

MULTIPLE IMPORTANCE SAMPLING WITH OVERLAPPING SETS OF PROPOSALS

Víctor Elvira[‡], Luca Martino[†], David Luengo[‡], Mónica F. Bugallo[◇]

[‡] Dep. of Signal Theory and Communic., Universidad Carlos III de Madrid, Leganés (Spain).

[†] Institute of Mathematical Sciences and Computing, Universidade de São Paulo, São Carlos (Brazil).

[‡] Dep. of Signal Theory and Communic., Universidad Politécnica de Madrid, Madrid (Spain).

[◇] Dept. of Electrical and Computer Eng., Stony Brook University, Stony Brook, NY 11794 (USA).

ABSTRACT

In this paper, we introduce multiple importance sampling (MIS) approaches with overlapping (i.e., non-disjoint) sets of proposals. We derive a novel weighting scheme, based on the deterministic mixture methodology, that leads to unbiased estimators. The proposed framework can be seen as a generalization of other well-known MIS algorithms available in the literature. Furthermore, it allows us to achieve any desired trade-off between the variance of the estimators and the computational complexity through the definition of the sets of proposals. Simulations using a bimodal target density show the good performance of the proposed approach.

Index Terms— Multiple importance sampling, variance reduction, Monte Carlo methods, Bayesian inference

1. INTRODUCTION

Importance sampling (IS) methods are usually employed for approximating moments of the distribution of a variable of interest [1, 2]. In standard IS, several samples are drawn from a single proposal distribution. The samples are then properly weighted in order to account for the mismatch between the proposal and the targeted distribution. Since the variance of the estimator can be substantially increased by a large mismatch, several proposal densities are often used to draw the samples. This variant of the method, called multiple importance sampling (MIS), is implicitly used in many known sequential and static algorithms, such as particle filtering [3], Population Monte Carlo (PMC) [4], adaptive MIS (AMIS) [5] or adaptive population importance sampling (APIS) [6].

In a recent work [7], we have analyzed MIS methods in a systematic way, defining a set of proper sampling and weighting mechanisms that yield unbiased and consistent natural estimators (see [7, Section IV] for further details). The schemes

described in [7] encompass all the MIS algorithms available in the literature (as far as we know), as well as some novel alternatives. Among all these weighting-sampling strategies, the so-called *deterministic mixture* (DM) method is shown to provide the best performance in terms of the estimator's variance. This approach evaluates the whole set (mixture) of proposals at the denominator of each weight, thus potentially leading to a remarkable increase in the computational cost.

In another recent work [8], we proposed to cluster the set of proposals into disjoint sets, applying at each denominator only the proposals that belong to each cluster. This procedure can attain most of the variance reduction of the DM approach, while keeping a limited computational cost. In this paper, we propose a novel weighting scheme where the sets of proposals can be non-disjoint (i.e., can overlap). This means that some proposals (i.e., the good ones) may be used in more than one mixture to reduce the variance of the weights and the resulting estimators. This novel method can be seen as a generalization of several methods, including the DM-MIS method, the standard MIS approach, and the partial DM-MIS scheme [8].

The rest of the paper is structured as follows. In Section 2, we first state the problem and then review IS and MIS methods. Section 3 introduces the new weighting scheme, justifies its validity and advantages, and illustrates it through an example. Finally, a numerical simulation is presented in Section 4, while Section 5 contains some concluding remarks.

2. PROBLEM STATEMENT AND BACKGROUND

The interest lies in computing a moment of a posterior probability density function (pdf), $\tilde{\pi}(\mathbf{x})$. In many applications, the pdf can be evaluated only up to a normalization factor, Z . In this way,

$$\tilde{\pi}(\mathbf{x}) = \frac{\pi(\mathbf{x})}{Z}, \quad (1)$$

where $\pi(\mathbf{x})$ is the unnormalized target, and $Z = \int \pi(\mathbf{x})d\mathbf{x}$. The computation of a particular moment of $\tilde{\pi}(\mathbf{x})$ is given by

$$I = \frac{1}{Z} \int f(\mathbf{x})\pi(\mathbf{x})d\mathbf{x}, \quad (2)$$

This work has been supported by the ERC grant 239784 and AoF grant 251170; the Spanish government through the OTOSiS (TEC2013-41718-R), DISSECT (TEC2012-38058-C03-01) and MIMOD-PLC (TEC2015-64835-C3-3-R) projects; the BBVA Foundation (MG-FIAR project); Grant 2014/23160-6 of the São Paulo Research Foundation (FAPESP); Grant 305361/2013-3 of the National Council for Scientific and Technological Development (CNPq); and the NSF under Award CCF-0953316.

where f can be any square-integrable function w.r.t. $\pi(\mathbf{x})$. In practical scenarios, we cannot obtain an analytical solution of (2) and/or Z is not known, and therefore, Monte Carlo methods are commonly used to approximate I .

2.1. Importance Sampling (IS)

Let us consider a set of N samples $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ drawn from a proposal pdf, $q(\mathbf{x})$, with heavier tails than the target, $\pi(\mathbf{x})$. The samples have associated importance weights given by

$$w_n = \frac{\pi(\mathbf{x}_n)}{q(\mathbf{x}_n)}, \quad n = 1, \dots, N. \quad (3)$$

Using the samples and weights, the moment of interest can be approximated as

$$\tilde{I}_{\text{IS}} = \frac{1}{N\hat{Z}} \sum_{n=1}^N w_n f(\mathbf{x}_n), \quad (4)$$

where $\hat{Z} = \frac{1}{N} \sum_{j=1}^N w_j$ is an unbiased estimator of Z [1]. Eq. (4) is a consistent estimator of I , and its variance is directly related to the discrepancy between $\pi(\mathbf{x})|f(\mathbf{x})|$ and $q(\mathbf{x})$ (for a specific choice of $f(\mathbf{x})$) or to the mismatch between the target $\pi(\mathbf{x})$ and the proposal $q(\mathbf{x})$ for a general $f(\mathbf{x})$ [9].

2.2. Multiple Importance Sampling (MIS)

Since finding a unique good proposal pdf, $q(\mathbf{x})$, is usually unfeasible, a more robust strategy is using a population of proposal pdfs. This scheme is often known in the literature as *multiple importance sampling* (MIS) [10]. Let us consider a set of N proposal pdfs, $\{q_1(\mathbf{x}), \dots, q_N(\mathbf{x})\}$, with exactly one sample drawn from each of them, i.e.,

$$\mathbf{x}_n \sim q_n(\mathbf{x}), \quad n = 1, \dots, N. \quad (5)$$

Note that we use this sampling procedure in the whole paper. Several valid importance weights can be associated to these samples [7]. In particular, two usual approaches are:

1. *Standard MIS* (s-MIS) [2, Section 2.5]:

$$w_n = \frac{\pi(\mathbf{x}_n)}{q_n(\mathbf{x}_n)}, \quad n = 1, \dots, N. \quad (6)$$

2. *Deterministic mixture MIS* (DM-MIS) [10]:

$$w_n = \frac{\pi(\mathbf{x}_n)}{\psi(\mathbf{x}_n)} = \frac{\pi(\mathbf{x}_n)}{\frac{1}{N} \sum_{j=1}^N q_j(\mathbf{x}_n)}, \quad n = 1, \dots, N, \quad (7)$$

where $\psi(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^N q_j(\mathbf{x})$ is the mixture pdf, composed of all the proposal pdfs. Several interpretations of this strategy are discussed in [7].

The consistency of the estimators is ensured in both cases. In [7] it is proved that the DM-MIS weights provide a superior performance than the s-MIS weights when Z is known (i.e., the estimator of Eq. (4) always has a lower or equal variance with the DM-MIS weights than with the s-MIS weights). However, the DM-MIS estimator is computationally more expensive, since it requires N evaluations of the proposal to obtain each weight instead of just one, i.e., N^2 in total instead of just N . Note that the number of evaluations of the target $\pi(\mathbf{x})$ is the same regardless of whether the weights are calculated according to (6) or (7), so this increase may be negligible in some situations. However, in some other scenarios, where evaluating the target and the proposals has a similar cost or the number of proposals grows (e.g., in AMIS), the increase in computational load can be substantial. In these cases, alternative efficient solutions, like the one described in the following section, must be devised.

2.3. Partial DM-MIS

In [8], a *partial* DM-MIS (p-DM-MIS) scheme is proposed. Since the excessive computational cost of the full DM-MIS comes from evaluating each sample in the whole set of proposals, in p-DM-MIS the set of proposal pdfs, $\{q_n(\mathbf{x})\}_{n=1}^N$, are grouped into P mixtures composed of M mixands, with $PM = N$.¹ A partition of the set $\{1, \dots, N\}$ is thus formed, yielding P disjoint subsets of M indices, \mathcal{S}_p with $p = 1, \dots, P$, s.t.

$$\mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots \cup \mathcal{S}_P = \{1, \dots, N\}, \quad (8)$$

with $\mathcal{S}_k \cap \mathcal{S}_\ell = \emptyset$ for all $k, \ell = 1, \dots, P$ and $k \neq \ell$. Each subset, $\mathcal{S}_p = \{j_{p,1}, j_{p,2}, \dots, j_{p,M}\}$, contains M indices, $j_{p,m} \in \{1, \dots, N\}$ for $m = 1, \dots, M$ and $p = 1, \dots, P$. The weights of the p -th mixture are then computed as

$$w_n = \frac{\pi(\mathbf{x}_n)}{\psi_p(\mathbf{x}_n)} = \frac{\pi(\mathbf{x}_n)}{\frac{1}{M} \sum_{j \in \mathcal{S}_p} q_j(\mathbf{x}_n)}, \quad n \in \mathcal{S}_p. \quad (9)$$

The partial DM-MIS approach can be seen as a generalization of the two alternatives of Section 2.2. Indeed, the particular cases $P = 1$ and $P = N$ correspond to the *full* DM-MIS (f-DM-MIS) and the *standard* MIS (s-MIS) approaches, respectively.² In p-DM-MIS the number of evaluations of the proposal pdfs is PM^2 . Since $N \leq PM^2 = NM \leq N^2$, the computational cost is larger than that of s-MIS approach (M times larger), but lower than that of the f-DM-MIS approach (since $M \leq N$). When Z is known, it can be proved that

$$\text{Var}(\hat{I}_{\text{f-DM-MIS}}) \leq \text{Var}(\hat{I}_{\text{p-DM-MIS}}) \leq \text{Var}(\hat{I}_{\text{s-MIS}}), \quad (10)$$

¹For the sake of simplicity, we assume that all the mixtures contain the same number of proposal pdfs. However, any strategy that clusters the N proposals into P disjoint mixtures (regardless of their size) is valid.

²From now on, we denote the deterministic mixture approach with the weights of Eq. (7) as f-DM-MIS.

for any choice of P and any strategy to group the original proposals $\{q_n(\mathbf{x})\}_{n=1}^N$ into mixtures (see [8, Appendix B]). There is then a clear tradeoff between performance and computational cost: a smaller number of mixtures (P) reduces the variance, but at the expense of an increase in the number of evaluations of the proposal densities.

3. BEYOND DISJOINT SETS

3.1. The intuition for the variance reduction in DM-MIS

Let us consider a multi-modal target $\pi(\mathbf{x})$, and let us assume that we want to estimate the normalizing constant ($Z = 1$), by sampling from N proposals. Some of the proposals (“good” proposals) are covering at least one of the modes, whereas others (“bad” proposals) are placed in areas without relevant probability mass of the target. Let us consider the s-MIS approach with the weights of Eq. (6). Since the estimator \hat{Z} is unbiased (i.e., $E[\hat{Z}] = 1$), the increase in the variance comes from uncommon realizations where some weights take a high value. This typically happens when a sample is drawn from a “bad” proposal, this sample falls in the tail of the proposal (i.e., we have small value at the weight’s denominator), and the target has a relevant mass in that location (i.e., we have a large value at the weight’s numerator). This situation can be easily avoided if there are other proposals covering that mode: if we use the DM-MIS weights of Eq. (7), the denominator will increase substantially when there is at least one “good” proposal placed close to the sample (since the evaluation of that “good” proposal will dominate). Note that the numerator of the weight is the same in both s-MIS and DM-MIS.

The p-DM-MIS with random allocation increases the chance w.r.t. the s-MIS weights of having at least one proposal within the mixture covering each mode. However, the p-DM-MIS setup can be generalized by allowing intersection among the subsets (mixtures) of proposals. The theoretical justification is the following: the “good” proposals might be exploited in more than one mixture in order to reduce the variance of the weights of the samples generated from the “bad” proposals in those mixtures. In order to successfully implement this novel overlapping partial DM-MIS scheme, there are two major challenges: “good” proposals must be identified, and the weights of the samples must be modified so that the resulting estimators are still consistent. Finding a good set of clusters is a delicate task that we plan to address in the future. In the rest of the paper, we assume that the clustering is given and focus on the second issue: finding a valid set of weights that yields consistent estimators.

3.2. Novel weighting procedure

Let us consider the overlapping sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_P$, s.t. $S_j \cup S_k \neq \emptyset$ for some $j, k \in 1, \dots, P$ and $j \neq k$. Since the n -th proposal now belongs to more than one mixture, the weight w_n of \mathbf{x}_n must be computed as the average of the *sub-weights*

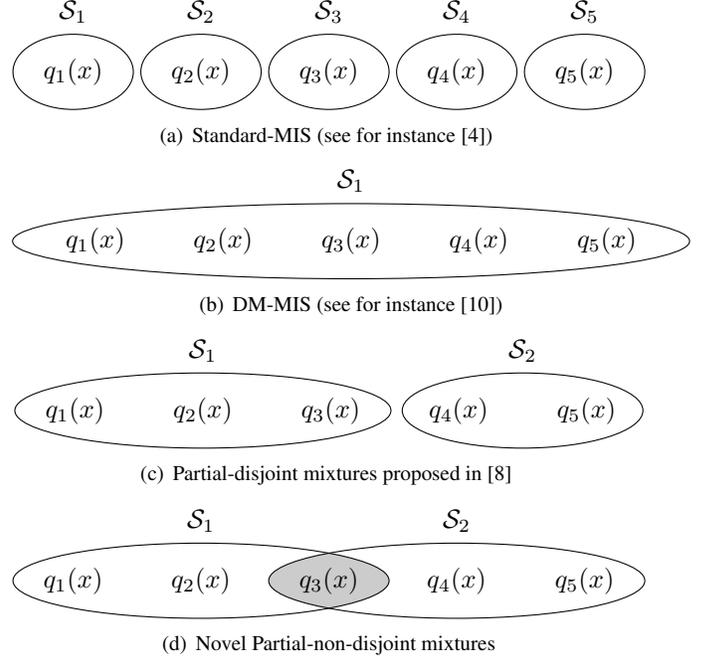


Fig. 1. (Example of Section 3.3) Proposal allocation to the subsets in different weighting schemes.

$w_n^{(j)}$ corresponding to each of those mixtures, i.e.,

$$w_n = \frac{1}{m_n} \sum_{j \in \mathcal{M}_n} w_n^{(j)}, \quad n = 1, \dots, N, \quad (11)$$

where $\mathcal{M}_n = \{j : n \in \mathcal{S}_j\}$ and m_n is the number of mixtures where q_n appears, i.e., $m_n = |\mathcal{M}_n|$. The *sub-weights* of the n -th sample in the p -th mixture are computed as

$$w_n^{(p)} = \frac{\pi(\mathbf{x}_n)}{\phi_p(\mathbf{x}_n)} = \frac{\pi(\mathbf{x}_n)}{\frac{1}{\sum_{k \in \mathcal{S}_p} \lambda_k} \sum_{i \in \mathcal{S}_p} \lambda_i q_i(\mathbf{x}_n)}, \quad n \in \mathcal{S}_p, \quad (12)$$

where $\lambda_j = \frac{1}{m_j}$. Note that, if the subsets are disjoint, then $m_j = 1 \forall j = 1, \dots, N$, which yields the weights of Eq. (9) [8]. Therefore, this weighting scheme can be seen as a generalization of the p-DM-MIS method. Note also that the factors λ_j imply that the proposals that are present in more mixtures have less impact on each one. In this way, the estimator \hat{Z} with the weights of (11) is consistent (\hat{Z} is also unbiased).

3.3. Example with $N = 5$ proposals

Let us illustrate the novel approach through a unidimensional example with $N = 5$ Gaussian proposals with equal variance σ^2 , i.e., $q_n(x) = \mathcal{N}(x; \mu_n, \sigma^2)$ for $n = 1, \dots, 5$. For the sake of clarity in the explanation, let us assume (without loss of generality) that $\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4 \leq \mu_5$. Fig. 1 shows a graphical representation of different clusterings of the proposals for both the disjoint and the non-disjoint approaches.

	Partial-disjoint mixtures [8]		Partial-non-disjoint mixtures	
	w_n	\mathcal{M}_n	w_n	\mathcal{M}_n
$x_1 \sim q_1$	$\frac{\pi(x_1)}{\frac{1}{3}(q_1(x_1)+q_2(x_1)+q_3(x_1))}$	{1}	$\frac{\pi(x_1)}{\frac{1}{2.5}(q_1(x_1)+q_2(x_1)+\frac{1}{2}q_3(x_1))}$	{1}
$x_2 \sim q_2$	$\frac{\pi(x_2)}{\frac{1}{3}(q_1(x_2)+q_2(x_2)+q_3(x_2))}$	{1}	$\frac{\pi(x_2)}{\frac{1}{2.5}(q_1(x_2)+q_2(x_2)+\frac{1}{2}q_3(x_2))}$	{1}
$x_3 \sim q_3$	$\frac{\pi(x_3)}{\frac{1}{3}(q_1(x_3)+q_2(x_3)+q_3(x_3))}$	{1}	$\frac{1}{2}\frac{\pi(x_3)}{\frac{1}{2.5}(q_1(x_3)+q_2(x_3)+\frac{1}{2}q_3(x_3))} + \frac{1}{2}\frac{\pi(x_3)}{\frac{1}{2.5}(\frac{1}{2}q_3(x_3)+q_4(x_3)+q_5(x_3))}$	{1, 2}
$x_4 \sim q_4$	$\frac{\pi(x_4)}{\frac{1}{2}(q_4(x_4)+q_5(x_4))}$	{2}	$\frac{\pi(x_4)}{\frac{1}{2.5}(\frac{1}{2}q_3(x_4)+q_4(x_4)+q_5(x_4))}$	{2}
$x_5 \sim q_5$	$\frac{\pi(x_5)}{\frac{1}{2}(q_4(x_5)+q_5(x_5))}$	{2}	$\frac{\pi(x_5)}{\frac{1}{2.5}(\frac{1}{2}q_3(x_5)+q_4(x_5)+q_5(x_5))}$	{2}

Table 1. (Example of Section 3.3) Weights applied to each sample in different weighting schemes.

On the one hand, Figs. 1(a) and 1(b) show respectively the s-MIS (i.e., as many clusters as proposals) and f-DM-MIS (i.e., a single cluster) approaches. On the other hand, Fig. 1(c) shows an example of a p-DM-MIS clustering: if we want to form two disjoint subsets, then a reasonable choice could be $\mathcal{S}_1 = \{1, 2, 3\}$ and $\mathcal{S}_2 = \{4, 5\}$. Finally, in a non-disjoint approach, like that shown in Fig. 1(d), the proposal q_3 can belong to both subsets, $\mathcal{S}_1 = \{1, 2, 3\}$ and $\mathcal{S}_2 = \{3, 4, 5\}$. Table 1 shows the weights of each sample with p-DM-MIS and with the novel approach.

4. NUMERICAL RESULTS

In order to illustrate the performance of the novel approach, let us consider a unidimensional bimodal target which is a mixture of two Gaussian pdfs:

$$\pi(x) = \frac{1}{2}\mathcal{N}(x; \nu_1, c^2) + \frac{1}{2}\mathcal{N}(x; \nu_2, c^2), \quad (13)$$

with $\nu_1 = -1$, $\nu_2 = 1$ and $c^2 = 1$.

The goal is to approximate, via Monte Carlo, the expected value of $x \sim \pi(x)$, i.e., $E[x] = \int_{\mathbb{R}} x\pi(x)dx$, and also the normalizing constant, $Z = \int_{\mathbb{R}} \pi(x)dx$. We apply the MIS algorithm in a setup with $N = 5$ Gaussian proposal pdfs, $\{q_n(x) = \mathcal{N}(x; \mu_n, \sigma^2)\}_{n=1}^N$, where the means are $\mu_1 = -3$, $\mu_2 = -2$, $\mu_3 = 0$, $\mu_4 = 2$, and $\mu_5 = 3$, testing two different values of the variance of the proposals, $\sigma^2 \in \{1, 2\}$. We draw $M = 50$ samples, i.e., exactly 10 samples per proposal. We compare the following weighting schemes:

- **standard-MIS** [4]: The standard MIS approach, where each sample takes into account only the proposal from which it was drawn. This setup corresponds to $\mathcal{S}_1 = \{1\}$, $\mathcal{S}_2 = \{2\}$, $\mathcal{S}_3 = \{3\}$, $\mathcal{S}_4 = \{4\}$, and $\mathcal{S}_5 = \{5\}$.
- **full-DM** [10]: The full-DM algorithm, which corresponds to a unique set $\mathcal{S}_1 = \{1, 2, 3, 4, 5\}$.
- **partial-DM-1** [8]: The partial-DM algorithm with disjoint clusters, $\mathcal{S}_1 = \{1, 2\}$, $\mathcal{S}_2 = \{3\}$ and $\mathcal{S}_3 = \{4, 5\}$.
- **partial-DM-2** [8]: The partial-DM algorithm with disjoint clusters, $\mathcal{S}_1 = \{1, 2, 3\}$ and $\mathcal{S}_2 = \{4, 5\}$.

Method	standard-MIS	full-DM	partial-DM-1	partial-DM-2	ND-partial-DM
MSE(\hat{Z})	16.773	0.0078	15.653	4.2529	0.0161
MSE(\hat{I})	55.22	0.0181	113.15	22.637	0.1928
MSE(\hat{I})	0.1163	0.0185	0.1099	0.1651	0.1374

Table 2. Bimodal target $\pi(x)$ with $\nu_1 = -2$, $\nu_2 = 1$, and $c^2 = 1$. Variance of the proposals $\sigma^2 = 1$. $M = 50$ samples.

Method	standard-MIS	full-DM	partial-DM-1	partial-DM-2	ND-partial-DM
MSE(\hat{Z})	0.6265	0.0103	0.1800	0.0966	0.0100
MSE(\hat{I})	4.2392	0.0239	0.8701	0.4734	0.0683
MSE(\hat{I})	0.1944	0.0245	0.2313	0.1630	0.0681

Table 3. Bimodal target $\pi(x)$ with $\nu_1 = -2$, $\nu_2 = 1$, and $c^2 = 1$. Variance of the proposals $\sigma^2 = 2$. $M = 50$ samples.

- **ND-partial-DM:** The novel non-disjoint approach, with $\mathcal{S}_1 = \{1, 2, 3\}$ and $\mathcal{S}_2 = \{3, 4, 5\}$.

Table 2 shows the results of the mean squared error (MSE) of \hat{Z} , the natural estimator of the normalizing constant; \hat{I} , the natural estimator of the target mean (i.e., $f(x) = x$ and assuming Z is known); and \tilde{I} , the self-normalized estimator of the target mean. The variance of all the proposal pdfs is $\sigma^2 = 1$, and the results are averaged over $5 \cdot 10^4$ runs. Note that the novel approach provides a lower MSE than the two p-DM-MIS schemes for all the estimators. Table 3 shows the same results for $\sigma^2 = 2$, leading to similar conclusions.

5. CONCLUSIONS

In this work, we have introduced a novel weighting scheme for multiple importance sampling (MIS). Based on the deterministic mixture weighting approach, the method forms subsets of proposals, allowing some of them to belong to more than one subset. The weight function for the samples drawn from each proposal is calculated in order to ensure that the resulting estimators are consistent. The novel approach is able to reduce the variance of these estimators while keeping a low computational cost, as shown in a numerical simulation example with a bimodal target and five proposals. In the future, we plan to address the creation of the clusters.

6. REFERENCES

- [1] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004.
- [2] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer, 2004.
- [3] N. Gordon, D. Salmond, and A. F. M. Smith, “Novel approach to nonlinear and non-Gaussian Bayesian state estimation,” *IEE Proceedings-F Radar and Signal Processing*, vol. 140, pp. 107–113, 1993.
- [4] O. Cappé, A. Guillin, J. M. Marin, and C. P. Robert, “Population Monte Carlo,” *Journal of Computational and Graphical Statistics*, vol. 13, no. 4, pp. 907–929, 2004.
- [5] J. M. Cornuet, J. M. Marin, A. Mira, and C. P. Robert, “Adaptive multiple importance sampling,” *Scandinavian Journal of Statistics*, vol. 39, no. 4, pp. 798–812, December 2012.
- [6] L. Martino, V. Elvira, D. Luengo, and J. Corander, “An adaptive population importance sampler: Learning from the uncertainty,” *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4422–4437, 2015.
- [7] V. Elvira, L. Martino, D. Luengo, and M. F. Bugallo, “Generalized multiple importance sampling,” *arXiv:1511.03095*, 2015.
- [8] V. Elvira, L. Martino, D. Luengo, and M. F. Bugallo, “Efficient multiple importance sampling estimators,” *Signal Processing Letters, IEEE*, vol. 22, no. 10, pp. 1757–1761, 2015.
- [9] A. Doucet and A. M. Johansen, “A tutorial on particle filtering and smoothing: Fifteen years later,” *Handbook of Nonlinear Filtering*, vol. 12, no. 656-704, pp. 3, 2009.
- [10] A. Owen and Y. Zhou, “Safe and effective importance sampling,” *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 135–143, 2000.