Reducing Variance by Reweighting Samples

Mathias Rousset, Yushun Xu, Pierre-André Zitt

REDUCING VARIANCE BY REWEIGHTING SAMPLES

MATHIAS ROUSSET, YUSHUN XU, PIERRE-ANDRÉ ZITT

Abstract. We devise methods of variance reduction for the Monte Carlo estimation of an expectation of the type $\mathbb{E}[\phi(X,Y)]$, when the distribution of X is exactly known. The key general idea is to give each individual of a sample a weight, so that the resulting weighted empirical distribution has a marginal with respect to the variable X as close as possible to its target. We prove several theoretical results on the method, identifying settings where the variance reduction is guaranteed. We perform numerical tests comparing the methods and demonstrating their efficiency.

Contents

2010 *Mathematics Subject Classification.* 65C05: Numerical analysis, Monte Carlo methods.
Key words and phrases. Variance Reduction, Optimal Transport, Control Variates, Post-Stratification.

1. Introduction

1.1. **The framework.** Let (X, Y) be a couple of random variables, and say that we are interested in computing the expected value $\mathbb{E}[\phi(Y)]$ — or more generally $\mathbb{E}[\phi(X, Y)]$ — for each ϕ in some class of test functions. Since the distribution of $\phi(X, Y)$ is most often impossible to obtain in closed analytic form, a classical approach is to resort to Monte-Carlo integration: given an iid sample $(\mathbf{X}; \mathbf{Y}) = (X_1, ..., X_N; Y_1, ..., Y_N)$, the usual "naïve" Monte Carlo estimator is

$$\mathbf{\Phi}_{MC}(\mathbf{X}, \mathbf{Y}) = \frac{1}{N} \sum_{n=1}^{N} \phi(X_n, Y_n). \tag{1}$$

This estimator is unbiased, and its mean square error is given by its variance:

$$\mathrm{MSE}(\mathbf{\Phi}_{MC}) := \mathbb{E}\left[(\mathbf{\Phi}_{MC}(\mathbf{X}, \mathbf{Y}) - \mathbb{E}[\phi(X, Y)])^2\right] = \frac{1}{N} \mathrm{Var}(\phi(X, Y)).$$

The behaviour in N is inescapable and given by the CLT; however, over the years, many *variance reduction* techniques have been devised to reduce the constant multiplicative factor, using various kinds of additional hypotheses on the couple (X, Y). For a general overview of these techniques, see for example the survey paper of Glynn [Gly94], or the book [Ros13]. We introduce in this paper new techniques for reducing variance, which can be seen as a variation on the classical post-stratification method, except that we do not have to fix strata. The method is based on two assumptions on the distribution of the couple (X, Y).

Assumption 1.1. *The distribution of the first marginal X is exactly known: $X \sim \gamma := \mathcal{N}(0, 1)$, the standard Gaussian distribution.*

We note that we could easily accomodate other distributions than the standard Gaussian, which we consider for simplicity; the main point is that we know the distribution of X.

Before introducing the second assumption, let us first recall a classical decomposition of the variance. If we denote by

$$M_\phi(X) := \mathbb{E}[\phi(X, Y)|X], \qquad V_\phi(X) := \mathbb{E}\left[(\phi(X, Y) - M_\phi(X))^2 \middle| X\right],$$

the mean and variance of $\phi(X, Y)$ conditionally on X, then the variance of $\phi(X, Y)$ may be rewritten as the sum of the expected conditional variance and the variance of the conditional expectation, so that the mean square error reads:

$$\mathrm{MSE}(\mathbf{\Phi}_{MC}) = \frac{1}{N}\mathbb{E}[V_\Phi(X)] + \frac{1}{N} \mathrm{Var}(M_\Phi(X)). \tag{2}$$

We now state informally and unprecisely the second assumption, see Corollary 1.16 and Remark 1.17 below for possible more precise statements.

Assumption 1.2. *For the considered test function ϕ, the (random) conditional variance $V_\phi(X)$ is sufficiently 'small' compared to the variance of the conditional expectation* $\mathrm{Var}(M_\phi(X))$.

Under our two assumptions, we devise a *generic method* that estimates $\mathbb{E}[\phi(X, Y)]$ with a smaller variance than the naïve method (1).

Since we do not have the liberty of choosing the values of $(X_n)_{1 \le n \le N}$, but we know exactly their distribution γ, our main idea is to use the samples $\mathbf{X} = (X_1, ..., X_N)$ to devise *random weights* $(w_n(\mathbf{X}))_{1 \le n \le N}$ such that the empirical measure $\sum_n w_n(\mathbf{X})\delta_{X_i}$ is "as close as possible" to the true distribution γ. For instance, for some distance dist () between distributions – the choice of which will be discussed below — we may look for solutions of the minimization problem:

$$\text{minimize: } \mathrm{dist}\left(\gamma, \sum_{n=1}^{N} w_n \delta_{X_n}\right) \quad \text{subject to: } \begin{cases} w_n \ge 0, \\ \sum_n w_n = 1. \end{cases} \tag{3}$$

This minimization problem typically admits a unique solution $(w_1(\mathbf{X}), ..., w_N(\mathbf{X}))$, which can be used instead of the naïve uniform weights $(1/N)$ to estimate $\mathbb{E}[\phi(X, Y)]$ by:

$$\mathbf{\Phi}_W(\mathbf{X}, \mathbf{Y}) = \sum_{n=1}^{N} w_n(\mathbf{X})\phi(X_n, Y_n).$$

In the remainder of this paper, we study this estimator to show, both theoretically and empirically, that it can indeed succeed in reducing the variance with respect to the naïve Monte Carlo method.

1.2. **A decomposition of the mean square error.** The mean square error of our estimator is given by:

$$\begin{aligned}
&\mathrm{MSE}(\mathbf{\Phi}_W) := \mathbb{E}\left[\left(\mathbf{\Phi}_W(\mathbf{X},\mathbf{Y}) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_n w_n(\mathbf{X})\phi(X_n,Y_n) - \sum_n w_n(\mathbf{X})M_\phi(X_n) + \sum_n w_n(\mathbf{X})M_\phi(X_n) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_n w_n(\mathbf{X})(\phi(X_n,Y_n) - M_\phi(X_n))\right)^2\right] + \mathbb{E}\left[\left(\sum_n w_n(\mathbf{X})M_\phi(X_n) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right]
\end{aligned}$$

since the cross terms vanish because $\mathbb{E}\left[\phi(X_n,Y_n) - M_\phi(X_n)|\mathbf{X}\right] = 0$. In the first term, we expand the square, condition on $\mathbf{X}$ and use the conditional independence of the (Y_n); we rewrite the second term using the notation $\eta_N^\star = \sum_n w_n(\mathbf{X})\delta_{X_n}$ and get

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbf{\Phi}_W(\mathbf{X},\mathbf{Y}) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \\
&\qquad = \mathbb{E}\left[\sum_n w_n(\mathbf{X})^2 V_\phi(X_n)\right] + \mathbb{E}\left[\left(\int M_\phi(x)d\eta_N^\star(x) - \int M_\phi(x)\gamma(dx)\right)^2\right] \qquad (4)
\end{aligned}$$

Let us note here that for the naïve choice $w_n(\mathbf{X}) = 1/N$, Equation (4) reduces to the decomposition (2) in two terms of the same order $1/N$. By choosing weights that minimize the distance between the reweighted measure $\eta_N^\star$ and γ, our goal is to make the second term of the right hand side of (4) negligible; for this we pay a price by increasing the first term. If $V_\phi(X)$ is small enough in a suitable sense, this price is expected to be small enough to still be able to decrease the global mean square error. It is the purpose of the main theoretical results of this paper to make this informal statement precise; see Section 1.5, in particular Corollary 1.13 (conditional on the validity of Conjecture 1.11) as well as Corollary 1.16 and Remark 1.17.

In order to give rigorous statements, we make two additional assumptions concerning the test function ϕ and the distance we will use.

Assumption 1.3 (The distance)**.** *The distance dist*() *may be written in operator norm form*

$$dist\,(\eta,\gamma) = \sup_{f\in\mathcal{D}} |\eta(f) - \gamma(f)| \qquad (5)$$

where $\mathcal{D}$ is a set of functions (typically a unit ball of test functions).

Assumption 1.4 (The test function)**.** *There exist two constants m_ϕ, v_ϕ such that:*

- *The conditional mean $x \mapsto M_\phi(x)$ is in the set $\mathcal{D}$ defined in the previous assumption, up to an affine transformation; more precisely, there exists c such that $(M_\phi(\cdot) - c)/m_\phi \in \mathcal{D}$. If $\mathcal{D}$ is the unit ball associated with a norm, the optimal constant m_φ is the associated distance between the line $\{M_\phi - c,\, c \in \mathbb{R}\}$ and 0.*
- *The conditional variance V_ϕ satisfies*

$$V_\phi(X_n) \le v_\phi \qquad a.s. \qquad (6)$$

Assuming this, we get, as an immediate consequence of (4):

$$\mathbb{E}\left[\left(\mathbf{\Phi}_W(\mathbf{X},\mathbf{Y}) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \le v_\phi \mathbb{E}\left[\sum_n w_n(\mathbf{X})^2\right] + m_\phi^2 \mathbb{E}\left[\mathrm{dist}\left(\eta_N^\star,\gamma\right)^2\right]. \qquad (7)$$

We consequently propose to define and compute the weights $(w_n(\mathbf{X}))_{1\le n\le N}$ according to

$$\text{minimize: } \mathrm{dist}\left(\gamma, \sum_{n=1}^N w_n\delta_{X_n}\right) + \delta\sum_n w_n^2, \quad \text{subject to: } \begin{cases} w_n \ge 0, \\ \sum_n w_n = 1, \end{cases} \qquad (8)$$

or more simply to (3) which is obtained from the former by taking $\delta = 0$.

Remark 1.5 (On Assumption 1.4). One could weaken the almost sure bound (6) to a moment condition at the price of stronger constraints on the weights, using for instance Hölder's inequality. Such cases won't be treated here and are left for future work.

Remark 1.6 (On the choice $\delta = 0$). Depending on the choice of the distance dist( ), solving (8) for $\delta \neq 0$, instead of $\delta = 0$ — which is exactly (3) — can be almost free or quite costly numerically, as will be detailed in the next section. Moreover it requires the tuning of another parameter δ. As a consequence, for simplicity and homogoneity, we will mainly focus on the choice $\delta = 0$. From the discussion above, this choice is formally appropriate in the limiting case of observables ϕ for which $v_\phi \ll m_\phi^2$.

To specify completely the algorithm, we need to choose an appropriate distance between probability measures to use in the minimization problem (3)-(8). Among the many possible choices, see e.g. [GS02] for a review, we focus on two choices.

1.3. **A regularized L^2 distance.** The first distance uses the Hilbert structure of the space $L^2(\gamma)$. In this setting, a natural choice for comparing a probability measure η with the target Gaussian distribution γ would be the χ^2 divergence $\int (\frac{d\nu}{d\gamma} - 1)^2 d\gamma$. Since this is degenerate if η is discrete, we need to mollify η in some way before taking this divergence. A natural way of doing this in $L^2(\gamma)$ is to use the Mehler kernel

$$\begin{aligned} K_h(x,y) &:= h^{-1/2} \exp(-(x - \sqrt{1-h}y)^2/(2h)) \exp(x^2/2) \\ &= h^{-1/2} \exp\left(-\frac{\sqrt{1-h}}{2h} (x-y)^2 + \frac{\sqrt{1-h}}{2 + 2\sqrt{1-h}} \left(x^2 + y^2\right) \right). \end{aligned} \tag{9}$$

and map η to $\int K_h(x,y) d\eta(y)$: in probabilistic terms, we replace η by ηP_t where P_t is the Ornstein–Uhlenbeck semigroup, before taking the χ^2 divergence. More formally, for any

$$h = 1 - e^{-2t} \in (0,1),$$

we will see below in Section 3.2 that the formula

$$\|\nu\|_h = \left\| \frac{d(\nu P_t)}{d\gamma} \right\|_{L^2(\gamma)} = \left\| \int K_h(x,y)\, \nu(dy) \right\|_{L^2(\gamma)}, \tag{10}$$

defines a norm on signed measures ν satisfying some moment conditions, and we can set

$$\mathrm{dist}(\eta_1, \eta_2) = \|\eta_1 - \eta_2\|_h$$

in (3). The latter distance has a variational representation as follows

$$\|\nu\|_h = \sup_{\|f\|_{L^2(\gamma)} \leq 1} \nu P_t f,$$

so that the set $\mathcal{D}$ in (5) is the 'regularized' image by the Ornstein–Uhlenbeck semigroup of the unit ball of the Hilbert space $L^2(\gamma)$, and the optimal constant m_ϕ in Assumption 1.4 is defined by (see also Remark 3.5):

$$m_\phi := \sup_{\|\nu\|_h \leq 1 \, \nu(1)=0} \nu M_\phi = \left\| P_t^{-1} \left(M_\phi - \int M_\phi \gamma \right) \right\|_{L^2(\gamma)}. \tag{11}$$

Finally, we will detail in Section 3.4 how the minimization problem (8) turns out to be a quadratic programming convex optimization problem, which can be solved using standard methods, typically with a cubic polynomial complexity in terms of the sample size N. Solving the case $\delta \neq 0$ is in fact easier than the case $\delta = 0$ since larger δ simply improve the conditioning of the symmetric matrix underlying the quadratic porgramming problem.

1.4. **An optimal transport distance.** The second choice we investigate is the Wasserstein distance $\mathcal{W}_1$, defined classically as follows:

Definition 1.7 (Wasserstein distance)**.** *Let (E,d) be a Polish metric space. For any two probability measures η_1, η_2 on E, the Wasserstein distance between η_1 and η_2 is defined by the formula*

$$\begin{aligned} \mathcal{W}_1(\eta_1, \eta_2) &= \left(\inf_{\pi \in \Pi} \int_E d(x_1, x_2) d\pi(x_1, x_2) \right) \\ &= \inf \left\{ \mathbb{E}\left[d(X_1, X_2) \right], \quad Law(X_1) = \eta_1, Law(X_2) = \eta_2 \right\}, \end{aligned} \tag{12}$$

where Π *is the set of all couplings of* η_1 *and* η_2*.*

Kantorovitch duality (see [Vil08]) implies that the latter distance is in fact an operator norm of the form

$$\mathcal{W}_1(\eta_1, \eta_2) = \|\eta_1 - \eta_2\|_{\mathrm{Lip}} = \sup_{\|f\|_{\mathrm{Lip}} \leq 1} \eta_1(f) - \eta_2(f),$$

where $\|f\|_{\mathrm{Lip}} = \sup_{x,y} \frac{f(x)-f(y)}{d(x,y)}$ is the Lipschitz seminorm. This is consistent with choosing for $\mathcal{D}$ in (5) the set of 1-Lipschitz functions.

Finally, we will see in Proposition 4.1 that, at least in dimension 1 for the choice $d(x_1, x_2) = |x_1 - x_2|$, the latter distance leads to an *explicit formula* for the optimal weights $(w_n(\mathbf{X}))_{1 \leq n \leq N}$, that can be computed with a complexity proportional to the sample size N. This leads to faster algorithms and more explicit bounds on the mean square error as compared to the L^2 case of the last section. However, for this optimal transport method, solving the case $\delta \neq 0$ is non explicit and thus harder (although still a convex optimization problem).

Remark 1.8 (On other Wasserstein distances)*.* We do not really lose generality here by only considering the $\mathcal{W}_1$ distance. Indeed, as can be seen from the proof of Proposition 4.1 below, the optimal weights would be the same for any distance $\mathcal{W}_p$, $p \geq 1$.

1.5. **Theoretical results.** Recall that $\mathbf{X} = (X_1, ..., X_N)$ is an i.i.d. $\mathcal{N}(0, 1)$ sequence in $\mathbb{R}$. We denote by

$$\overline{\eta}_N = \frac{1}{N} \sum_n \delta_{X_n}.$$

the empirical measure of the sample $\mathbf{X}$. The reweighted measure $\sum_n w_n(\mathbf{X}) \delta_{X_n}$ will be denoted by:

- $\eta^\star_{h,N}$, if the $w_n(\mathbf{X})$ solve (8) for the L^2 distance with parameter h and δ;
- $\eta^\star_{\mathrm{Wass},N}$, if the $w_n(\mathbf{X})$ solve (3) for the Wasserstein distance (we will only consider the case $\delta = 0$).

We first focus on results on these optimally reweighted measures, shedding light on the behaviour of the bound (7) and especially the second term in it. We start by the L^2 minimization method.

Theorem 1.9 (The L^2 method)**.** *For any fixed* h*,* N *and any* $\delta \geq 0$*, the optimization problem* (8) *with the distance* $\|\cdot\|_h$ *has almost surely a unique solution. The distance of the optimizer* $\eta^\star_{h,N}$ *to the target* γ *satisfies:*

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] \leq \mathbb{E}\left[\|\overline{\eta}_N - \gamma\|_h^2\right] = \frac{1}{N}\left(\frac{1}{h} - 1\right). \tag{13}$$

Moreover, in the case $\delta = 0$*, there exists a numerical* h_0 *such that, for all* $h > h_0$*,*

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] = o(1/N) \tag{14}$$

as N *goes to infinity.*

The following result, where the window size h is allowed to depend on h, is an easy consequence of (13).

Corollary 1.10. *If* $(h_N)_{N \geq 1}$ *is bounded away from* 1 *and satisfies* $Nh_N \to \infty$*, then*

$$\eta^\star_{h_N,N} \xrightarrow{(d)} \gamma \quad \textit{in probability.}$$

The bound given by Equation (14) justifies our strategy in the sense that we managed to decrease significantly the second term in the decomposition (7) of the mean square error. Considering the first term in that decomposition leads naturally to the following conjecture, which is supported by numerical tests.

Conjecture 1.11. *For any* h*, there exists a constant* C_h *such that the optimal weights for the* L^2 *method with* $\delta = 0$ *satisfy*

$$\limsup_N N\mathbb{E}\left[\sum_n w_n(\mathbf{X})^2\right] \leq C_h.$$

Remark 1.12 (The conjecture holds true for $h = 0$). A quick computation based on (9) and Section 3.4 shows that the quadratic minimisation problem (8) in the case $\delta = 0$ and $h = 0$ is equivalent to the minimization over the simplex of the diagonal quadratic form

$$(w_1, \dots, w_N) \mapsto \sum_{n=1}^{N} w_n^2 \exp\left(\frac{1}{2} X_n^2\right),$$

which is solved by $w_n(\mathbf{X}) = Y_n/(\sum_m Y_m)$, where $Y_m = \exp\left(-\frac{1}{2}X_m^2\right)$. Therefore

$$\mathbb{E}\left[N \sum_n w_n^2(\mathbf{X})\right] = \mathbb{E}\left[\frac{N \sum_n Y_n^2}{\left(\sum_m Y_m\right)^2}\right] = \mathbb{E}\left[\frac{N^2 Y_1^2}{(\sum_m Y_m)^2}\right].$$

The random variable $S_N = N^2 Y_1^2/(\sum_n Y_n)^2 \leq N^2$ converges almost surely to $Y_1^2/\mathbb{E}\left[Y_1\right]^2$ by the law of large numbers. Moreover, the S_N are uniformly integrable. Indeed,

$$\mathbb{E}\left[S_N \mathbf{1}_{S_N > K}\right] \leq N^2 \mathbb{P}\left[S_N > K\right] \leq N^2 \mathbb{P}\left[\frac{1}{N}\sum_n Y_n < K^{-1/2}\right],$$

where in the last inequality we have used $Y_1 \leq 1$. If $K^{-1/2} < \mathbb{E}\left[Y_1\right]$, the last probability is exponentially small in N by Hoeffding's inequality so that the uniform integrability follows. Consequently

$$\lim_{N \to +\infty} \mathbb{E}\left[N \sum_n w_n(\mathbf{X})^2\right] = \frac{\mathbb{E}\left[Y_1^2\right]}{\mathbb{E}\left[Y_1\right]^2} = \frac{2}{\sqrt{3}},$$

and the conjecture holds with $C_0 = 2/\sqrt{3}$.

Corollary 1.13. *Assume that Conjecture 1.11 holds true. Let h be larger than the numerical constant h_0 of Theorem 1.9, and assume that M_ϕ is regular enough so that* (11) *is finite, i.e. $m_\phi < +\infty$ – for instance M_ϕ is analytic. Assume also that $V_\phi(x)$ is bounded above by a constant v_ϕ. Then the L^2 method with $\delta \leq v_\phi/m_\phi^2$ satisfies*

$$MSE(\mathbf{\Phi}_W) = \mathbb{E}\left[\left(\Phi_W(\mathbf{X}, \mathbf{Y}) - \mathbb{E}\left[\phi(X, Y)\right]\right)^2\right] \leq \frac{C_h v_\phi}{N} + m_\phi^2 o(1/N)$$

for some numerical constant C_h and numerical $o(1/N)$.

Therefore, under the above assumptions, the L^2 method is asymptotically better than the naïve Monte Carlo approach in terms of MSE as soon as $v_\phi \leq \mathrm{Var}(M_\phi(X))/(C_h - 1)$.

Remark 1.14 (On the optimal choice of h). The question of the best choice for the smoothing parameter h is not easy to tackle: in the upper bound (7), h appears in the weights *via* C_h, in the distance and in m_ϕ. We will give below theoretical and empirical evidence that the best choice is related to the regularity of the test function ϕ, and that smaller h are needed if ϕ is very irregular.

We are able to prove similar but more complete results for the Wasserstein method.

Theorem 1.15 (The Wasserstein method)**.** *For any N, the optimization problem has almost surely a unique solution. The distance $D = \mathcal{W}_1\left(\eta^\star_{Wass,N}, \gamma\right)$ of the optimizer $\eta^\star_{Wass,N}$ to the target γ satisfies for all integer $p \geq 1$ the moment bounds:*

$$\mathbb{E}\left[D^p\right] = \mathcal{O}^\star\left(\frac{1}{N^p}\right),$$

where $\mathcal{O}^\star$ means $\mathcal{O}$ up to logarithmic factors. In particular,

$$\eta^\star_{Wass,N} \xrightarrow[N \to +\infty]{Law} \gamma \quad \textit{in} \quad \textit{in probability.}$$

Moreover, the optimal weights satisfy:

$$\mathbb{E}\left[\sum_n w_n(\mathbf{X})^2\right] \leq \frac{6}{N}.$$

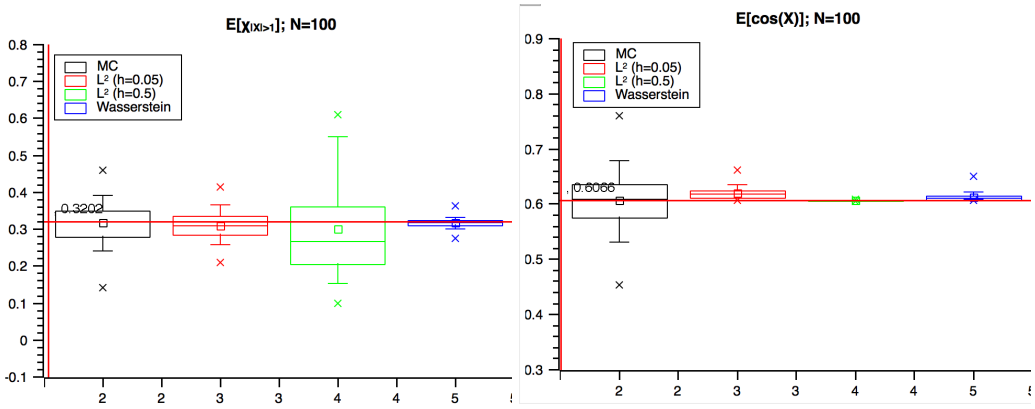


FIGURE 1. Comparison of methods for $\phi(X)$ and $\delta = 0$

Corollary 1.16. *Assume M_ϕ is Lipschitz – so that $m_\phi < +\infty$, and that $V_\phi(x)$ is bounded above by a constant v_ϕ. Then*

$$MSE(\mathbf{\Phi}_W) = \mathbb{E}\left[\left(\Phi_W(\mathbf{X},\mathbf{Y}) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \leq \frac{c_0 v_\phi}{N} + m_\phi^2 o(1/N)$$

for some numerical constant $c_0 \leq 6$ and numerical $o(1/N)$.

Therefore, under the above assumptions, the Wasserstein method is asymptotically better than the naïve Monte Carlo approach in terms of MSE as soon as $v_\phi \leq \mathrm{Var}(M_\phi(X))/(c_0 - 1)$.

Remark 1.17 (The optimal c_0). A careful look at the proof shows that, if the X_i follow the uniform distribution on $[0,1]$, the bound on $\mathbb{E}\left[\sum_n w_n(\mathbf{X})^2\right]$ may be divided by 4, leading to $c_0 = 3/2$. Numerical tests suggest that even in the Gaussian case, this bound still holds true asymptotically in the sense that $N\mathbb{E}\left[\sum_n w_n(\mathbf{X})^2\right] \to \frac{3}{2}$. Therefore, we conjecture that the Wasserstein method is better than the naïve Monte Carlo approach as soon as $v_\phi \leq 2\,\mathrm{Var}(M_\phi(X))$.

1.6. **Numerical experiments.** We supplement our theoretical findings with numerical tests. In the first series of tests, we compare numerically the naïve Monte Carlo method, the L^2 method with various choices of the bandwidth, and the Wasserstein method, in the toy case where X itself is the variable of interest. For simplicity, and for homogeneity between the two methods, we have chosen in this first serie of numerical tests $\delta = 0$ (up to numerical precision) in the the L^2-method. This case is an idealized case of concrete problems where $v_\phi \ll m_\phi^2$.

The full results may be found in Section 5. Figure 1 shows that both methods perform much better than the naïve Monte Carlo estimator. The L^2 method is often able to reduce significantly the statistical error, but the bandwidth parameter h must be chosen carefully, depending on N and the type of observable we are interested in. The parameter-free Wasserstein method is faster and more robust, but may be outperformed by a well-tuned L^2 method for very regular observables (for example the cosine function).

The second series of numerical tests in done in Section 6. We let G be a d-dimensional standard Gaussian vector, and assume we are interested in the distribution of a non-linear function $Y = F(G)$. We linearize F near 0 and let $X = (DF)_0 G$, so that the distribution of X is an explicit one dimensional Gaussian. We then use our method to estimate, for any fixed t, the cumulant generating function $\log \mathbb{E}\left[\exp(tF(G))\right]$, using X as our "control variable". In this more realistic setting, we focus on the more robust Wasserstein method, and show how it can be compared to, and combined with, a more classical control variate approach.

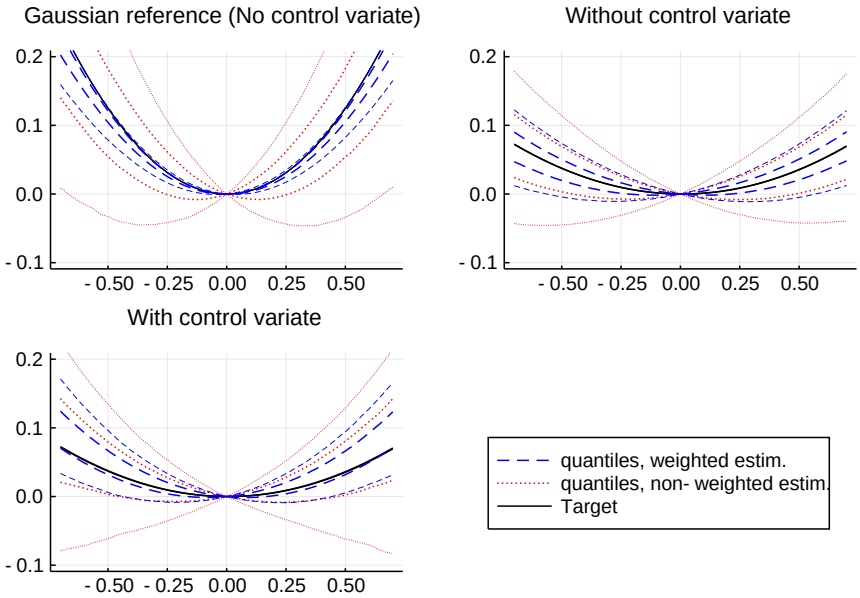


FIGURE 2. The figures above represents the $[.05, .25, .75, .95]$-quantile enveloppes of the different estimators of the cumulant generating functions for $F(G)$.

In Figure 2, we clearly see that our reweighting method always reduces variance in a substantial way, without using any prior information of F. This non-linear example also shows that a standard variance reduction by a linear control variate may be useless.

1.7. **Conclusion.** We have proposed *generic and robust* variance reduction techniques based on reweighting samples using a one dimensional Gaussian control variable. The latter can be seen as generalization of post-stratification methodologies and can outperform variance reduction by control variate even in simple situations.

Theoretically, the results, which prove effective variance reduction for both methods, are quite similar. The main difference is that in the Wasserstein setting, we are actually able to control the variance of the weights with an explicit constant $c_0 \leq 6$; in the L^2 case we only conjecture that a similar result holds. This difference essentially comes from the fact that in our one dimensional setting, the optimal Wasserstein weights are explicit, and therefore much easier to study.

Numerically, we have observed (as theoretically suggested) that the L^2 approach, as compared to the Wasserstein approach, requires more regular obervables and some tuning, and is more costly when the sample size become large. Note however, that the L^2 approach may be amenable to control variables X in higher dimension, where Wasserstein optimization — optimal transport — problems are known to be very cumbersome. This issue is left for future work.

1.8. **Outline of the paper.** In Section 2, we briefly discuss how our method fits in the landscape of variance reduction techniques, and how it can be seen as complementing control variates and generalizing post-stratification. In Section 3 we discuss the L^2 method; the results on the Wasserstein approach are established in Section 4. The first numerical tests, considering only the gaussian variable X, are presented in Section 5. Finally, the tests on more realistic models are presented in Section 6.

2. Comparison with classical methods

2.1. **Comparison to variance reduction with control variates.** The method presented in this work can be seen as an alternative to control variates. More precisely, as we are about to explain now, they can be interesting as a complement to control variates when the latter is either not efficient or too expensive.

Within the framework of Section 1.1, a control variate is a computable function ψ of X such that the mean square error of $\mathbb{E}\left[(\phi(X,Y) - \psi(X))^2\right]$ is as small as possible (see [Owe13, Gla13]

for a general introduction). Recent works have studied various techniques to find an optimized ψ using basic functions and a Monte Carlo approach (see for instance [Jou09, PS18]). The associated estimator is then

$$\mathbf{\Phi}_{CV}(\mathbf{X},\mathbf{Y}) = \frac{1}{N}\sum_{n=1}^{N}\phi(X_n,Y_n) - \psi(X_n) + \int \psi d\gamma,$$

whose mean square error (identical to variance since it is unbiased), satisfies

$$\mathrm{MSE}(\mathbf{\Phi}_{CV}(\mathbf{X},\mathbf{Y})) = \frac{1}{N}\mathbb{E}\left[\left(\phi(X,Y) - \psi(X) + \int \psi d\gamma - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right].$$

This quantity is classically minimized by choosing $\psi(X) = \mathbb{E}\left[\phi(X,Y)|X\right]$, up to an irrelevant additive constant.

As a consequence, if a good control variate, close to the conditional expectation $\psi(X) \sim \mathbb{E}\left[\phi(X,Y)|X\right]$ is available, and if we try to apply our method to $\tilde{\phi}(X,Y) = \phi(X,Y) - \psi(X)$, then Assumption 1.2 *will not hold* for $\tilde{\phi}$. In the numerical experiment done in Section 6 — see also Figure 2 — we compare the reweighting method with a natural affine control variate $\Psi(X) = aX + b$ which is not sufficient to approximate correctly $\mathbb{E}\left[\phi(X,Y)|X\right]$; interestingly the reweighting method is able to overcome this issue in a generic way, without specific analytic approximation of $\mathbb{E}\left[\phi(X,Y)|X\right]$ contrary to what is required to improve the control variate.

The latter discussion suggests that the present variance reduction method based on re-weighting samples will be useful in one of the following two situations:

- The available control variates behaves very poorly.
- One is interested in estimating $\mathbb{E}\left[\phi(X,Y)\right]$ for a large class of test functions ϕ, making the calculation of control variates very costful.

2.2. **Comparison to post-stratification variance reduction.** The present work may be interpreted as a generalization of post-stratification methods to continuous state spaces. In the framework of Section 1.1, post-stratification can be defined by first choosing a finite partition of $\mathbb{R}$, given by K 'strata', for instance the K-quantiles $x_{1/2} < \ldots < x_{K-1/2}$ defined by $\int_{x_{k-1/2}}^{x_{k+1/2}} d\gamma = 1/K$.

In that context, the post-stratification weights are defined by

$$\begin{cases} w_n(\mathbf{X}) = \frac{1}{KB_n(\mathbf{X})}, \\ B_n(\mathbf{X}) = \mathrm{card}\left\{X_m \in \mathrm{strat}(X_n),\, 1 \le m \le N\right\}, \end{cases} \tag{15}$$

where in the above $\mathrm{strat}(X_n)$ is the interval $[x_{k-1/2}, x_{k+1/2}[$ containing X_n. The latter post-stratification weights are defined so that the sum of the weights of particles in a given stratum is constant an equal to $1/K$.

We can then check the following:

Lemma 2.1. *Let us denote by $\mathcal{D}_K$ the space of functions on $\mathbb{R}$ that are constant on the srata $[x_{k-1/2}, x_{k+1/2}[$, for $k = 1 \ldots K$. Consider the semi-norm over finite measures*

$$p_K(\mu) = \sup_{\psi\in\mathcal{D}_K,\, \|\psi\|_\infty \le 1} \mu(\psi).$$

Then the post-stratification weights (15) *is the solution to the minimization problem obtained by setting $dist\,(\eta,\gamma) = p_K\left(\eta - \gamma\right)$ in* (3) *that is*

$$\textit{minimize: } p_K\left(\gamma - \sum_{n=1}^{N} w_n\delta_{X_n}\right), \quad \textit{subject to: } \begin{cases} w_n \ge 0, \\ \sum_n w_n = 1, \end{cases} \tag{16}$$

that moreover minimize the variance of the weight $\sum_n w_n^2 - 1$.

Proof. It's easy to check that, by definition,

$$p_K\left(\gamma - \sum_{n=1}^{N} w_n\delta_{X_n}\right) = \sum_{k=1}^{K}\left|\frac{1}{K} - \sum_n w_n \mathbf{1}_{X_n\in[x_{k-1/2},x_{k+1/2}[}\right|$$

As a consequence, the weights that are solution to the minimization (16) are exactly those such that for all $1 \le m \le n$

$$\sum_n w_n \mathbf{1}_{X_n\in\mathrm{strat}(X_m)} = 1/K$$

which means that the sum of the weights of particles in the same stratum are equal to $1/K$. Now the unique minimum of $\sum_{n\in I} w_n^2$ under the constraint that $\sum_{n\in I)} w_n = c_0$ is constant is given by uniform weights $w_n = c_0/\mathrm{card}(I)$ since by Jensen

$$\sum_{n\in I} (c_0/\mathrm{card}(I))^2 = \left(\sum_{n\in I} w_n\right)^2 /\mathrm{card}(I) \leq \sum_{n\in I} w_n^2. \qquad \square$$

As a consequence, the methods presented in this work can be interpreted as extensions of the post-stratification methods from the semi-norm p_K to the norms $\|\,.\,\|_h$ or to the Wasserstein distance (which is in fact a norm) $\mathcal{W}_1$.

3. The L^2 method

In this section, after recalling a few classical facts and formulae on the Hilbert space $L^2(\gamma)$, we define precisely the h-norm on signed-measures, and study in Section 3.4 the corresponding optimization problem. Finally, we prove in Section 3.5 and 3.6 the results announced in Theorem 1.9.

3.1. Useful tools in L^2.

We start by recalling a few useful definitions and results concerning the standard Gaussian Hilbert space $L^2(\gamma)$.

Orthogonalizing the standard polynomial basis with respect to the scalar product $\langle f, g\rangle_\gamma = \int f(x)g(x)d\gamma(x)$ gives rise to the classical family of *Hermite polynomials* (H_n), see e.g. [AS92, Chapter 22] for details: a Hilbert basis of $L^2(ga)$, where H_n is a polynomial of degree n, with the normalization $\langle H_m, H_n\rangle = m!\mathbf{1}_{m=n}$. We write h_n the corresponding orthonormal basis $h_n = (n!)^{-1/2}H_n$.

Recall the definition of the Mehler kernel (9):

$$\begin{aligned} K_h(x,y) &= h^{-1/2}\exp(-(x-\sqrt{1-h}y)^2/(2h))\exp(x^2/2) \\ &= h^{-1/2}\exp\left(-\frac{\sqrt{1-h}}{2h}(x-y)^2 + \frac{\sqrt{1-h}}{2+2\sqrt{1-h}}\left(x^2+y^2\right)\right). \end{aligned}$$

It will be useful to introduce another parameter t such that $h = 1-e^{-2t}$; we let

$$k_t(x,y) = K_h(x,y) = K_{1-e^{-2t}}(x,y).$$

The classical formula of Mehler gives the spectral decomposition of this kernel.

Lemma 3.1 (Mehler's formula)**.** *For all $(x,y)\in\mathbb{R}^2$ and $t>0$,*

$$k_t(x,y) = \sum_{n=0}^{\infty} e^{-nt}h_n(x)h_n(y) = k_t(y,x). \tag{17}$$

It is also classical to interpret this kernel as the probability density kernel of the Ornstein-Uhlenbeck semigroup with respect to the standard Gaussian.

Lemma 3.2. *Let P_t denotes the semigroup of probability transitions of the Ornstein–Uhlenbeck process solution to the SDE $dX_t = -X_t dt + \sqrt{2}B_t$ where B_t is a standard Brownian motion, that is $\mathbb{E}[f(X_t)|X_0\sim\eta] = \int P_t f(x)d\eta(x)$ for all bounded continuous test function f and any probability measure η. Then it holds*

$$P_t(x,dz) = k_t(x,z)\gamma(dz).$$

Proof. X_t has the same distribution as $e^{-t}X_0 + \sqrt{1-e^{-2t}}G$ for a standard Gaussian random variable G. Hence recalling (9) and $h = 1-e^{-2t}$ it yields $\mathbb{E}[f(X_t)|X_0 = x] = \int f(z)k_t(x,z)\gamma(dz)$. $\square$

Let us collect a few consequences of this representation.

Lemma 3.3. *Let $\gamma = \mathcal{N}(0,1)$, then*

$$\int K_h(x,y)\,\gamma(dy) = \int K_h(x,y)\gamma(dx) = 1; \tag{18}$$

$$\int k_s(x,y)k_t(y,z)\gamma(dy) = k_{s+t}(x,z); \tag{19}$$

$$\int\int K_h(x,y)^2\gamma(dx)\gamma(dy) = \frac{1}{h}. \tag{20}$$

Proof. For the first equality, we integrate (17) with respect to one of the variables:

$$\begin{aligned}\int k_t(x,y)\gamma(dy) &= \int \sum_{n=0}^{\infty} e^{-nt} h_n(x) h_n(y)\gamma(dy) \\ &= \sum_{n=0}^{\infty} e^{-nt} h_n(x) \int h_n(y)\gamma(dy) = h_0(x) = 1.\end{aligned}$$

The second equality is another way of expressing the semigroup property for the Ornstein Uhlenbeck process: For all $x,\ z \in \mathbb{R}$

$$\begin{aligned}\int k_t(x,y)k_s(y,z)\gamma(dy) &= \int \sum_{m,n} e^{-mt-ns} h_m(x)h_m(y)h_n(y)h_n(z)\gamma(dy) \\ &= \sum_{n} e^{-n(t+s)} h_n(x)h_n(z) \\ &= k_{t+s}(x,z).\end{aligned}$$

Applying this to the special case $x = z$ and $t = s$, for $h = 1 - e^{-2t}$, and integrating with respect to $\gamma(dx)$ yields

$$\begin{aligned}\int\int K_h(x,y)^2\gamma(dx)\gamma(dy) &= \int k_{2t}(y,y)\gamma(dy) = \int \sum_{n=0}^{\infty} e^{-2nt} h_n(y)^2\gamma(dy) \\ &= \sum_{n=0}^{\infty} e^{-2nt} = \frac{1}{1-e^{-2t}} = \frac{1}{h}.\end{aligned}$$

□

3.2. **The h-norm: theoretical properties.** We gather the definition and main properties of the h-norm in the following result.

Theorem 3.4. *Let*

$$h = 1 - e^{-2t}, \qquad t > 0.$$

Let $\mathcal{M}$ be the set of signed measures on $\mathbb{R}$ with a finite total mass, and let

$$\mathcal{S} = \left\{ \nu \in \mathcal{M}, \int \exp\left(\frac{y^2}{4}\right) |\eta|\,(dy) < +\infty \right\}.$$

(1) For any $\nu \in \mathcal{S}$, the function $x \mapsto \int K_h(x,y)d\nu(y)$ is in $L^2(\gamma)$, so that

$$\|\nu\|_h := \left\| \frac{d\,(\nu P_t)}{d\gamma} \right\|_{L^2(\gamma)} = \left\| \int K_h\,(x,y)\,\nu(dy) \right\|_{L^2(\gamma)} < \infty.$$

(2) Let us denote $\mathcal{D} := \{P_t\varphi; \varphi \in L^2(\gamma), \|\varphi\|_{L^2(\gamma)} \le 1\}$ the unit ball of the space $\{\psi = P_t\varphi; \varphi \in L^2(\gamma)\}$ endowed with the norm $\left\|P_t^{-1}\psi\right\|_{L^2(\gamma)}$. Then $\|\nu\|_h$ admits the dual representation

$$\|\nu\|_h = \sup_{\|\varphi\|_{L^2(\gamma)} \le 1} \int \varphi \frac{d\,(\nu P_t)}{d\gamma} d\gamma = \sup_{\left\|P_t^{-1}\psi\right\|_{L^2(\gamma)} \le 1} \int \psi d\nu = \sup_{\psi \in \mathcal{D}} \int \psi d\nu. \tag{21}$$

(3) The map $\nu \mapsto \|\nu\|_h$ is a norm on $\mathcal{S}$.
(4) $\|\cdot\|_{h'} \le \|\cdot\|_h$, when $h \le h'$.
(5) If $(\eta_k)_{k\in\mathbb{N}}$ and η are probability measures in $\mathcal{S}$, and if $\|\eta_k - \eta\|_h \to 0$, then η_k converges weakly to η.

Remark 3.5 (On the set $\mathcal{D}$). The set $\mathcal{D}$ is the image of the unit ball of L^2 by the Ornstein-Uhlenbeck semigroup and consists of very regular functions. Indeed, its coefficents on the Hermite basis must decrease geometrically: $\psi \in \mathcal{D}$ if and only if

$$\|\psi\|_{L^2(\gamma)} = \sum_{k\ge 0} \mathrm{e}^{kt} \left(\int_{\mathbb{R}} h_k \psi d\gamma \right)^2 \le 1.$$

The set $\mathcal{D}$ contains of course all conveniently normalized polynomials, as well as many explicit non-polynomial functions. For instance, the cosine function is in $\mathcal{D}$, as can be checked thanks to the computations of $P_t e^{i\lambda\cdot}$ in the proof below.

Proof. (1) By Minkowski's integral inequality, we have

$$\begin{aligned}\|\nu\|_h &= \left(\int\left(\int K_h(x,y)\nu(dy)\right)^2\gamma(dx)\right)^{\frac{1}{2}}\\ &\le \int\left(\int (K_h(x,y))^2\gamma(dx)\right)^{\frac{1}{2}}|\nu|(dy).\end{aligned}$$

By (19) from Lemma 3.3 applied with $x=z$ and $s=t$, and a quick computation using the explicit formula for k_{2t}, we can rewrite the innermost integral as follows:

$$\int K_h(x,y)^2\gamma(dx) = k_{2t}(y,y) = (1-e^{-4t})^{-1/2}\exp\left(\left(\frac{e^{-2t}}{1+e^{-2t}}\right)y^2\right).$$

Therefore, $\|\nu\|_h$ is finite whenever $\exp\left(\left(\frac{e^{-2t}}{2(1+e^{-2t})}\right)y^2\right)$ is $|\nu|$ integrable. In particular it is finite for all h if $|\nu|$ integrates $\exp(x^2/4)$.

(2) The first equation uses the Hilbert structure of $L^2(\gamma)$ and the second one follows from the fact that the Ornstein–Uhlenbeck semigroup P_t is self adjoint in $L^2(\gamma)$.

(3) Homogeneity and sub-additivity follow easily from the dual expression (21). Since the positivity $\|\eta\|_h \ge 0$ is obvious, it is enough to prove that $\|\eta\|_h = 0$ implies $\eta = 0$. We prove this fact using characteristic functions. Denote by $\mathcal{F}(f)$ the Fourier transform, for any function $f:\mathbb{R}\to\mathbb{C}$

$$\mathcal{F}(f)(\xi) := \int e^{-2\pi i x\cdot\xi} f(x)dx.$$

We recall that, for any $a>0$,

$$\mathcal{F}(e^{-ax^2})(\xi) = \sqrt{\frac{\pi}{a}}\exp\left(-\frac{\xi^2}{\frac{a}{\pi^2}}\right). \tag{22}$$

Now, remark that the Ornstein Uhlenbeck semigroup P_t leaves the set of functions $\{x\mapsto ce^{i\lambda x}\}$ invariant: indeed, for any $\lambda\in\mathbb{R}$, let $\tilde\lambda = \frac{\lambda}{\sqrt{1-h}}$ and $c = e^{\frac{h\lambda^2}{2(1-h)}}$, we have

$$\begin{aligned}P_t\left(ce^{i\tilde\lambda\cdot}\right)(x) &= \mathbb{E}\left[ce^{i\tilde\lambda\left(\sqrt{1-h}x+\sqrt{h}G\right)}\right] = ce^{i\sqrt{1-h}\tilde\lambda x}\mathbb{E}\left[e^{i\tilde\lambda\sqrt{h}G}\right]\\ &= ce^{i\lambda x}\int e^{-i(-\sqrt{h}\tilde\lambda)x}\gamma(dx),\end{aligned}$$

by Fourier transform (22), we get

$$\begin{aligned}P_t\left(ce^{i\tilde\lambda\cdot}\right)(x) &= ce^{i\lambda x}\int e^{-i(-\sqrt{h}\tilde\lambda)x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\\ &= ce^{i\lambda x}e^{-\frac{h\tilde\lambda^2}{2}}\\ &= e^{i\lambda x}.\end{aligned}$$

Since $\left\|ce^{i\tilde\lambda\cdot}\right\|_{L^2(\gamma)} = ce^{-\frac{\tilde\lambda^2}{2}} \le |c|$, then for any $\lambda\in\mathbb{R}$

$$\begin{aligned}\left|\nu(e^{i\lambda\cdot})\right| &= \left|\nu\left(P_t\left(ce^{i\tilde\lambda\cdot}\right)\right)\right| \le |c|\sup_{\|\phi\|_{L^2}\le 1}\left|(\nu)\left(P_t(\phi)\right)\right|\\ &= |c|\,\|\nu\|_h.\end{aligned} \tag{23}$$

Therefore, if $\|\nu\|_h = 0$, then $\left|\nu(e^{i\lambda\cdot})\right| = 0$ for all λ, which implies $\nu = 0$.

(4) Let $h' = 1-e^{-2t'}$, then $t\le t'$. By definition of $\|\cdot\|_h$, we have

$$\begin{aligned}\|\nu\|_{h'} &= \sup_{\|\varphi\|_{L^2}\le 1}\nu(P_{t'}\varphi) = \sup_{\|\varphi\|_{L^2}\le 1}\nu(P_tP_{t'-t}\varphi)\\ &\le \sup_{\varphi:\left\|P_{t'-t}\varphi\right\|\le 1}\nu(P_tP_{t'-t}\varphi)\\ &= \sup_{\|\varphi\|_{L^2}\le 1}\nu(P_t\varphi)\\ &= \|\nu\|_h,\end{aligned}$$

where we have use that $\|P_{t'-t}\varphi\|_{L^2} \leq \|\varphi\|_{L^2}$ by Jensen's inequality.

(5) Let (η_k) and η be probability measures in $\mathcal{S}$ such that $\|\eta_k - \eta\| \to 0$. For any λ and any k, we apply (23) to $\nu = \eta_k - \eta$ and let k go to infinity. This implies that $\eta_k(e^{i\lambda\cdot})$ converges to $\eta(e^{i\lambda\cdot})$ for all λ, so η_k converges to η in distribution. □

3.3. **Choice of the bandwidth h.** We have seen in Theorem 3.4 that the mapping $h \mapsto \|\,\|_h$ is decreasing, whereas the mapping $h \mapsto m_\phi := \sup_{\|\nu\|_h \leq 1} \nu(M_\phi)$ is increasing. If one tries to minimize the second term $m_\phi^2 \mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right]$ in the upper bound (7), one can easily check that its derivative with respect to h has the same sign as

$$\frac{d}{dh}\ln m_\phi^2 + \frac{d}{dh}\ln \mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] \tag{24}$$

On the other hand, in the Hermite polynomials orthonormal basis, we have the simple formula (see Remark 3.5) $m_\phi^2 = \sum_{k\geq 1} \mathrm{e}^{kt}\left(\int h_k M_\phi d\gamma\right)^2$ with $h = 1 - \mathrm{e}^{-2t}$, so that

$$\frac{d}{dh} m_\phi^2 = \frac{dt}{dh}\sum_{k\geq 1} k\mathrm{e}^{kt}\left(\int h_k M_\phi d\gamma\right)^2,$$

and thus the ratio $\frac{d}{dh}\ln m_\phi^2 = m_\phi^{-2}\frac{dm_\phi^2}{dh}$ can be interpreted as a strong measure of the irregularity of $x \mapsto M_\phi(x)$ — the more the observable M_ϕ is 'irregular', the more the high frequency modes are relatively large and the larger $\frac{d}{dh}\ln m_\phi^2$ is. As a consequence, for any fixed h, a less regular observable M_ϕ renders the gradient (24) strictly positive, showing that the minimizer of $h \mapsto m_\phi^2 \mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right]$ is attained for *smaller* h — that is, as expected, for smaller kernel bandwidths. This monotony between the best choice of bandwidth h and the regularity of the observable will be observed numerically in Section 5.

3.4. **The L^2 method as a quadratic programming problem.** We now discuss the minimization problem (3) when dist () is the h-norm, first from a deterministic point of view. Let $\mathbf{x} = (x_1, ..., x_N)$ be a vector in $\mathbb{R}^N$. We want to solve the following minimization problem :

$$\text{minimize: } \left\|\sum_n w_n \delta_{x_n} - \gamma\right\|_h^2 + \delta\sum_n w_n^2, \quad \text{subject to: } \begin{cases} w_n \geq 0, \\ \sum_n w_n = 1. \end{cases}$$

Let us denote by Ω the simplex $\{w = (w_1, ..., w_N) | w_i \geq 0,\ \sum_{n=1}^N w_n = 1\}$, and let $F(w) = \|\sum_n w_n \delta_{x_n} - \gamma\|_h^2$. By definition, F may be rewritten as follows:

$$\begin{aligned} F(w) &= \left\|\sum_{n=1}^N w_n K_h(y, x_n) - 1\right\|^2_{L^2(\gamma(dy))} \\ &= \int\left(\sum_{n=1}^N w_n K_h(y, x_n) - 1\right)^2 \gamma(dx) \\ &= \sum_{n,m}^N w_n w_m \int K_h(y, x_n) K_h(y, x_m)\gamma(dx) - 2\sum_{n=1}^N w_n \int K_h(y, x_n)\gamma(dx) + 1 \end{aligned}$$

By (19) in Lemma 3.3, we have

$$\begin{aligned} F(w) &= \sum_{n,m}^N w_n w_m k_{2t}(x_n, x_m) - 2\sum_{n=1}^N w_n + 1 \\ &= w^\top Q w - 1, \end{aligned}$$

where Q is the $N \times N$ matrix whose components are given by

$$Q_{n,m} = k_{2t}(x_n, x_m), \quad \text{for any } 1 \leq n, m \leq N. \tag{25}$$

For future reference, let us note that using Mehler's formula, $k_{2t}(x_m, x_n) = \sum_k e^{-2kt} h_k(x_m) h_k(x_n)$, so that we can also write, for any weight vector $(w_1, ..., w_N)$,

$$\begin{aligned} \left\| \sum_{n=1}^N w_n \delta_{x_n} - \gamma \right\|_h^2 = w^\top Q w - 1 = \sum_{k \geq 0} e^{-2kt} \left(\sum_n w_n h_k(x_n) \right)^2 - 1 \\ = \sum_{k \geq 1} e^{-2kt} \left(\sum_n w_n h_k(x_n) \right)^2 . \end{aligned} \tag{26}$$

The minimization problem is therefore reduced to the following quadratic problem over a convex set:

$$\text{minimize: } w^\top (Q + \delta \mathrm{Id}) w \quad \text{subject to: } w \in \Omega. \tag{27}$$

Proposition 3.6 (The quadratic programming problem)**.** *If the x_n are pairwise distinct, then Q is positive definite, and the minimization problem* (27) *has a unique solution even if $\delta = 0$.*

This holds in particular with probability one if the $(x_n)_{1 \leq n \leq N} = (X_n)_{1 \leq n \leq N}$ are iid samples of γ.

Remark 3.7. Here the solution may or may not be in the interior of the simplex: there are vectors $\mathbf{x} = (x_1, ..., x_N)$ for which some of the components of the optimal weight vector are zero.

Proof. For any column vector $a = (a_1, ..., a_N)$, we need to prove that $a^T Q a = 0$ if and only if $a = 0$. The same calculation leading to (26) yields

$$a^T Q a = \sum_k e^{-2kt} \left(\sum_m a_m h_k(x_m) \right)^2 = 0,$$

which implies that, for all integer k, $\sum_{m=1}^N a_m h_k(x_m) = 0$. Let P_n be the Lagrange cardinal polynomial that satisfies $P_n(x_m) = \delta_{nm}$. Since P_n may be decomposed on the basis of the h_k, it holds that $\sum_m a_m P_n(x_m) = 0$, so a_n must be zero. Since n is arbitrary, $a = 0$. This shows that Q is positive definite.

We are therefore optimizing a strictly convex function over a compact convex set: the minimizer exists and is unique. □

3.5. **A first comparison with the naïve empirical measure.** Recall that $\overline{\eta}_N = \frac{1}{N} \sum_n \delta_{X_n}$ and $\eta^\star_{h,N} = \sum_n w_n(\mathbf{X}) \delta_{X_n}$ denote respectively the naïve and L^2-reweighted empirical measure.

Proof of Theorem 1.9. The existence and uniqueness of the minimizer follow from Proposition 3.6.

We now want to establish (13), that is,

$$\mathbb{E}\left[\left\| \eta^\star_{h,N} - \gamma \right\|_h^2 \right] \leq \mathbb{E}\left[\left\| \overline{\eta}_N - \gamma \right\|_h^2 \right] = \frac{1}{N} \left(\frac{1}{h} - 1 \right).$$

The first inequality follows from Jensen's inequality. Indeed, by definition, the (w_n) solve (27), so that almost surely,

$$\left\| \eta^\star_{h,N} - \gamma \right\|_h^2 + \delta \sum_n w_n(\mathbf{X})^2 \leq \left\| \overline{\eta}_N - \gamma \right\|_h^2 + \delta / N.$$

Jensen's inequality on the weights (w_n) then implies

$$1/N^2 = \left(\sum_n w_n(\mathbf{X})/N \right)^2 \leq \sum_n w_n^2(\mathbf{X})/N$$

so that

$$\left\| \eta^\star_{h,N} - \gamma \right\|_{h_N} \leq \left\| \overline{\eta}_N - \gamma \right\|_h .$$

To compute the expected value of $\left\| \overline{\eta}_N - \gamma \right\|_h^2$, we use the spectral decomposition (26) to write

$$\mathbb{E}\left[\left\| \overline{\eta}_N - \gamma \right\|_h^2 \right] = \frac{1}{N^2} \sum_{n,m}^N \mathbb{E}\left[\sum_{k=1}^\infty e^{-2kt} h_k(X_n) h_k(X_m) \right],$$

where t satisfies $h = 1-e^{-2t}$. Since $X_1, ..., X_N$ are i.i.d. $\mathcal{N}(0,1)$ and the $(h_k)_{k\geq 1}$ are all orthogonal to $h_0 = 1$ in $L^2(\gamma)$,

$$\begin{aligned}\mathbb{E}\left[\|\overline{\eta}_N - \gamma\|_h^2\right] &= \frac{1}{N^2}\sum_{k=1}^{\infty} e^{-2kt}\Big\{\sum_{n=m}\mathbb{E}\left[(h_k(X_n))^2\right] + \sum_{n\neq m}\mathbb{E}\left[h_k(X_n)\right]\mathbb{E}\left[h_k(X_m)\right]\Big\}\\ &= \frac{1}{N}\left(\sum_{k=1}^{\infty} e^{-2kt}\right) = \frac{1}{N}\left(\frac{e^{-2t}}{1-e^{-2t}}\right)\\ &= \frac{1}{N}\left(\frac{1}{h}-1\right),\end{aligned}$$

concluding the proof of Equation (13). □

We end this section by proving the weak convergence result of Corollary 1.10. The proof uses the following classical result (see e.g. [Kal02, Lem. 3.2]), which implies in particular that convergence in probability is a topological notion that does not depend on the choice of a metric.

Lemma 3.8 (Subsequence criterion)**.** *Let $Y_1, Y_2, ...$ be random elements in a metric space (S,d). Then $Y_n \xrightarrow{\mathbb{P}} Y$ iff for all sub-sequence $(k_n) \subset \mathbb{N}$, there exists a further subsequence $(l_{k_n}) \subset (k_n)$ such that $Y_n \to Y$ a.s. along (l_{k_n}).*

Proof of Corollary 1.10. We let $N \to +\infty$ with $\frac{1}{N} \ll h_N \leq h_0 < 1$, and show that in probability, the random measure $\eta^\star_{h,N}$ converges weakly to γ.

Let $h_N = 1 - e^{-2t}$. By Lem. 3.8, it is enough to prove that from any subsequence of $\eta^*_{h_N,N}$, we can extract a further subsequence along which $\eta^*_{h_N,N}$ converges in distribution to γ. Let us consider an arbitrary subsequence of $\eta^*_{h_N,N}$. By Equation (13),

$$\mathbb{E}\left[\left\|\eta^\star_{h_N,N} - \gamma\right\|^2_{h_N}\right] \leq \frac{1}{N}\left(\frac{1}{h_N}-1\right) \xrightarrow[N\to\infty]{} 0,$$

so the random variable $\left\|\eta^\star_{h_N,N} - \gamma\right\|_{h_N}$ converges in $L^2(\mathbb{P})$ to 0. Convergence in L^2 implies convergence in probability so that by Lemma 3.8, there is a further subsequence along which

$$\left\|\eta^\star_{h_N,N} - \gamma\right\|_{h_N} \xrightarrow[N\to\infty]{a.s.} 0.$$

Since $h_N \leq h_0$, we get by Theorem 3.4, item (4), that along the sub-subsquence,

$$\left\|\eta^\star_{h_N,N} - \gamma\right\|_{h_0} \xrightarrow{a.s.} 0.$$

Since $\|\cdot\|_{h_0}$-convergencence implies weak convergence by item (5) of Theorem 3.4, $\eta^\star_{h,N} \xrightarrow[(d)]{a.s.} \gamma$ along the sub-subsequence. □

3.6. **Fast convergence of the weighted measure and a conjecture.** In this Section we prove the second part of Theorem 1.9: in the $\delta = 0$ case, for h sufficiently large,

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|^2_h\right] = o(1/N).$$

3.6.1. *Strategy of proof.* Recall that $\eta^\star_{h,N}$ is defined by minimizing $\|\sum w_n\delta_{X_n} - \gamma\|^2_h$ over all weight vectors. The main difficulty here is that this minimizer is not explicit. However, for any integer K, the spectral decomposition giving (26) can be used to split the cost function in two terms:

$$\|\eta - \gamma\|_h^2 = \sum_{k=1}^{K} e^{-2kt}\left(\sum_n w_n h_k(X_n)\right)^2 + \sum_{k>K} e^{-2kt}\left(\sum_n w_n h_k(X_n)\right)^2. \tag{28}$$

Let $w_n^K(\mathbf{X})$ be an optimizer of the *first*, finite dimensional term. If N is large enough with respect to K, then it is reasonable to expect that, with high probability, the value of this finite dimensional problem is zero.

Definition 3.9 (K-good vectors)**.** *A vector* $\mathbf{x} = (x_1, ..., x_N)$ *is said to be* K-good *if there exists a weight vector* $(w_1, ..., w_N)$ *in the simplex* Ω *such that*

$$\forall 1 \leq k \leq K, \quad \sum_n w_n h_k(x_n) = 0.$$

Let G be the "good event" $G = \{\mathbf{X} \text{ is } K\text{-good}\}$. On G we compare $w_n(\mathbf{X})$ to $w_n^K(\mathbf{X})$; on the bad event we simply use the naïve empirical measure:

$$\begin{aligned}\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] &\leq \mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2 \mathbf{1}_G\right] + \mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2 \mathbf{1}_{G^c}\right] \\ &\leq \mathbb{E}\left[\left\|\sum_n w_n^K(\mathbf{X})\delta_{X_n} - \gamma\right\|_h^2 \mathbf{1}_G\right] + \mathbb{E}\left[\left\|\overline{\eta}_N - \gamma\right\|_h^2 \mathbf{1}_{G^c}\right]\end{aligned}$$

For the first term, on the good event G, we apply (28): by definition the first term vanishes and we get

$$\begin{aligned}\left\|\sum_n w_n^K(\mathbf{X})\delta_{X_n} - \gamma\right\|_h^2 \mathbf{1}_G &\leq \sum_{k>K} e^{-2tk}\left(\sum_n w_n^K(\mathbf{X})h_k(X_n)\right)^2 \\ &\leq \sum_{k>K}\sum_n e^{-2tk} w_n^K(\mathbf{X}) h_k^2(X_n).\end{aligned}$$

where we used Jensen's inequality with the weights w_n^K in the last line. We now take the expectation, bounding w_n^K by one, to get

$$\begin{aligned}\mathbb{E}\left[\left\|\sum_n w_n^g(\mathbf{X})\delta_{X_n} - \gamma\right\|_h^2 \mathbf{1}_G\right] &\leq \sum_{k>K}\sum_n e^{-2tk}\mathbb{E}\left[h_k^2(X_n)\right] \\ &\leq N\sum_{k>K} e^{-2tk} \\ &\leq N\frac{e^{-2t(K+1)}}{1-e^{-2t}}.\end{aligned}$$

On the bad event we use Hölder's inequality:

$$\mathbb{E}\left[\left\|\overline{\eta}_N - \gamma\right\|_{h_N}^2 \mathbf{1}_{G^c}\right] \leq \mathbb{E}\left[\left\|\overline{\eta}_N - \gamma\right\|_{h_N}^4\right]^{1/2} \mathbb{P}\left[G^c\right]^{1/2}.$$

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] \leq N\frac{e^{-2t(K+1)}}{1-e^{-2t}} + \mathbb{E}\left[\left\|\overline{\eta}_N - \gamma\right\|_{h_N}^4\right]^{1/2} \mathbb{P}\left[G^c\right]^{1/2}. \tag{29}$$

In order to bound the 4th moment of the h-norm, we proceed as follows. Recall that $h = 1-e^{-2t}$, and suppose that $e^{2t} \geq 3$, so that we can write $t = s + u$ with s satisfying $e^{2s} = 3$. Then

$$\begin{aligned}\left\|\overline{\eta}_N - \gamma\right\|_h^4 &= \left\|\frac{d\overline{\eta}_N P_t}{d\gamma} - 1\right\|_2^2 = \left\|P_s\left(\frac{d\overline{\eta}_N P_u}{d\gamma} - 1\right)\right\|_2^4 \\ &\leq \left\|P_s\left(\frac{d\overline{\eta}_N P_u}{d\gamma} - 1\right)\right\|_4^2 \\ &\leq \left\|\left(\frac{d\overline{\eta}_N P_u}{d\gamma} - 1\right)\right\|_2^2\end{aligned}$$

where the last line uses Nelson's theorem, that is, the hypercontractivity of the Ornstein-Uhlenbeck semigroup (see *e.g.* [Gro93]) which here holds true between L^4 and L^2 for time greater than $s = (\ln 3)/2$.

Taking expectations and reusing (13), we get

$$\mathbb{E}\left[\left\|\overline{\eta}_N - \gamma\right\|_h^4\right] \leq \frac{1}{N}\left(\frac{1}{1-e^{-2u}} - 1\right) = \frac{1}{N}\left(\frac{1}{1-3e^{-2t}} - 1\right).$$

Putting everything together, we have for $t \geq (\ln 3)/2$:

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] \leq N\frac{e^{-2t(K+1)}}{1-e^{-2t}} + \frac{1}{\sqrt{N}}\left(\frac{1}{1-3e^{-2t}} - 1\right)^{1/2}\mathbb{P}\left[G^c\right]^{1/2}. \tag{30}$$

To go forward, the main challenge is to get a bound on the probability of the bad event.

3.6.2. *Control on the bad event by coupon collecting.* Let M be an integer and decompose the real line $\mathbb{R}$ in M segments between the quantiles $(z_i)_{0\leq i\leq M}$, where $F_\gamma(z_i) = \int_{-\infty}^{z_i}\gamma(dx) = i/M$.

Definition 3.10 (Well spread vector)**.** *A vector $\mathbf{x} = (x_1, ..., x_N)$ is said to be* M-well-spread *if it visits each of the M quantiles of the Gaussian:*

$$\forall 1 \leq j \leq M, \exists 1 \leq i \leq N \quad x_i \in (z_{i-1}, z_i).$$

The main results of this section are the two following lemmas.

Lemma 3.11 (M-well-spread implies K-good)**.** *There exists a universal constant C such that, if $\mathbf{x}$ is M-well-spread, then it is K-good for all K such that*

$$M > CK^{5/2}8^K.$$

Lemma 3.12 (Large samples are well-spread)**.** *Suppose that $N > (2p+2)M\ln(M)$. For $\mathbf{X} = (X_1, ..., X_N)$, an iid gaussian sample, the probability that $\mathbf{X}$ is not M-well-spread is small:*

$$\mathbb{P}\left[\exists i, \forall n, X_n \notin (z_{i-1}, z_i)\right] \leq \frac{M}{M-1}\frac{1}{M^{2p+1}}.$$

We start by the short proof of this second lemma.

Proof of Lemma 3.12. We interpret the question as a coupon collecting problem. For M coupons, the number of trials T needed to get a complete collection admits the following classical deviation bound, see for example [MR95, Section 3.6.1, p. 58]:

$$\forall l \in \mathbb{N}, \quad \mathbb{P}\left[T > l\right] \leq M(1-1/M)^l \leq M\exp(-l/M),$$

obtained by expressing $\{T > l\}$ as the union of the M events "the k^{th} coupon never appears in the l trials". Thus

$$\forall t, \quad \mathbb{P}\left[T > M\ln(M) + Mt\right] \leq \frac{M}{M-1}\exp(-t),$$

where the $M/(M-1)$ factor comes from the fact that $M\ln(M) + Mt$ might not be an integer. We choose $t = (2p+1)\ln(M)$, and recall that by assumption $(2p+2)M\ln(M) < N$. This yields a bound on the probability of not being well-spread:

$$\begin{aligned}\mathbb{P}\left[\exists i, \forall n, X_n \notin (x_{i-1}, x_i)\right] = \mathbb{P}\left[T > N\right] &\leq \mathbb{P}\left[T > (2p+2)M\ln(M)\right] \\ &\leq \frac{M}{M-1}\frac{1}{M^{2p+1}}.\end{aligned}$$

□

The proof of Lemma 3.11 is a bit more involved. Let us first state and prove three additional lemmas.

Lemma 3.13. $\mathbf{x}$ *is K-bad if and only if there exists a polynomial P such that* $\deg(P) \leq K$*, P is orthogonal to h_0, and*

$$\forall 1 \leq n \leq N, \quad P(x_i) > 0.$$

Proof. By definition, $\mathbf{x}$ is K-bad if and only if the origin of $\mathbb{R}^K$ is not in the convex hull of the N points $(h_1(X_n), ..., h_K(X_n))$. If this is the case, then by the hyperplane separation theorem there exists an $\alpha = (\alpha_1, ..., \alpha_K)$ that has a positive scalar products with the N points, that is,

$$\forall 1 \leq n \leq N, \quad \sum_{k=1}^K \alpha_k h_k(X_n) > 0.$$

In other words, the polynomial $P = \sum_{k=1}^K h_k$ takes positive values on each of the X_n for $1 \leq n \leq N$. Since the (h_k) are orthogonal, P is indeed orthogonal to h_0.

Conversely if such a $P = \sum_{k=1}^K \alpha_k$ exists then $\alpha = (\alpha_1, ..., \alpha_K)$ has a (strictly) positive scalar product with the N points $(h_1(X_n), ..., h_K(X_n))_{1\leq n\leq N}$, so it has a positive scalar product with any convex combination of these points, therefore 0 cannot be in the convex hull of these points. □

Lemma 3.14. *There is a universal constant C such that, if $P = \sum_{k=1}^{K} a_k h_k$ is a polynomial of degree at most K orthogonal to h_0, then*

$$\mathbb{E}\left[|P(Z)|^3\right]^{1/3} \leq CK^{1/4}2^{K/2}\mathbb{E}\left[P(Z)^2\right]^{1/2}.$$

Proof. Without loss of generality we may assume $\sum_{k=1}^{K} a_k^2 = 1$.

$$\begin{aligned}\mathbb{E}\left[|P(Z)|^3\right]^{1/3} &\leq \sum |a_k|\,\mathbb{E}\left[|h_k(Z)|^3\right]^{1/3}\\ &\leq CK^{-1/4}2^{K/2}\sum |a_k|\\ &\leq CK^{1/4}2^{K/2},\end{aligned}$$

where the second line follows from Theorem 2.1, eq. (2.2) in [LC02], remarking that our h_k are normalized in L^2 instead of monic, and the last line from the bound $\sum |a_k| \leq \sqrt{K}(\sum_k |a_k|^2)^{1/2}$. □

Lemma 3.15. *If $X \in L^3$ satisfies $\mathbb{E}[X] = 0$, then*

$$\mathbb{P}[X > 0] \geq \frac{\mathbb{E}\left[X^2\right]^3}{4\mathbb{E}\left[X^3\right]^2}.$$

Proof. Since $\mathbb{E}[X] = 0$, $\mathbb{E}[X_+] = \mathbb{E}[X_-] = \frac{1}{2}\mathbb{E}[|X|]$. Therefore by Hölder's inequality,

$$\frac{1}{4}\mathbb{E}[|X|]^2 = \mathbb{E}[X+]^2 = \mathbb{E}[X\mathbf{1}_{X>0}]^2 \leq \mathbb{E}\left[X^2\right]\mathbb{P}[X > 0].$$

Moreover, another application of Hölder's inequality yields

$$\mathbb{E}\left[X^2\right] \leq \mathbb{E}[|X|]^{1/2}\,\mathbb{E}\left[|X|^3\right]^{1/2}.$$

Putting these two inequalities together, we get

$$\mathbb{P}[X > 0] \geq \frac{\mathbb{E}[|X|]^2}{4\mathbb{E}\left[X^2\right]} \geq \frac{\mathbb{E}\left[X^2\right]^3}{4\mathbb{E}\left[|X|^3\right]^2}.$$ □

Proof of Lemma 3.11. Suppose that $\mathbf{x}$ is M-well-spread but K-bad. By Lemma 3.13, there exists a $P_{\mathbf{x}}$ of degree at most K that takes positive values on each of the x_i. This $P_{\mathbf{x}}$ has $L \leq K$ real roots $r_1 \leq \cdots \leq r_L$. To fix ideas, suppose that $P_{\mathbf{x}}$ is negative at $-\infty$. Setting $r_0 = -\infty$ and $r_{L+1} = \infty$, the open set $\{z : P_{\mathbf{x}}(z) < 0\}$ may therefore be written as the union of disjoint, possibly empty, intervals $\bigcup_{m \text{ even}, m\leq L}(r_m, r_{m+1})$. These intervals cannot contain the x_i, so each one is included in the union of two adjacent interquantiles intervals, so that for m even, $m \leq L$,

$$\int_{r_m}^{r_{m+1}} \gamma(dz) \leq \frac{2}{M}.$$

Furhtermore, the number of intervals is at most $\lceil (K+1)/2 \rceil \leq (K+3)/2$. Rewriting the gaussian integral as a probability, we get, for Z a standard Gaussian random variable,

$$\mathbb{P}[P_{\mathbf{x}}(Z) < 0] \leq (K+3)/M.$$

The key point now is that $P_{\mathbf{x}}$ is orthogonal to $h_0 = 1$, that is, in probabilistic terms, $\mathbb{E}[P_{\mathbf{x}}(Z)] = 0$, so that we can use the concentration lemma 3.15, to bound the left hand side from below and get:

$$\frac{K+3}{M} \geq \frac{\mathbb{E}\left[P(Z)^2\right]^3}{4\mathbb{E}\left[|P(Z)^3|\right]^2} \geq \frac{c}{K^{3/2}8^K}.$$

This bound implies the claim. □

3.6.3. *End of the proof of Theorem 1.9.* Let us recall the bound (30) for $t \geq (\ln 3)/2$:

$$\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] \leq N\frac{e^{-2t(K+1)}}{1-e^{-2t}} + \frac{1}{\sqrt{N}}\left(\frac{1}{1-3e^{-2t}} - 1\right)^{1/2}\mathbb{P}\left[G^c\right]^{1/2}. \tag{31}$$

For each N choose M and K the *largest* possible integers such that

$$N > 4M\ln(M), \qquad\qquad M > CK^{5/2}8^K. \tag{32}$$

Note that in particular $N = \mathcal{O}^\star(M)$. This relation between M and K ensures that, by Lemma 3.11, the sample $\mathbf{X}$ is K-good as soon as it is M-well-spread, so that by Lemma 3.12,

$$\mathbb{P}\left[G^c\right] = O(1/M^3) = O^\star(1/N^3),$$

and the last term in (31) is $o(1/N)$. For the first term of (31), the bound is in Ne^{-2tK}. The definitions of K and M ensure that $N = O((8+\varepsilon)^K)$, so that $N^2e^{-2tK} = o(1)$ if t is large enough, or in other words $Ne^{-2tK} = o(1/N)$. Putting everything together, we get $\mathbb{E}\left[\left\|\eta^\star_{h,N} - \gamma\right\|_h^2\right] = o(1/N)$, concluding the proof of Equation (14) and of Theorem 1.9.

3.6.4. *Proof of Corollary 1.13.* Let us denote by $w_n^\delta(\mathbf{X})$ and by $\eta_N^{\delta,\star}$ the optimal weights and the associated weighted empirical distribution obtained by the optimization problem (8) for a given δ (we drop the subscript h in notation for simplicity). By construction,

$$\begin{aligned}\mathbb{E}\left[\mathrm{dist}\left(\eta_N^{0,\star},\gamma\right)^2\right] + \delta\sum_n w_n^0(\mathbf{X})^2 &\geq \mathbb{E}\left[\mathrm{dist}\left(\eta_N^{\delta,\star},\gamma\right)^2\right] + \delta\sum_n w_n^\delta(\mathbf{X})^2\\ &\geq \mathbb{E}\left[\mathrm{dist}\left(\eta_N^{0,\star},\gamma\right)^2\right] + \delta\sum_n w_n^\delta(\mathbf{X})^2\end{aligned}$$

so that $\sum_n w_n^\delta(\mathbf{X})^2 \leq \sum_n w_n^0(\mathbf{X})^2$. As a consequence, the MSE obtained with a given δ can be first bounded using (7) and then using the wieghts $w_n^0(\mathbf{X})$ so that

$$\mathbb{E}\left[\left(\mathbf{\Phi}_W^\delta(\mathbf{X},\mathbf{Y}) - \mathbb{E}\left[\phi(X,Y)\right]\right)^2\right] \leq (v_\phi - \delta m_\phi^2)\mathbb{E}\left[\sum_n w_n^0(\mathbf{X})^2\right] + m_\phi^2\mathbb{E}\left[\mathrm{dist}\left(\eta_N^{0,\star},\gamma\right)^2 + \delta\sum_n w_n^0(\mathbf{X})^2\right].$$

The corallary then follows from Theorem 1.9 and Conjecture 1.11.

4. The Wasserstein method

4.1. **An exact expression for the optimal weights.** The fact that the Wasserstein method is both easier to analyze and faster in practice stems from the fact that the minimization problem can be solved explicitely.

Proposition 4.1. *Let $\mathbf{x} = (x_1, ..., x_N)$ be a set of N distinct points in $\mathbb{R}$, let $(x_{(1)} < x_{(2)} \cdots < x_{(N)}$ be their ordered relabelling, and let $(y_n)_{0\leq n\leq N}$ be the middle points $(1/2)(x_{(n)} + x_{(n+1)})$, with the convention $y_0 = -\infty$ and $y_N = \infty$.*

For $w = (w_1, ..., w_N)$ in the simplex $\Omega = \{(w_1, ..., w_N) \in \mathbb{R}_+^N, \sum_n w_n = 1\}$, let $F(w)$ be the cost

$$F(w) = \mathcal{W}_1\left(\sum_{n=1}^N w_n\delta_{x_n}, \gamma\right).$$

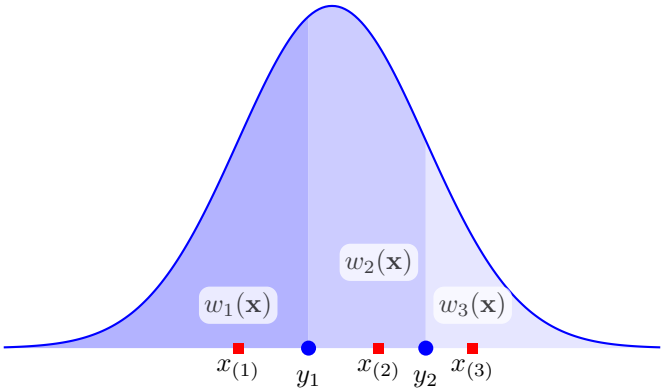
The optimization problem

$$\textit{minimize: } F(w) \quad \textit{subject to: } w \in \Omega$$

has a unique solution $w(\mathbf{x}) = (w_1(\mathbf{x}), ..., w_N(\mathbf{x}))$, given by

$$w_n(\mathbf{x}) = \int_{y_m - 1}^{y_m} \gamma(dz),$$

where m is the unique integer such that $x_n = x_{(m)}$.



Given the sample $(\mathbf{x})$, the optimal Wasserstein weights are obtained by computing the middle points $y_n = (x_{(n)} + x_{(n+1)})/2$, and letting $w_n = \gamma([y_{n-1}, y_n])$.

FIGURE 3. The optimal weights $w(\mathbf{x})$.

Proof. First, note that thanks to the relabelling in the last part of the statement, is enough to prove the result when the (x_n) are already ordered; we assume from now on that $x_1 < \cdots < x_N$.

Let η be any probability measure on $\mathbb{R}$, and recall that γ is the standard Gaussian measure; denote by F_η, F_γ their respective cumulative distribution functions. The Wasserstein distance $\mathcal{W}_1$ between η and γ admits the following classical representation, see for example [Vil03, Remark 2.19 item (iii)] :

$$\mathcal{W}_1(\gamma, \eta) = \int_{\mathbb{R}} |F_\eta(x) - F_\gamma(x)|\, dx.$$

Consider now the discrete measure $\eta(w) = \sum_{n=1}^N w_n \delta_{x_n}$. By cutting the integral at the points x_n and isolating the first and last terms, we get the explicit formula

$$\begin{aligned} F(w) &= \mathcal{W}_1\left(\sum_{n=1}^N w_n \delta_{x_n}, \gamma\right) \\ &= \sum_{n=1}^{N-1} \int_{x_n}^{x_{n+1}} \left|\sum_{m=1}^n w_m - F_\gamma(z)\right| dz + \int_{-\infty}^{x_1} |F_\gamma(z)|\, dz + \int_{x_N}^{\infty} |1 - F_\gamma(z)|\, dz. \end{aligned} \tag{33}$$

Note that the extremal terms do not depend on the weight vector w. For $1 \le n \le N-1$, consider now the n^{th} term in this sum, and write it as $\phi_n(\sum_{m=1}^n w_m)$, where

$$\phi_n(c) = \int_{x_n}^{x_{n+1}} |c - F_\gamma(z)|\, dz.$$

Writing $\phi_n(c) = (x_{n+1} - x_n)\mathbb{E}\left[|c - F_\gamma(U)|\right]$ for U a uniform variable on $[x_n, x_{n+1}]$, we see by classical properties of medians, see e.g. [Str11, p. 43], that ϕ_n attains its minimal value at the unique median of the distribution of $F_\gamma(U)$, that is, at the point p where $\mathbb{P}\left[F_\gamma(U) \le p\right] = 1/2$. Since

$$\mathbb{P}\left[F_\gamma(U) \le p\right] = \mathbb{P}\left[U \le F_\gamma^{-1}(p)\right] = (F_\gamma^{-1}(p) - x_n)/(x_{n+1} - x_n),$$

the minimum of ϕ_n is attained at the unique point $F_\gamma(y_n)$, where we recall that y_n is the midpoint $(x_n + x_{n+1})/2$.

To conclude the proof, it is now enough to remark that letting $w_n = \int_{y_{n-1}}^{y_n} \gamma(dz)$, we get $\sum_{m=1}^n w_m = F_\gamma(y_n)$, so that $(w_1, ..., w_N)$ minimizes all the terms in the sum (33). □

4.2. **Probabilistic properties of the optimal weights.** Let X_1, ..., X_n be *i.i.d.* $\mathcal{N}(0,1)$. In this section we investigate the behaviour of the $\mathcal{W}_1$ distance $D = D(\mathbf{X}) = \mathcal{W}_1(\sum_n w_n(\mathbf{X})\delta_{X_n}, \gamma)$ between the optimally reweighted sample and the target Gaussian measure. We start by proving the first part of Theorem 1.15: for any integer p,

$$\mathbb{E}\left[D^p\right] = \mathcal{O}^\star\left(\frac{1}{N^p}\right)$$

where $\mathcal{O}^\star$ means $\mathcal{O}$ up to logarithmic correction terms.

Proof of Theorem 1.15, first part. Let us first note that, since γ is absolutely continuous with respect to the Lebesgue measure, classical results on optimal transportation in dimension 1 for the usual distance (see for example [Vil03, Theorem 2.18] and the remarks that follow it) imply that the Monge-Kantorovitch problem (12) defining $\mathcal{W}_1(\eta,\gamma)$ distance has an explicit minimizer, given by the deterministic coupling $(T(Z),Z)$, where $Z\sim\gamma$ and T is the monotone transport map

$$T(z)=F_\eta^{-1}(F_\gamma(z)).$$

Therefore, the optimal coupling between a Gaussian random variable X and the optimally reweighted empirical measure $\sum_n w_n(\mathbf{x})\delta_{x_n}$ is given by the piecewise constant transport map that sends each interval $]y_n,y_{n+1}[$ to x_n, so D has the explicit expression

$$D=\int \min_n |x-X_n|\,\gamma(dx).$$

We start by a rough bound: for any $\lambda>0$, the Laplace transform $\exp\exp(\lambda D)$ may be bounded as follows using Jensen's inequality:

$$\begin{aligned}\mathbb{E}\left[\exp(\lambda D)\right]&=\mathbb{E}\left[\exp\left(\lambda\int\min_n|x-X_n|\,\gamma(dx)\right)\right]\\&\le\mathbb{E}\left[\exp\left(\lambda\min_n|X-X_n|\right)\right],\end{aligned}$$

where $X\sim\gamma$ is independent of $\mathbf{X}=(X_1,...,X_n)$. Then

$$\begin{aligned}\mathbb{E}\left[\exp(\lambda D)\right]&\le\mathbb{E}\left[\exp\left(\lambda|X-X_1|\right)\right]\\&\le\mathbb{E}\left[\exp\left(\lambda|X|\right)\right]^2.\end{aligned}$$

Since the last expression is finite, we have established

$$\forall\lambda,\exists C_\lambda,\forall N,\quad \mathbb{E}\left[\exp\left(\lambda D\right)\right]\le C_\lambda. \tag{34}$$

We now let $M<N$ be an integer and decompose the real line $\mathbb{R}$ in M segments between the quantiles $(x_i)_{0\le i\le M}$, where $F_\gamma(x_i)=\int_{-\infty}^{x_i}\gamma(dx)=i/M$. We let G be the "M-well-spread event" (Definition 3.10) that there is at least one of the $(X_n)_{1\le n\le N}$ in each of the M "bins" $(]x_{i-1},x_i[)_{1\le i\le M}$. We then proceed in three steps.

Step 1: D is small on the well-spread event. Indeed, on G, there exist $N(1),...,N(M)$ such that $X_{N(i)}\in]x_{i-1},x_i[$. Therefore

$$\begin{aligned}D\mathbf{1}_G&=\mathbf{1}_G\int\min_n|x-X_n|\,\gamma(dx)\\&=\mathbf{1}_G\sum_{i=1}^M\int_{x_{i-1}}^{x_i}\min_n|x-X_n|\,\gamma(dx)\\&\le\mathbf{1}_G\sum_{i=1}^M\int_{x_{i-1}}^{x_i}\left|x-X_{N(i)}\right|\gamma(dx)\\&\le\sum_{i=2}^{M-1}|x_i-x_{i-1}|\int_{x_{i-1}}^{x_i}\gamma(dx)+2\int_{x_{M-1}}^{\infty}|x-x_{M-1}|\,\gamma(dx)\\&\le\frac{2}{M}x_{M-1}+2\int_{x_{M-1}}^{\infty}|x-x_{M-1}|\,\gamma(dx)\\&\le\frac{2}{M}x_{M-1}+2\int_{x_{M-1}}^{\infty}x\gamma(dx)\\&\le\frac{2}{M}x_{M-1}+\frac{2}{\sqrt{2\pi}}\exp(-x_{M-1}^2/2).\end{aligned}$$

From the classical gaussian tail estimate

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{t}-\frac{1}{t^3}\right)\exp(-t^2/2)\le 1-F_\gamma(t)\le\frac{1}{\sqrt{2\pi}}\left(\frac{1}{t}\right), \tag{35}$$

applied to $t = x_{M-1}$, it is easily seen by taking logarithms that $x_{M-1} \sim \sqrt{2\log(M)}$. Using the first inequality in (35) again, we get $\exp(-x_{M-1}^2/2) = \mathcal{O}^\star(1/M)$, and finally

$$D\mathbf{1}_G = \mathcal{O}^\star\left(\frac{1}{M}\right).$$

Step 2: the well-spread event is very likely. Assuming from now on that M satisfies $N > (2p+2)M\ln(M)$, we get thanks to Lemma 3.12 that

$$\mathbb{P}\left[G^c\right] = \mathcal{O}(1/M^{2p+1}).$$

Step 3: conclusion. We decompose $\mathbb{E}\left[D^p\right]$ in two parts, depending on whether the sample X is well-spread or not. On G we use the result from Step 1; on G^c we apply Hölder's inequality, the bound on $\mathbb{P}\left[G^c\right]$ from step 2, and the *a priori* control on $\mathbb{E}\left[D^{2p}\right]$ given by the preliminary bound (34):

$$\begin{aligned}
\mathbb{E}\left[D^p\right] &= \mathbb{E}\left[D^p\mathbf{1}_G\right] + \mathbb{E}\left[D^p\mathbf{1}_{G^c}\right] \\
&\leq \mathcal{O}^\star\left(\frac{1}{M^p}\right) + \sqrt{\mathbb{E}\left[D^{2p}\right]}\sqrt{\mathbb{P}\left[G^c\right]} \\
&\leq \mathcal{O}^\star\left(\frac{1}{M^p}\right) + \mathcal{O}\left(\frac{1}{M^{p+1/2}}\right) \\
&\leq \mathcal{O}^\star\left(\frac{1}{M^p}\right).
\end{aligned}$$

Since M may be chosen large enough to guarantee $N = \mathcal{O}^\star(M)$, this implies $\mathbb{E}\left[D^p\right] = \mathcal{O}^\star\left(\frac{1}{N^p}\right)$. □

Proof of Theorem 1.15, second part. We now turn to the proof of the control in l^2 of the optimal weights, and show that

$$\mathbb{E}\left[\sum_{n=1}^{N} w_n(\mathbf{X})^2\right] \leq \frac{6}{N}.$$

By definition,

$$w_n(\mathbf{X}) = F_\gamma(Y_{n+1}) - F_\gamma(Y_n),$$

where the Y_n are the middle points of the reordered sample and F_γ is the cdf of the standard Gaussian distribution. By a rough upper bound, for $2 \leq n \leq N-1$,

$$w_n \leq F_\gamma(X_{(n+1)}) - F_\gamma(X_{(n-1)}).$$

The cdf F_γ maps the ordered sample $(X_{(1)}, ..., X_{(N)})$ to an ordered sample $(U_{(1)}, ..., U_{(N)})$ of the uniform distribution on $[0,1]$, so

$$w_n \leq U_{(n+1)} - U_{(n-1)},$$

for $2 \leq n \leq N-1$, $w_1 \leq U_{(2)}$ and $w_n \leq 1 - U_{(N-1)}$.

Let us upper bound $\mathbb{E}\left[w_n^2\right]$ for $2 \leq n \leq N-1$, using known results on order statistics for uniform variables that may be found e.g. in [Das11, Chapter 6, Theorem 6.6]. Conditionnally on $U_{(n+1)} = u$, $U_{(n-1)}$ is distributed like the second largest value in a sample of n uniform variables on $[0,u]$, that is, like

$$uU^{1/n}V^{1/(n-1)}$$

where U and V are iid uniform on $[0,1]$. Therefore

$$\begin{aligned}
\mathbb{E}\left[w_n^2\right] &\leq \mathbb{E}\left[U_{(n+1)}^2(1 - U^{1/n}V^{1/(n-1)})^2\right] \\
&= \mathbb{E}\left[U_{(n+1)}^2\right]\left(1 - 2\int u^{1/n}v^{1/(n-1)}dudv + \int u^{2/n}v^{2/(n-1)}dudv)\right) \\
&= \frac{6}{(n+1)(n+2)}\mathbb{E}\left[U_{(n+1)}\right]^2.
\end{aligned}$$

N	Number of samples
M	Number of repetitions
h	Bandwidth

TABLE 1. Notation for the numerical tests

Now $U_{(n+1)}$ follow a Beta$(n+1, N-n)$ distribution, so

$$\begin{aligned}\mathbb{E}\left[U_{(n+1)}^2\right] &= \mathrm{Var}(U_{(n+1)}) + \mathbb{E}\left[U_{(n+1)}\right]^2 \\ &= \frac{(n+1)(N-n)}{(N+1)^2(N+2)} + \frac{(n+1)^2}{(N+1)^2} \\ &= \frac{(n+1)(n+2)}{(N+1)(N+2)},\end{aligned}$$

so that $\mathbb{E}\left[w_n^2\right] \leq \frac{6}{(N+1)(N+2)}$, for all $2 \leq n \leq N-1$. One easily checks that this bound also holds for $n = 1$ and N, and by summing we get

$$\mathbb{E}\left[\sum_{n=1}^{N} w_n^2(\mathbf{X})\right] \leq \frac{6N}{(N+1)(N+2)} \leq \frac{6}{N}. \qquad \square$$

5. Numerical experiments I

In this section we focus on the comparison between the weighted empirical measures $\overline{\eta}_N$, $\eta^\star_{h,N}$ and $\eta^\star_{\mathrm{Wass},N}$.

5.1. **Implementation.** The implementation of the Wasserstein method is straightforward: given $(\mathbf{x})$, we only need to sort it, compute the middle points (y_n) and deduce the weights by applying F_γ.

For the L^2 method, the quadratic programming optimization problem 3 ($\delta \simeq 0$ case) is solved using a standard Scilab library based on the dual iterative method detailed in [GI83].

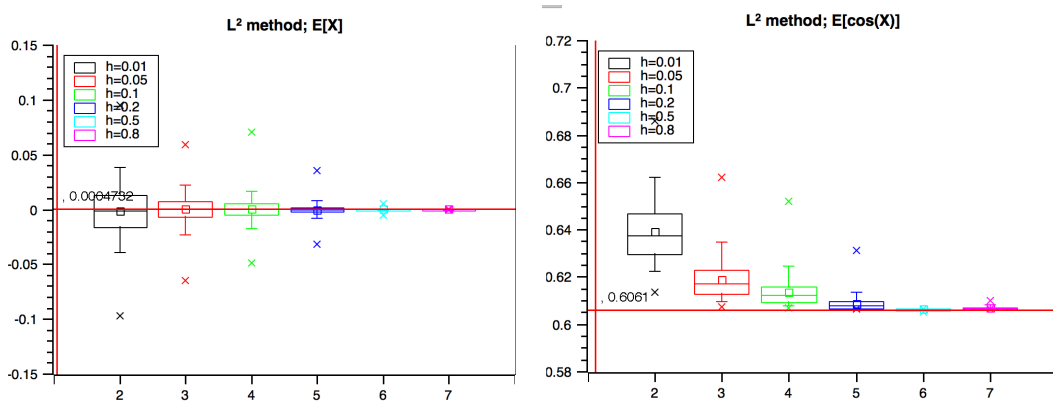
The methods are then tested by computing estimators for the expected value of three functions of X: $\mathbb{E}[X]$, $\mathbb{E}[\cos(X)]$ and $\mathbb{E}[\mathbf{1}_{X>1}]$. The estimators are computed on samples of size N, and the experiment is repeated M times. We present the results as boxplots representing the quantiles on the M repetitions.

5.2. **Regularity of the test function and choice of the bandwidth.** We first investigate the influence of the bandwidth parameter h on the L^2 method, by testing various values of $h \in \{0.01, 0.05, 0.1, 0.2, 0.5, 0.8\}$ on the three test functions ϕ: a) $x \mapsto x$, b) $x \mapsto \cos(x)$ and c) $x \mapsto \mathbf{1}_{|x|>1}$.

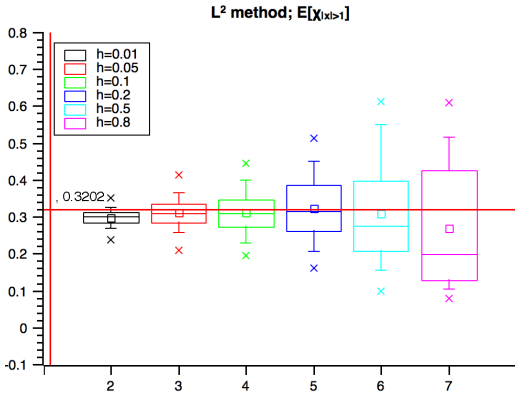
Figure 4(a) corresponds to the test function $x \mapsto x$ which is very specific, the symmetry ensures that the estimator is unbiased, and the method seems to be better the larger h is. Figure 4(b) corresponds to the test function cos; a bias clearly appears in that case, and the estimator is better when h is quite large, with a trade-off at $h = .5$ ($h = .8$ is not as good). In both cases, the fact that the estimators are better when h is quite large may be linked to two remarks made above:

- Remark 3.5 where it is recalled that $x \mapsto x$ and cos are regular test functions that belongs to the image by the Orstein-Uhlenbeck semi-group P_t of an L^2 function on which the optimization is based on;
- Remark 1.14 where it is suggested that the more 'regular' this test function is, the larger the optimal h should be.

However, when we apply the method to estimate the expectation of a discontinuous function of X, here $x \mapsto \mathbf{1}_{|x|>1}$, which does not belong to the appropriate class of regularity, the picture is completely different and the best estimator is obtained for a much smaller $h \approx 0.05$, as can be seen in Figure 4(c).

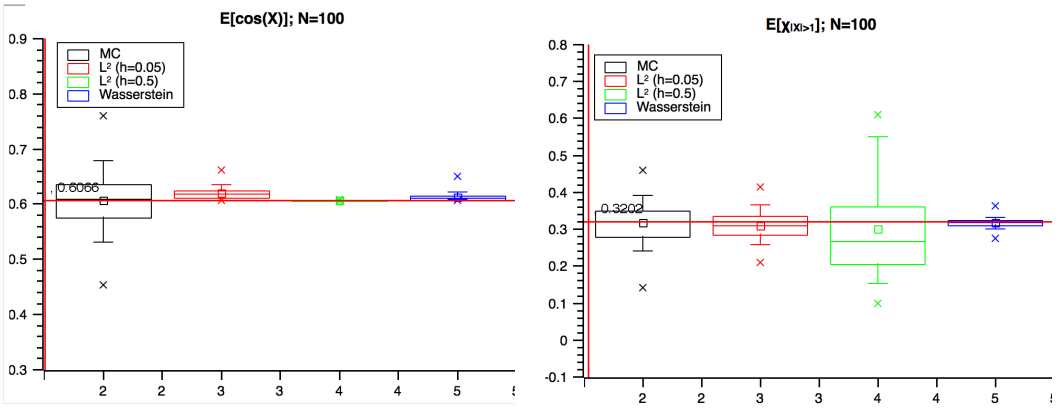


(a) $x \mapsto x$ (b) $x \mapsto \cos x$



(c) $x \mapsto x \mapsto \mathbf{1}_{|x|\geq 1}$

FIGURE 4. $M = 1000$, $N = 100$



(a) $x \mapsto \cos x$ (b) $x \mapsto \mathbf{1}_{|x|\geq 1}$

FIGURE 5. $M = 1000$, $N = 100$

5.3. **Comparison between naïve, L^2 and Wasserstein.** Next, we compare the naïve Monte Carlo method, the L^2 reweighting ($h = .5$ and $h = .05$) and the Wasserstein reweighting.

Figure 5(a) corresponds to the cos test function case, and the naïve Monte Carlo approach is outperformed by all reweighting methods, even with sub-optimal tuning (L^2 for $h = .05$). On the contrary, in Figure 5(b) which corresponds to the step test function the naïve Monte Carlo approach is much better than the L^2 reweighting methods with sub-optimal tuning ($h = .5$), and

similar to the L^2 reweighting methods with quasi-optimal tuning ($h = .05$). This is consistent with the fact that the L^2 reweighting method has been derived for regular test functions, which excludes the step function case.

The Wasserstein reweighting is in both cases (cos and step) much better than the naïve Monte Carlo, and better than the L^2 reweighting in the step function case. For the cos case, the Wasserstein reweighting is similar to the L^2 reweighting method with sub-optimal $h = .05$ and much worse than the L^2 reweighting method with quasi-optimal $h = .5$.

5.4. **Conclusion.** The Wasserstein reweighting is more robust (no parameter to tune) than the L^2 reweighting, and outperforms the latter for irregular test functions. However, for sufficiently regular test functions and with well-chosen bandwidth h, the L^2 reweighting is much better.

6. Numerical experiments II

In this section, we present numerical results exhibiting the variance reduction obtained with the reweighting method.

For simplicity, and having in mind the various drawbacks of the L^2 method in terms of speed and parameter tuning, we will only focus on weights computed with a Wasserstein distance in the minimization problem (3) — that is, the minimization problem with $\delta = 0$.

6.1. **Exchangeable functions of Gaussian vectors.** Let $(G_1, \dots, G_N)$ denotes a sequence of N i.i.d. centered Gaussian vectors in $\mathbb{R}^d$ with identity covariance matrix. We consider the problem of reducing the variance of Monte Carlo estimators of the distribution of $F(G)$ where

$$F : \mathbb{R}^d \to \mathbb{R}$$

is a smooth non-linear function, which is invariant by permutation of the d coordinates (exchangeability). We assume for simplicity the following normalization:

$$F(0) = 0, \qquad D_0F = (1/\sqrt{d}, \dots, 1/\sqrt{d})$$

and set for each $n = 1, \dots, N$:

$$X_n := (D_0F) \cdot G_n \sim \mathcal{N}(0, 1), \qquad Y_n := F(G_n).$$

We are then interested in estimating the cumulant generating function of the distribution of $F(G)$ denoted by

$$k_Y(t) := \log \mathbb{E}\left(\mathrm{e}^{tY}\right) = \log \mathbb{E}\left(\mathrm{e}^{tF(G)}\right),$$

and possibly to compare it to the cumulant generating function of the distribution of the standard Gaussian distribution

$$j_X(t) := \log \mathbb{E}\left(\mathrm{e}^{tX}\right) = \log \mathbb{E}\left(\mathrm{e}^{t(D_0F)\cdot G}\right) = \frac{t^2}{2}.$$

We will consider, compare and combine various estimators. The first two are the naïve and Wasserstein reweighted estimators of k_Y, defined by

$$\mathbf{k}_{\mathrm{MC}}(\mathbf{Y})(t) = \log \frac{1}{N} \sum_{n=1}^{N} \mathrm{e}^{tY_n} \qquad \mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})(t) = \log \sum_{n=1}^{N} w_n(\mathbf{X}) \mathrm{e}^{tY_n},$$

where the weights $w(\mathbf{X})$ are computed with the control variables $\mathbf{X}$ through the minimization problem (3) associated with the (Euclidean-based) Wasserstein distance.

We define similarly two estimators for j_X,

$$\mathbf{j}_{\mathrm{MC}}(\mathbf{Y})(t) = \log \frac{1}{N} \sum_{n=1}^{N} \mathrm{e}^{tX_n} \qquad \mathbf{j}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})(t) = \log \sum_{n=1}^{N} w_n(\mathbf{X}) \mathrm{e}^{tX_n}.$$

Since $j_X(t)$ is explicit, it is quite natural to try and use e^{tX_n} as a control variate, leading to a new estimator:

$$\mathbf{k}_{\mathrm{CV}}(\mathbf{Y})(t) = \mathbf{k}_{\mathrm{MC}}(\mathbf{Y})(t) - \mathbf{j}_{\mathrm{MC}}(\mathbf{X})(t) + j_X(t),$$

Finally, we combine the reweighting and the control variate idea by defining

$$\mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y})(t) = \mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})(t) - \mathbf{j}_{\mathrm{W}}(\mathbf{X})(t) + j_X(t),$$

Note that the control variate has been defined as the best linear approximation of F around the mean $0 \in \mathbb{R}^d$.

We run our tests with the following particular choice of a non-linear function:

$$F_r(g) = \frac{1}{r} \sin\left(\frac{1}{\sqrt{d}} \sum_{i=1}^{d} \sin(r \times g^i) \right)$$

with the parameters $d = 10$ and $r \in \{0.1, 1\}$. Note that r encodes the strength of the nonlinearity, in the sense that

$$\lim_{r \to 0} F_r(g) = g.$$

In all this section, we have taken samples of size $N = 30$. This choice has been made so that the quantiles of the estimators scales appropriately with the target function k_Y to be estimated.

6.1.1. *The almost linear case,* $r = .1$. This case corresponds to a function F which is close to the identity function. In Fig. 6, we have represented the quantiles of the different estimators of $k_Y(t)$ for $t \in [-0.7, 0.7]$.

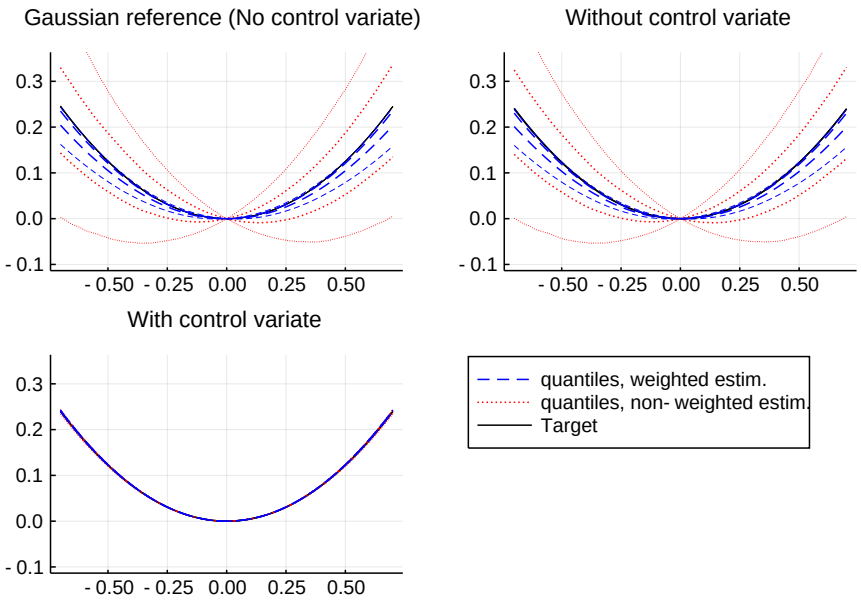


FIGURE 6. **Case** $r = .1$. The figures above represents the $[.05, .25, .75, .95]$-quantiles of the different estimators of the cumulant generating functions for $F(G)$, both without weights (dotted), and with weight (dashed). The figure in the upper left corner represents the estimators of the Gaussian reference $\mathbf{j}_{\mathrm{MC}}(\mathbf{X})$ and $\mathbf{j}_{\mathrm{W}}(\mathbf{X})$. The figure in the upper right corner represents the estimators $\mathbf{k}_{\mathrm{MC}}(\mathbf{Y})$ and $\mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})$ (without control variate). Finally, the figure in the lower left corner represents the estimators of $\mathbf{k}_{\mathrm{CV}}(\mathbf{X}, \mathbf{Y})$ and $\mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y})$ (with control variate).

In Fig. 7, we zoom in on the figure the lower left corner of Fig. 6 where a linear control variate is used, by plotting the difference $\mathbf{k}_{\mathrm{CV}}(\mathbf{X}, \mathbf{Y}) - j_X = \mathbf{k}_{\mathrm{MC}}(\mathbf{Y})(t) - \mathbf{j}_{\mathrm{MC}}(\mathbf{X})(t)$ as well as $\mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y}) - j_X = \mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y}) - \mathbf{j}_{\mathrm{W}}(\mathbf{X})$.

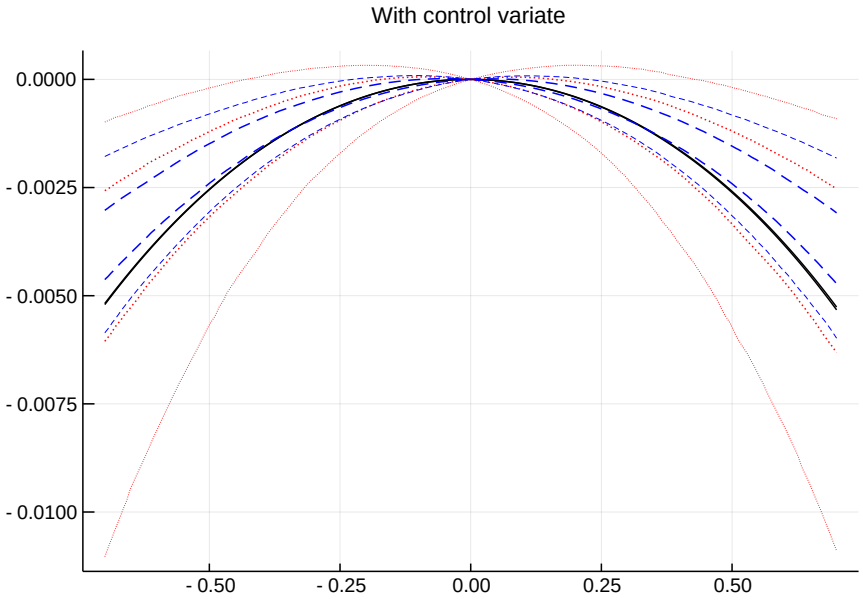


FIGURE 7. **Case** $r = .1$. The figure above represents the $[.05, .25, .75, .95]$-quantiles of $\mathbf{k}_{\mathrm{CV}}(\mathbf{X}, \mathbf{Y}) - j_X$ (dotted) and $\mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y}) - j_X$ (dashed).

6.1.2. *The nonlinear case,* $r = 1$. This case corresponds to a function F with a significant non-linear behavior. In Fig. 8, we have represented the quantile envelopes of the different estimators of $k_Y(t)$ for $t \in [-0.7, 0.7]$.

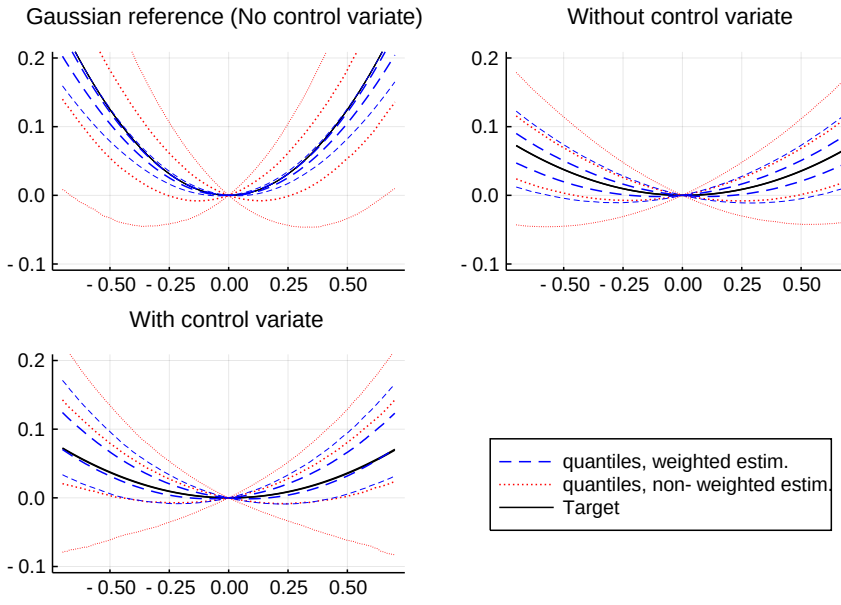


FIGURE 8. **Case** $r = 1$. The figures above represents the $[.05, .25, .75, .95]$-quantiles of the different estimators of the cumulant generating functions for $F(G)$, both without weights (dotted), and with weight (dashed). The figure in the upper left corner represents the estimators of the Gaussian reference $\mathbf{j}_{\mathrm{MC}}(\mathbf{X})$ and $\mathbf{j}_{\mathrm{W}}(\mathbf{X})$. The figure in the upper right corner represents the estimators $\mathbf{k}_{\mathrm{MC}}(\mathbf{Y})$ and $\mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})$ (without control variate). Finally, the figure in the lower left corner represents the estimators of $\mathbf{k}_{\mathrm{CV}}(\mathbf{X}, \mathbf{Y})$ and $\mathbf{k}_{\mathrm{CV+W}}(\mathbf{X}, \mathbf{Y})$ (with control variate).

6.1.3. *Interpretation.* First note that the non-linearity of the function F in the case $r = .1$ has a non-negligible influence on the distribution of $F(G)$, as can be seen in the upper right and lower

left figures of Fig. 8, where the cumulant generating function $k_Y(t)$ (the 'target', represented with a full line) is substantially different from the Gaussian reference $j_X(t)$ (the 'Gaussian reference', represented with a full thin line), and has a much smaller variance.

We first immediately observe that in all cases (the Gaussian reference, the estimator of k_y without control variate, and the estimator of k_y with control variate) the use of the studied weighting method substantially improve the estimation by:

(1) Significantly reducing the spread of the tail distribution (the $\{.1, .9\}$-quantiles) of the estimators.
(2) Significantly reducing the statistical error of the typical outcomes (the $\{.25, .75\}$–quantiles) of the estimators.

Then we can observe that as expected, the error reduction due to the weighting method is slightly better for the Gaussian reference. However, it is clear that the error reduction due to the weighted method is very significant in each case. For instance the typical error (as given by the $\{.25, .75\}$–quantiles) of the estimator $\mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})$ is reduced almost by a factor 2 as compared to $\mathbf{k}_{\mathrm{MC}}(\mathbf{Y})$. As a reference, the typical error on $\mathbf{j}_{\mathrm{W}}(\mathbf{X})$ is reduced by a factor 5 as compared to $\mathbf{j}_{\mathrm{MC}}(\mathbf{X})$.

Finally, it is remarkable to notice that in the case $r = 1$ the control variate method is useless and may even be counterproductive. On the contrary the weighting method behaves well and reduces the error (with or without control variate). It clear from Fig.8 that the weighting method outperforms the control variate method which is not useful here.

This experiment demonstrates that the weighting method can then very easily and very efficiently be used to reduce the statistical error caused by non-linear functions, without resorting to *ad hoc* analytic calculations.

6.2. **A physical toy example.**

6.2.1. *Model.* In this section, we illustrate the use of the weighting method with a more concrete, physical example. We consider a Langevin stochastic differential equation in $\mathbb{R}^d$

$$\begin{cases} dQ_t = P_t dt \\ dP_t = -Q_t dt + \varepsilon \mathcal{F}(Q_t) dt - P_t dt + \sqrt{2} dW_t \end{cases}$$

which is a toy model for a thermostatted linear mechanical system. The latter is perturbed out of equilibrium by an exterior force field $\mathcal{F} : \mathbb{R}^d \to \mathbb{R}^d$, and we are interested in computing the distribution of the long time stationary back reaction, that is to say the distribution of $\mathcal{F}(Q)$ where $Q \in \mathbb{R}^d$ is distributed according to the invariant distribution of the Langevin process, and this for ε small.

For simplicity we assume that $\mathcal{F}_i(Q) = F(Q) \in \mathbb{R}$ is independent of i with again

$$F(0) = 0, \qquad D_0 F = (1/\sqrt{d}, \dots, 1/\sqrt{d}).$$

We set $X_n = F(Q_{\tau n})$ where τ is a sufficiently large decorrelation time. We also set $Y_n = D_0 F \cdot \tilde{Q}_{\tau n}$ where $\tilde{Q}$ is solution to the coupled, non perturbed linear system

$$\begin{cases} d\tilde{Q}_t = \tilde{P}_t dt \\ d\tilde{P}_t = -\tilde{Q}_t dt - \tilde{P}_t dt + \sqrt{2} dW_t \end{cases}$$

so that Y_n is a Gaussian sequence of unit standard Gaussian variables that are approximately independent (for large τ). Using elementary calculations (see *e.g.* [BGM10]), one can check that the positive definite quadratic Lyapunov functional

$$D_t := \left|P_t - \tilde{P}_t\right|^2 + \left|Q_t - \tilde{Q}_t\right|^2 + \left(Q_t - \tilde{Q}_t\right) \cdot \left(P_t - \tilde{P}_t\right)$$

satisfies almost surely the following differential inequality:

$$\frac{d}{dt} D_t \leq -\frac{1}{2} D_t + 4 D_t^{1/2} \varepsilon \left|F(Q_t)\right|.$$

Assuming for simplicity that $\|F\|_\infty = 1$, a Gronwall-type integration yields that for any $t \geq 0$

$$\left|P_t - \tilde{P}_t\right|^2 + \left|Q_t - \tilde{Q}_t\right|^2 \leq c D_t \leq \varepsilon + \mathcal{O}\left(\mathrm{e}^{-t/2}\right),$$

for some numerical constant c. Hence (Q_t, P_t) converges when $\varepsilon \to 0$ to the Orstein-Uhlenbeck process $(\tilde{Q}_t, \tilde{P}_t)$ uniformly in time. This coupling calculation thus suggests that $(Q_{\tau n})_{n\geq 1}$ will be close to a i.i.d. Gaussian sequence when $\varepsilon \to 0$ and $\tau \gg 1$.

6.2.2. *Numerical experiment.* We present in Figure 9 some numerical results in a test case with the same non-linear function:

$$F(g) = \sqrt{d}\sin\left(\frac{1}{d}\sum_{i=1}^{d}\sin(g^i)\right)$$

with the parameters $(\varepsilon = .01, d = 10, N = 50)$. The methodology is the same as in the last section.

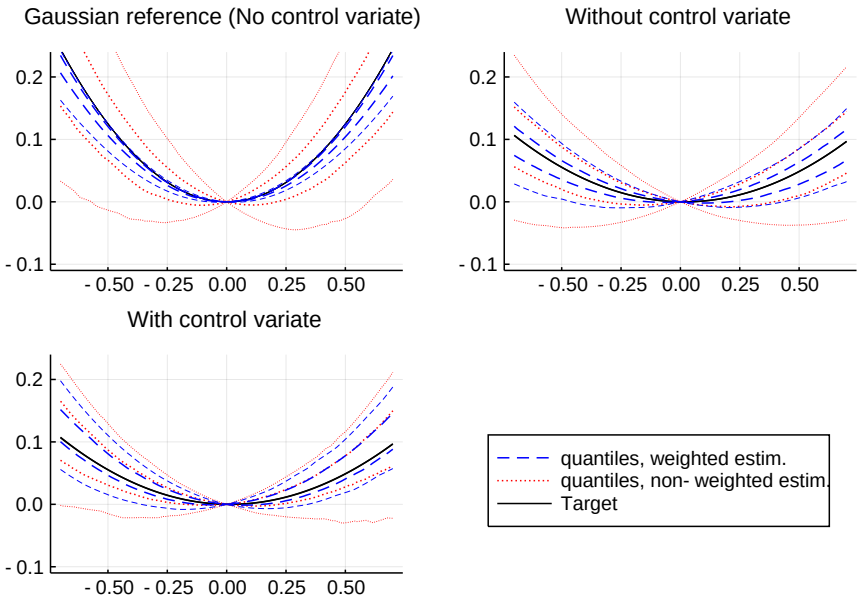


FIGURE 9. The figures above represents the $\{.05, .25, .75, .95\}$-quantiles of the different estimators of the cumulant generating functions for the stationary distribution of the exterior force $F(Q)$, both without weights (dotted), and with weight (dashed). The figure in the upper left corner represents the estimators of the Gaussian reference $\mathbf{k}_{\mathrm{MC}}(\mathbf{X})$ and $\mathbf{k}_{\mathrm{W}}(\mathbf{X})$. The figure in the upper right corner represents the estimators $\mathbf{k}_{\mathrm{MC}}(\mathbf{Y})$ and $\mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})$ (without control variate). Finally, the figure in the upper right corner represents the estimators of $\mathbf{k}_{\mathrm{Cv}}(\mathbf{X}, \mathbf{Y})$ and $\mathbf{k}_{\mathrm{W}}(\mathbf{X}, \mathbf{Y})$ (without control variate)

6.2.3. *Interpretation.* We first remark that the target distribution has an increased variance due to the presence of $\varepsilon \neq 0$. We then remark that the result are very similar as in previous section, except that the statistical error reduction in the case with weights is similar with control variate or without control variate. For any choice of estimator (with or without control variate), we see that the use of weighting substantially improve the statistical error.

6.3. **Conclusion.** In various non-trivial cases where a random quantity is approximated by a Gaussian one dimensional control variate, the Wasserstein reweighting method significantly reduces variance (as compared to a naïve Monte Carlo calculation), and outperforms a control variate variance reduction.

6.4. **Acknowledgments.** We thank P.-M. Samson for suggesting the short proof of Lemma 3.15, and J. Bigot for many constructive remarks that led to many clarifications, and a much nicer proof of Proposition 4.1. This work was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement number 614492.

References

[AS92] Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions with formulas, graphs, and handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover Publications, Inc., New York, 1992.

[BGM10] François Bolley, Arnaud Guillin, and Florent Malrieu, *Trend to equilibrium and particle approximation for a weakly selfconsistent vlasov-fokker-planck equation*, ESAIM: Mathematical Modelling and Numerical Analysis **44** (2010), no. 5, 867–884.

[Das11] Anirban DasGupta, *Probability for statistics and machine learning*, Springer Texts in Statistics, Springer, New York, 2011, Fundamentals and advanced topics. MR 2807365

[GI83] D. Goldfarb and A. Idnani, *A numerically stable dual method for solving strictly convex quadratic programs*, Mathematical Programming **27** (1983), no. 1, 1–33.

[Gla13] Paul Glasserman, *Monte carlo methods in financial engineering*, vol. 53, Springer Science & Business Media, 2013.

[Gly94] P.W. Glynn, *Efficiency improvement techniques*, Annals of Operations Research **53** (1994), no. 1, 175–197

[Gro93] Leonard Gross, *Logarithmic sobolev inequalities and contractivity properties of semigroups*, Dirichlet forms, Springer, 1993, pp. 54–88.

[GS02] Alison L Gibbs and Francis Edward Su, *On choosing and bounding probability metrics*, International statistical review **70** (2002), no. 3, 419–435.

[Jou09] Benjamin Jourdain, *Adaptive variance reduction techniques in finance*, Advanced Financial Modelling **8** (2009), 205.

[Kal02] Olav Kallenberg, *Foundations of modern probability*, Probability and its Applications (New York), Springer-Verlag, New York, 2002.

[LC02] Lars Larsson-Cohn, *L p -norms of hermite polynomials and an extremal problem on wiener chaos*, Ark. Mat. **40** (2002), no. 1, 133–144.

[MR95] Rajeev Motwani and Prabhakar Raghavan, *Randomized algorithms*, Cambridge University Press, Cambridge, 1995. MR 1344451

[Owe13] Art B. Owen, *Monte carlo theory, methods and examples*, 2013.

[PS18] François Portier and Johan Segers, *Monte carlo integration with a growing number of control variates*, arXiv preprint arXiv:1801.01797 (2018).

[Ros13] Sheldon M. Ross, *Simulation*, Elsevier/Academic Press, Amsterdam, 2013, Fifth edition [of 1433593]. MR 3294208

[Str11] Daniel W. Stroock, *Probability theory*, second ed., Cambridge University Press, Cambridge, 2011, An analytic view. MR 2760872

[Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR MR1964483 (2004e:90003)

[Vil08] ———, *Optimal transport: old and new*, vol. 338 Springer Science & Business Media, 2008.

Mathias Rousset, e-mail: mathias.rousset(AT)inria.fr

INRIA Rennes & IRMAR, Université de Rennes 1
France

Pierre-André Zitt, e-mail: Pierre-André.Zitt(AT)u-pem.fr

Yushun Xu, e-mail: yushun.xu(AT)u-pem.fr

Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne-la-Vallée
5, boulevard Descartes, Cité Descartes - Champs-sur-Marne
77454 Marne-la-Vallée cedex 2
France