prove a 'black box type theorem'. However, this theorem only guarantees a sufficient condition.

Let the approximating space be $\Sigma^{a,r}_N$, where the superscript $a$ stands for adaptive. Here $N$ denotes the size of the final partition, i.e. total number of leaf nodes if the cell division is viewed as a dyadic tree. Then the 'black box type theorem' says: if $f \in B^s_\gamma(L_\gamma(\Omega))$ with $\frac{1}{\gamma} < \frac{s}{d} + \frac{1}{p}$, then $\sigma^a_N(f)_{L_p(\Omega)} \lesssim N^{-s/d}|f|_{B^\alpha_\gamma(L_\gamma(\Omega))}$. It is instructive to note that the theorem requires an approximation space *strictly left to the Sobolev embedding line*.

**Remark 9.** *It can be noted from Theorem 3 that for any general initial density $\rho_0 \in L_1(\Omega)$, the nonlinear approximation can not improve the rate of convergence attained by linear approximation for approximating the*

*joint density, since we can not guarantee the existence of a Besov space on or to the left of the Sobolev embedding line, to which the joint density must belong to.*