Waste-free sequential Monte Carlo for light source detection in astronomical images

STATS 608 final project

Jaylin Lowe and Tim White

Due April 29th, 2024

Contents

1 Introduction

1.1 Background

Sequential Monte Carlo (SMC) methods are a flexible class of algorithms for approximating sequences of probability distributions via weighted samples, or particles. Many varieties of SMC algorithms exist, but they generally rely on the same core procedures: sampling particles from a tractable initial distribution, mutating them with Markov chain Monte Carlo (MCMC) moves, updating their weights to track the sequence of target distributions, and resampling them to avoid weight degeneracy [1]. SMC is often employed in settings where the entire sequence of distributions is of interest, including as a filtering algorithm for state-space models [2, 3]. However, the same underlying ideas can also be used to conduct inference for a single static target, which may be high-dimensional or multimodal. In such settings, these algorithms are referred to as SMC samplers [4, 5].

In a Bayesian regime with latent random variables z and observed random variables x, SMC samplers offer a potentially convenient method of approximating the posterior $p(z \mid x) \propto p(z)p(x \mid z)$, where $p(z)p(x \mid z)$ can be evaluated pointwise but $p(z \mid x)$ cannot be sampled from directly due to the intractable normalizing constant $p(x)$. A common approach in this setting is to define the prior $p(z)$ as the initial distribution and the posterior as the final target, and to bridge the two with a sequence of auxiliary targets $p(z)p(x \mid z)^{\tau_t}$ constructed according to a tempering schedule $0 = \tau_0 < \cdots < \tau_T = 1$. For problems of this nature, SMC samplers are a promising alternative to importance sampling and MCMC, as (i) they are amenable to parallelization over particles, (ii) they allow straightforward estimation of $p(x)$, and (iii) they can be adapted based on the current particles to facilitate exploration of the latent space [6].

The empirical performance of SMC samplers typically hinges on the design of the tempering, resampling, and mutation stages of the algorithm. There are well-established tempering and resampling strategies that tend to be suitable for many problems. For the former, one can use root-finding to select temperatures τ_t that minimize the chi-square pseudo-distance between consecutive targets [7], while for the latter it is common to use stratified or systematic resampling [8]. In contrast, many choices for the mutation procedure are less reliable, particularly for high-dimensional or otherwise complex models. A prevalent strategy is to apply a k-step Metropolis-Hastings kernel within each SMC iteration, using the empirical covariance of the weighted particles to inform the covariance of the proposal distribution. However, it is not trivial to select a suitable value of k, either a priori or adaptively. The performance of SMC samplers is known to degrade when k is too small due to insufficient mixing of the mutation kernels, but the algorithm is computationally costly when k is too large since its complexity is linear in this quantity [9].

A recent proposal by Dau and Chopin offers a principled solution to this trade-off with a "waste-free" variant of SMC that claims to (i) decrease the effort required to tune the Metropolis-Hastings mutation kernel and (ii) reduce the asymptotic variance of SMC estimators under some assumptions [10]. Whereas a standard SMC sampler resamples N of the N weighted particles in each iteration, mutates each of them k times, and discards the intermediate mutations, waste-free SMC resamples $M{<<}N$ particles and mutates each of them $P{-}1$ times, keeping all mutations and thus recovering $N{=}MP$ particles. Dau and Chopin demonstrate the advantages of this procedure compared to standard SMC in three numerical experiments: fitting a Bayesian logistic regression model, enumerating Latin squares, and evaluating Gaussian orthant probabilities. However, the finite-sample benefits of their approach (or potential lack thereof) in other complex inference settings have not yet been explored.

1.2 Objectives and outline

In all three of Dau and Chopin's experiments, the dimension of the target distribution is fixed within each SMC iteration even if the dimension of the state space increases with time. However, there exist many interesting model classes (e.g., finite mixture models, change-point models) for which the number of unknown parameters to infer is itself unknown [11]. Motivated by this, we will study the utility of embedding waste-free resampling and mutation within an SMC sampler designed for a challenging Bayesian inverse problem that requires transdimensional inference: detecting and distinguishing overlapping light sources in astronomical images [12]. While our SMC algorithm is tailored to the task of object detection, it nonetheless falls within the class of samplers considered by Dau and Chopin. Hence, we theorize that the benefits of their waste-free approach are potentially attainable in this setting.

The remainder of this report is guided by the following questions:

1. Why are waste-free SMC samplers purported to facilitate implementation and reduce asymptotic variance compared to standard SMC samplers?
2. How does the variability of posterior estimates change across mutation kernel parameters for standard and waste-free SMC samplers? Is this variability smaller for one method than the other?
3. How accurate and calibrated are the posterior estimates of a waste-free SMC sampler in the context of light source detection? How does this compare to a standard SMC sampler?

After introducing our model and SMC sampler for light source detection in subsection 2.1 and subsection 2.2, respectively, we will address the first of the above questions in subsection 2.3. We will highlight the algorithmic differences between the standard and waste-free versions of our sampler and discuss the favorable properties of the waste-free approach. In section 3, we will investigate the remaining two guiding questions through experiments involving small synthetic images of crowded starfields. First, in subsection 3.1, we will assess the variability of several light source detection metrics across many runs of the standard and waste-free samplers on the same set of four images. Then, in subsection 3.2, we will evaluate the accuracy and calibration of the standard and waste-free samplers across many images using similar metrics. Finally, we will conclude in section 4 with a discussion of the limitations of our analysis and the challenges of transdimensional inference.

2 Methods

2.1 Notation and Bayesian model

Astronomical cataloging is the task of inferring the positions and properties of light sources from images captured by large sky surveys. This task is challenging for images that contain visually overlapping light sources, as the number of sources in such an image is uncertain and the properties of the blended sources are ambiguous [13]. Here, we take a probabilistic approach to astronomical cataloging by considering a generative model for small images depicting many visually overlapping stars.

Let s denote the number of stars in an image of height H pixels and width W pixels. We will henceforth refer to s as the source count of the image, and we define $s \sim \text{Uniform}\{0, 1, 2, ..., D\}$ for some maximum number of stars D. Each star in the image has a location and a flux (i.e., brightness). Given s, we model i.i.d. locations $u_1, u_2, ..., u_s \sim \text{Uniform}([0, H] \times [0, W])$ and fluxes $f_1, f_2, ..., f_s \sim \mathcal{N}(\mu, \sigma^2)$, where μ and σ are center and scale parameters informed by astrophysics.

Together, the collection of latent variables defined above forms a catalog $z := \{s, \{u_j\}_{j=1}^s, \{f_j\}_{j=1}^s\}$ that describes the imaged stars. Given a catalog z, the intensity of the image at pixel (h, w) is $x_{hw} \,|\, z \sim \text{Poisson}(\lambda_{hw}(z) + \gamma)$, where γ is the background intensity of the image and $\lambda_{hw}(z)$ is the weighted sum of the fluxes at pixel (h, w), with weights determined by a bivariate Gaussian point-spread function.

2.2 Sequential Monte Carlo for light source detection

Given an image $x = \{\{x_{hw}\}_{h=1}^H\}_{w=1}^W$, we aim to characterize the posterior distribution $p(z \,|\, x)$ of possible catalogs that explain the image. To do so, we generate a collection of weighted catalogs from $p(z \,|\, x)$ using an SMC sampler that was recently proposed with the task of light source detection in mind [14]. This sampler, which uses likelihood tempering as introduced in subsection 1.1, offers a convenient approach to transdimensional inference. It does not require the user to define, or subsequently sample from, transdimensional proposals, which can hinder the scalability and mixing quality of potential alternatives such as reversible jump MCMC [15].

Instead, this SMC sampler initializes an equal number of catalogs with each candidate value of s and preserves the source count of each catalog for the duration of the algorithm. A set of importance weights is maintained for each "block" of catalogs that have the same source count, and the resampling step of the algorithm is performed within these blocks using the intra-block weights. The sampler otherwise progresses through the usual mutation and reweighting stages of SMC, exploring the space of light source fluxes and locations for catalogs with various source counts and iteratively assigning weights to catalogs based on their plausibility under the current target. A separate set of inter-block importance weights is also maintained throughout the algorithm, and the inter-block weights of the catalogs returned after the final iteration can be used to assess the posterior probabilities of different values of s.

Algorithm 1 `Sequential Monte Carlo sampler, stratified by source count`

Input: image x; prior $p(z)$; likelihood $p(x \mid z)$; Metropolis-Hastings kernel $\mathrm{MH}_\tau(z, dz)$ for $\tau \in [0, 1]$;
number of blocks B (indexed by b); number of catalogs per block N (indexed by n);
choice of resampling scheme (e.g., multinomial, stratified, systematic),
choice of standard mutation (with parameter k) or waste-free mutation (with parameters M and $P{=}N/M$).

Iteration $t \leftarrow 0$. Tempering exponent $\tau_t \leftarrow 0$.

while $\tau_t < 1$ **do**

$t \leftarrow t + 1$.

if $t = 0$ **then**

Catalogs $z_t^{bn} \sim p(z)$ such that $\forall\, b,\ \ s_t^{b1} = \cdots = s_t^{bN}$.
Unnormalized weights $w_t^{bn} \leftarrow 1$.
Intra-block normalized weights $\widetilde{W}_t^{bn} \leftarrow \frac{1}{N}$.
Inter-block normalized weights $W_t^{bn} \leftarrow \frac{1}{BN}$.

INITIALIZE

if $t > 0$ **then**

for block $b \in \{1, \ldots, B\}$ **do**

if `standard` **then**

Resample N indices $\{A_t^{bn}\}_{n=1}^{N} \leftarrow$ `resample`$(N,\ \{w_{t-1}^{bn}\}_{n=1}^{N})$.

For $n \in \{1, \ldots, N\}$, mutate z_t^{bn} with k-step kernel $\mathrm{MH}_{\tau_{t-1}}(z_{t-1}^{A_t^{bn}}, dz)$ and keep the final mutation.

if `waste-free` **then**

Resample M indices $\{A_t^{bn}\}_{n=1}^{M} \leftarrow$ `resample`$(M,\ \{w_{t-1}^{bn}\}_{n=1}^{N})$.

For $n \in \{1, \ldots, M\}$, mutate z_t^{bn} with $(P{-}1)$-step kernel $\mathrm{MH}_{\tau_{t-1}}(z_{t-1}^{A_t^{bn}}, dz)$ and keep all mutations.

For $n \in \{1, \ldots, N\}$, reset $\widetilde{W}_{t-1}^{bn} \leftarrow \frac{1}{N}$ and $W_{t-1}^{bn} \leftarrow \frac{1}{N} \sum_{n=1}^{N} W_{t-1}^{bn}$.

RESAMPLE AND MUTATE

Update $\tau_t \leftarrow \tau_{t-1} + \delta$, where $\delta \in [0, 1{-}\tau_{t-1}]$. TEMPER

Update $w_t^{bn} \leftarrow W_{t-1}^{bn} p(x \mid z_t^{bn})^{\tau_t - \tau_{t-1}}$.
Update $\widetilde{W}_t^{bn} \leftarrow w_t^{bn} / \sum_n w_t^{bn}$.
Update $W_t^{bn} \leftarrow w_t^{bn} / \sum_b \sum_n w_t^{bn}$.

UPDATE WEIGHTS

Output: Weighted catalog approximation $\{\{z_t^{bn}, W_t^{bn}\}_{b=1}^{B}\}_{n=1}^{N}$ of $p(z \mid x)$.

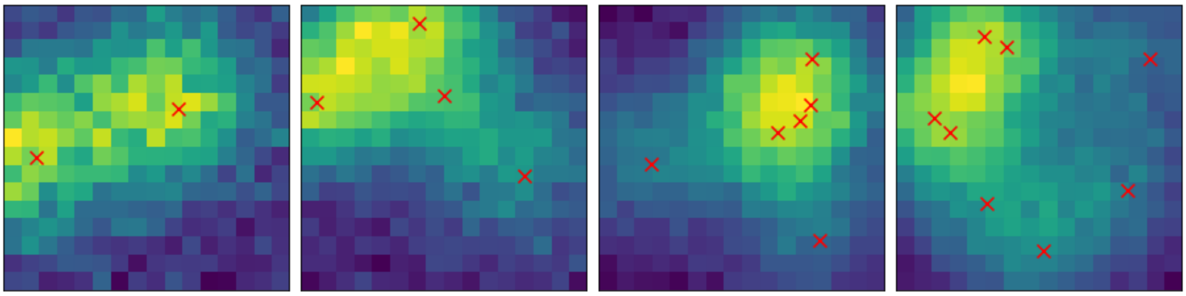
We formalize this procedure, which we will henceforth refer to as the "standard" version of our SMC sampler, in Algorithm 1. We emphasize that the mutation kernel for this standard sampler is assumed to be a k-step Metropolis-Hastings kernel that is invariant under the current target distribution, as this is the setting considered by Dau and Chopin. In the context of light source detection, this kernel comprises a Gaussian random walk for the flux of each light source and a truncated Gaussian random walk for the location, with proposal variances fixed to a fraction of the prior variance in both cases.

As discussed in subsection 1.1, the number of Metropolis-Hastings steps k per SMC iteration is a crucial user-specified parameter. A large k is favorable since it permits greater exploration of potential fluxes and locations, but too large a k may be computationally prohibitive since the mutation step requires skN likelihood evaluations per SMC iteration for each block of catalogs, where s is the source count of catalogs in that block. Methods have been proposed to set k adaptively — e.g., by stopping the mutation procedure when some condition on the squared jumping distance is satisfied — but Dau and Chopin assert that these strategies tend to be empirically unreliable or conceptually flawed.

2.3 Waste-free resampling and mutation

Dau and Chopin's waste-free formulation of the resampling and mutation steps, as introduced in subsection 1.1, is motivated by the trade-off between computational demand and latent space exploration that is inherent in the choice of k [10]. Their approach, which can be applied to a general class of SMC samplers, including ours, claims to increase the robustness of the algorithm to the choice of mutation kernel parameters. It requires only a minor modification to a standard SMC sampler: Instead of resampling N catalogs, mutating each catalog k times, and keeping the final mutated catalogs, we now resample $M{<<}N$ catalogs, mutate each of them $P{-}1$ times, and keep all $N{=}MP$ resulting catalogs. Thus, waste-free SMC essentially replaces the choice of k with the choice of M, and the latter is purported to exert less influence on the quality of the weighted catalogs produced by the algorithm. In the SMC sampler

Figure 1: Four 16×16 images from our Bayesian model with true source count $s \in \{2, 4, 6, 8\}$
(Pixel intensities reflect magnitude of flux and red crosses indicate locations of light sources)



introduced in subsection 2.2, this modification occurs within each block of catalogs, as demonstrated in Algorithm 1.

The authors prove that waste-free SMC samplers yield an unbiased estimate of the normalizing constant of each target distribution in the sequence, and that estimators computed under the weighted particle approximation of each target are consistent and asymptotically normal. While these properties are the same as those possessed by a standard SMC sampler, the waste-free procedure also provides a notable additional benefit: Under some assumptions in the so-called long-chain regime where $P \to \infty$, the output of a waste-free SMC sampler has a smaller asymptotic variance than that of a corresponding standard SMC sampler.

As Dau and Chopin discuss, the supposed stability and reduced variance of waste-free SMC samplers is attributable to (i) their reliance on only a small number of potentially highly correlated ancestor catalogs in each iteration and (ii) their utilization of the $P-2$ "intermediate" mutations that would be discarded by a standard SMC sampler. An additional insight that the authors do not allude to is the potential link between their waste-free procedure and product-form Monte Carlo estimators [16]. The mutation procedure in a waste-free SMC sampler is expected to behave similarly to M independent Markov chains with the same target distribution. Therefore, it is perhaps not surprising that waste-free SMC estimators, which are computed via repeated combination of the output of these quasi-independent chains, might have a smaller asymptotic variance than standard SMC estimators, which are based on repeated combination of the output of a single chain. In the following section, we will investigate whether these favorable results regarding waste-free SMC hold in the context of light source detection, a regime where the standard SMC sampler introduced above has previously demonstrated strong performance.

3 Experiments

3.1 Variability of posterior estimates for individual images

In all three of their experiments, Dau and Chopin demonstrate that waste-free SMC produces estimates that are more consistent and less variable across various choices of M and P than standard SMC estimates are across choices of k. We aim to determine if these same trends hold for our standard and waste-free SMC samplers, which are tailored to light source detection. As such, our first experiment focuses on the variability of the posterior mean source count, posterior mean total flux, and estimated normalizing constant of four images with different source counts across many runs of standard and waste-free SMC. We generate four synthetic images of size 16 pixels by 16 pixels from the model in subsection 2.1. Figure 1 displays these images, which have source counts of two, four, six, and eight, respectively.

We consider five combinations of mutation kernel parameters for each of the two samplers. These parameters are listed in Table 1. Following Dau and Chopin, we selected these parameters to ensure that each sampler was allowed the same number of calls to the likelihood function in each SMC iteration. This amounts to ensuring that $kN{=}MP$ for any choices of $\{k, N\}$ and $\{M, P\}$, as the standard and waste-free samplers require skN and sMP likelihood evaluations, respectively, per SMC iteration for each block of catalogs. For each of these parameter combinations and each of the four images, we run the standard and waste-free samplers 100 times each. We implemented this experiment (and the experiment

Table 1: Mutation kernel parameters for our SMC samplers in the experiments in subsection 3.1

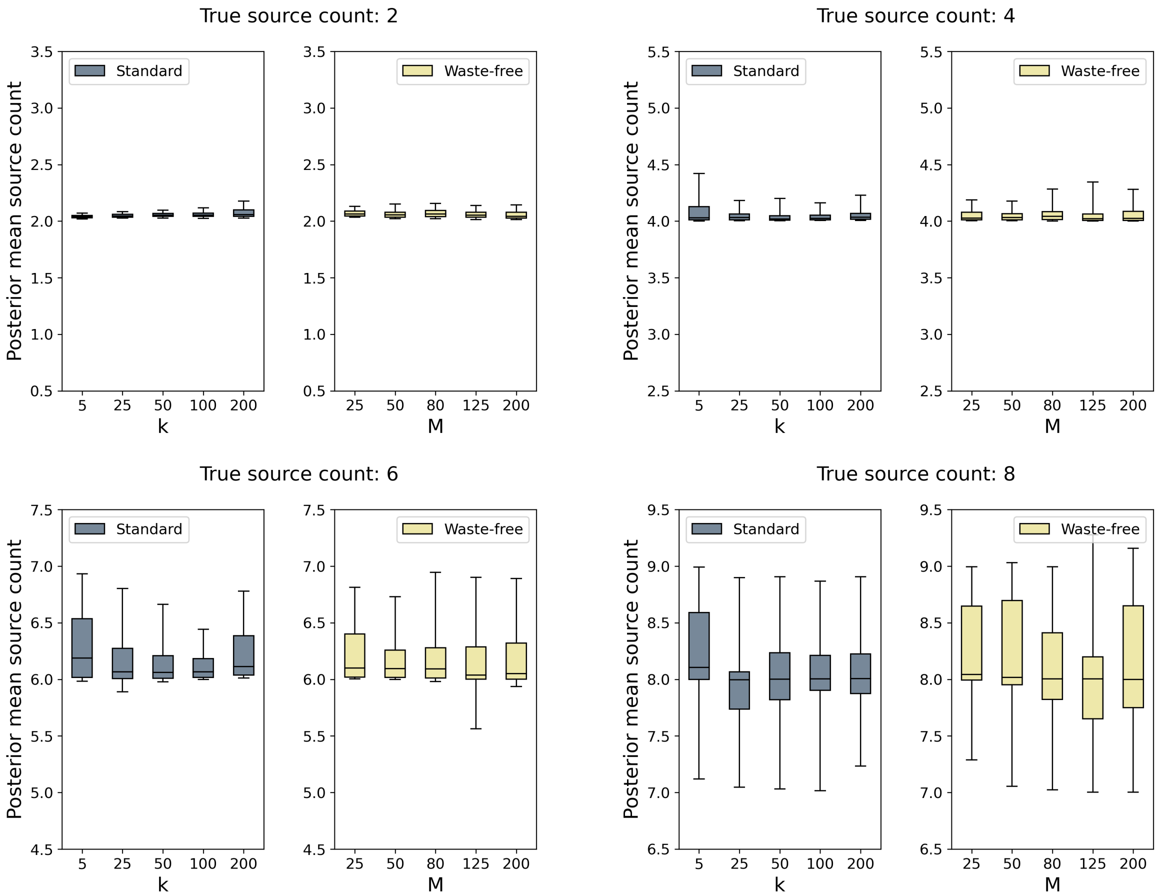
(a) Standard

Number of iterations k	Catalogs per block N
5	2000
25	400
50	200
100	100
200	50

(b) Waste-free

Resampled ancestors M	Number of iterations P
25	400
50	200
80	125
125	80
200	50

Figure 2: Posterior mean source count estimated in 100 runs of SMC for four images
(Whiskers cover the middle 90% of runs, extending to the 5th and 95th quantiles)



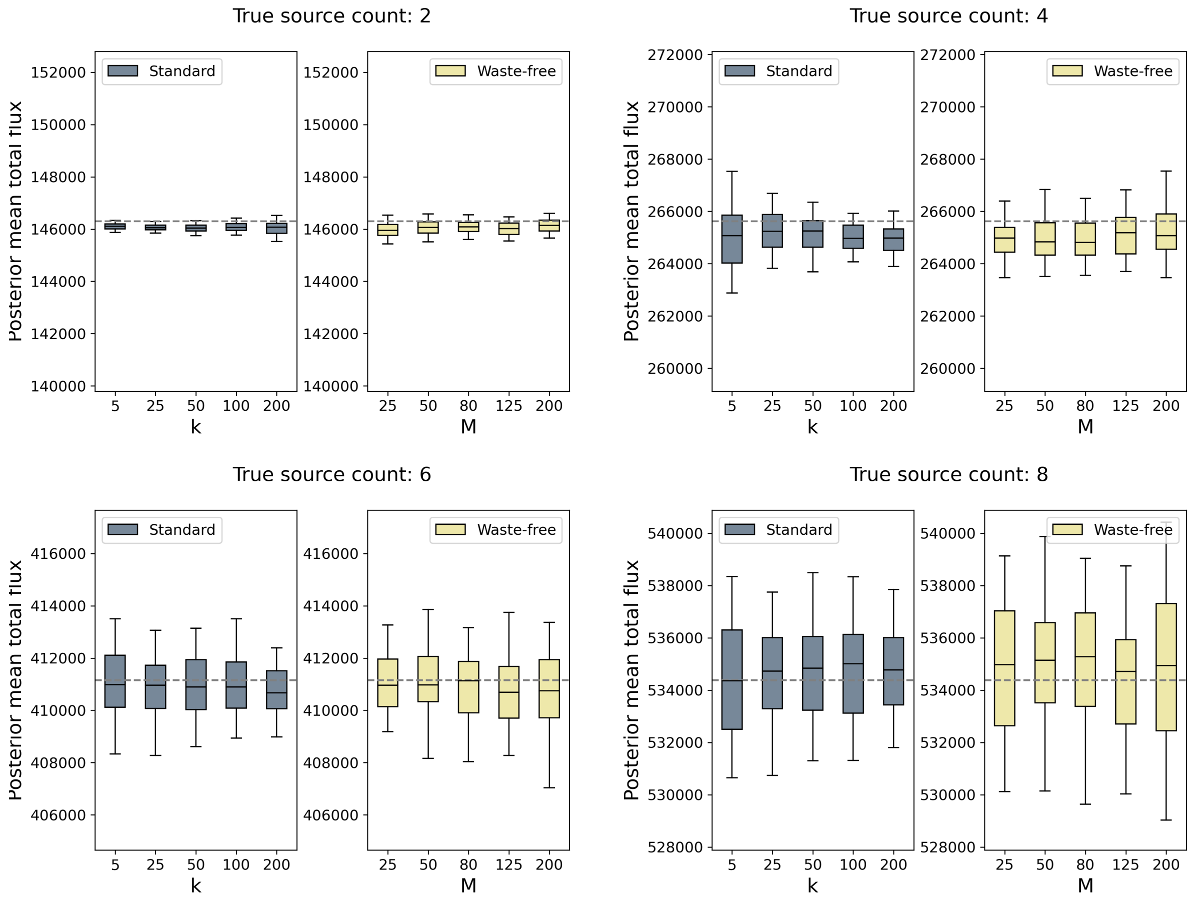
in subsection 3.2) in PyTorch on one NVIDIA GPU.[1]

Figure 2 displays the distribution of the posterior mean source counts across the 100 runs of the two samplers for each of the four images. Overall, the variability of the posterior mean source counts does not appear to be meaningfully impacted by the choice of the mutation kernel parameters, at least given the computational budget prescribed in Table 1. For the standard sampler, there is some indication that choosing too small or too large a k is suboptimal — variability is largest for k=200 for the image with two sources, while it is largest for k=5 for the image with four sources. However, the impact of a poor choice of k seems fairly minimal in this setting. We also find that the variability of the estimated source counts is similar in magnitude between the standard and waste-free samplers. For both methods, this variability is greater for images with higher source counts, which is not surprising since these images involve more overlap between sources and are thus more ambiguous.

Figure 3 and Figure 4 illustrate similar patterns for the posterior mean total flux and the log of the normalizing constant (i.e., $\log p(x)$), respectively. For the former, both samplers provide accurate estimates of the true observed total flux for all four images (as indicated by the gray dotted lines), and the

[1]Our code is available at `https://github.com/timwhite0/smc_object_detection/tree/wastefree`.

Figure 3: Posterior mean total flux estimated in 100 runs of SMC for four images
(Whiskers cover the middle 90% of runs, extending to the 5th and 95th quantiles)



variability of these estimates is larger for more crowded images. For the log of the normalizing constant, a quantity for which both samplers are known to provide unbiased and consistent estimates, we observe a similar trend. It is somewhat challenging to interpret these normalizing constants in our transdimensional setting, but they can essentially be viewed as a reflection of the marginal likelihood of the images under all configurations of source counts, fluxes, and locations considered by our prior. In both Figure 3 and Figure 4, the two methods appear to be robust to their respective mutation kernel parameters.

Ultimately, the results of our first experiment suggest that waste-free mutation does not meaningfully reduce the variability of posterior estimates in the context of our light source detection model. The asymptotic variance result proved by Dau and Chopin is not borne out in our simulations, likely because this a finite-sample setting where both samplers perform quite well.

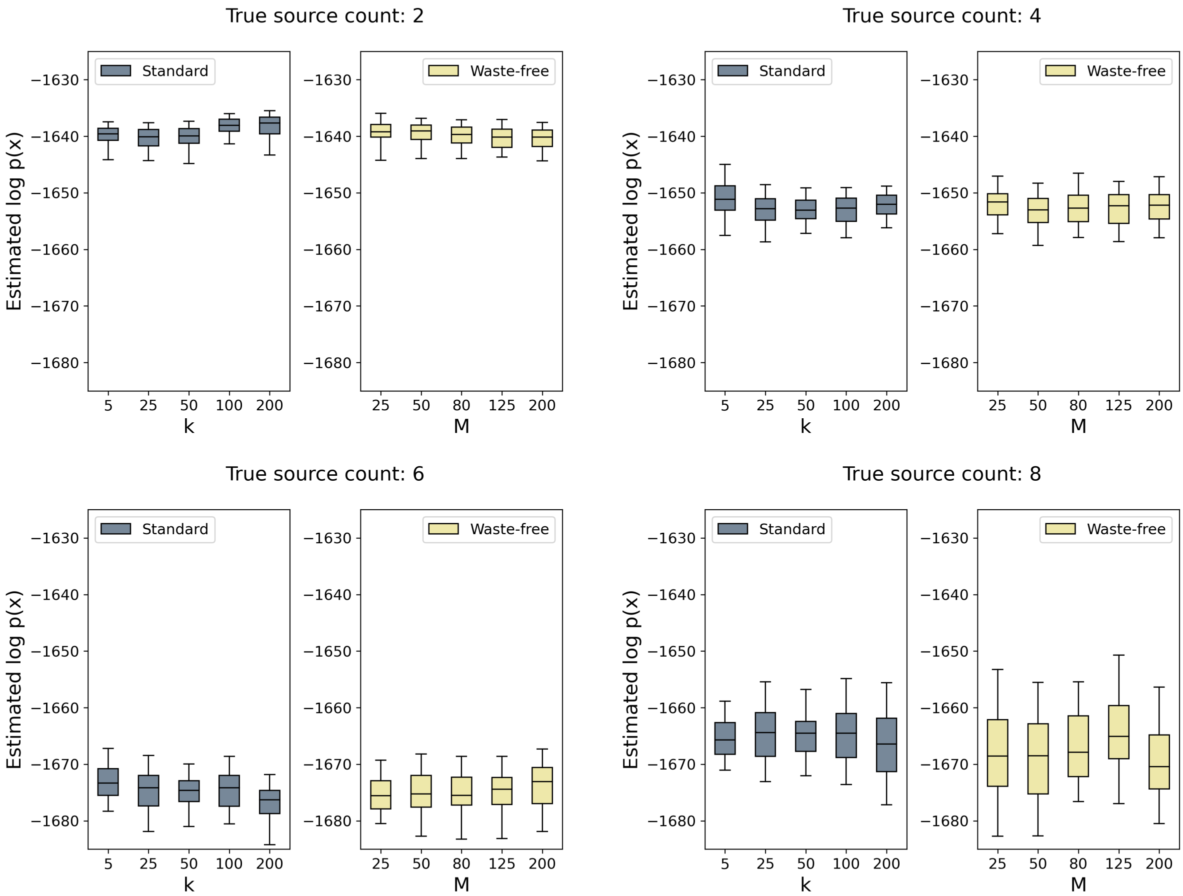
3.2 Accuracy and calibration of posterior estimates across many images

In our second experiment, we investigate the accuracy and calibration of the posterior mean source counts estimated by standard and waste-free SMC across a large collection of images. We generate 1,500 images of size 16 pixels by 16 pixels from the model in subsection 2.1 with true source counts ranging uniformly from zero to eight. We run our standard and waste-free samplers once for every image, and we fix their respective mutation kernel parameters at reasonable values based on the results of our first experiment ($\{k, N\} = \{100, 100\}$ for standard, $\{M, P\} = \{80, 125\}$ for waste-free).

Standard SMC yields a correct point estimate (i.e., posterior mean count rounded to the nearest integer) of the true source count in 1,246 of the 1,500 images (83.1%), with a mean absolute error of 0.197. Waste-free SMC achieves similar accuracy, yielding a correct point estimate in 1,212 of the 1,500 images (80.8%) with a mean absolute error of 0.217. Thus, both samplers are capable of estimating the true source counts of images from our model with relatively high accuracy, and they also capture the uncertainty surrounding

Figure 4: $\log p(x)$ estimated in 100 runs of SMC for four images

(Whiskers cover the middle 90% of runs, extending to the 5th and 95th quantiles)



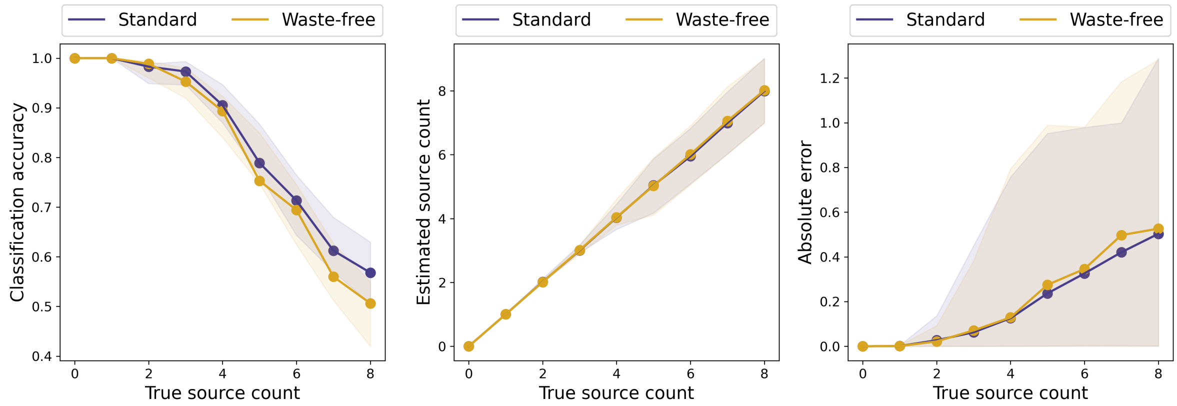
these estimates.

To illustrate the latter point, Figure 5 displays the accuracy, calibration, and absolute error of the standard and waste-free estimates among images with each true source count. For both samplers, classification accuracy (i.e., the proportion of images with a correct estimate among those with a particular source count) decreases as the true source count increases but remains above 50% even in crowded images containing eight stars. In such images, the absolute error of both samplers' posterior mean source counts is still only around 0.5, on average, and it ranges no higher than 1 for most images. The middle panel illustrates that both standard and waste-free SMC are well-calibrated in the sense that (i) their posterior mean source counts are close to the true source counts, on average, for $s \in \{0, 1, 2, ..., 8\}$ and (ii) these estimates exhibit greater variability in more ambiguous images. Similar to our first experiment, these results suggest that waste-free SMC does not hold a clear advantage over standard SMC for the task of detecting light sources in images.

4 Discussion

The results of the two experiments indicate that our waste-free SMC sampler achieves comparable performance to our standard SMC sampler in terms of the variability, accuracy, and calibration of its posterior estimates for several light source detection metrics. We find that the waste-free sampler is stable across several reasonable choices of its mutation kernel parameters M and P, as expected. However, we also find that the standard sampler is fairly robust to the choice of its mutation kernel parameter k, much more so than in the three numerical experiments conducted by Dau and Chopin. The performance of both samplers is quite strong for the class of models considered in this report, and we do not observe empirical evidence of waste-free SMC's purported advantage in robustness or variability.

Figure 5: Accuracy, calibration, and absolute error of posterior mean source counts across 1,500 images
(Error bands indicate middle 90% of bootstrap estimates (left) or middle 90% of images (center and right))



There are several limitations of our analysis that are worth mentioning, both regarding Dau and Chopin's work and the constraints we encountered in our transdimensional setting. First, we reiterate that we did not use the empirical covariance of catalogs from the previous SMC iteration to adapt the covariance of the Metropolis-Hastings proposals in the current SMC iteration. This is a common technique for many types of SMC samplers [1, 7]. However, it is not obvious how to implement it in settings like ours where the latent variables of interest represent sets of light sources; the index of each light source within a catalog is not necessarily fixed from one iteration to the next, and thus the covariances of the fluxes and locations between two indexes of a catalog are not directly interpretable. Our approach of fixing the proposal covariances for the fluxes and locations to a fraction of their respective prior variances is a reasonable alternative since the prior encodes information about the typical brightness of a star and the size of the image. We do not anticipate that this decision had a meaningful impact on our results.

Second, since our interest was to establish a baseline comparison between standard and waste-free SMC in the context of light source detection, we did not explore any of the methodological tweaks proposed by Dau and Chopin. Specifically, we did not consider their method of adapting the waste-free parameters M and P based on the autocorrelation of the chains in the mutation step. This extension would potentially be useful for our waste-free sampler. We observed in our experiments that the acceptance rates of the Metropolis-Hastings kernels tended to decrease as the number of SMC iterations increased, and Dau and Chopin recommended the adaptive approach for any problem where mixing deteriorates over time in this fashion. However, the empirical benefits of this approach would likely be negligible in the experiments considered above since our waste-free sampler already achieves strong performance.

Finally, recall that we determined the mutation kernel parameters for our experiments by allocating the same number of likelihood evaluations to the standard and waste-free samplers. This is the approach taken by Dau and Chopin in their experiments, and it amounts to equating the CPU budgets granted to the two samplers. However, we implemented our experiments on a GPU, and we found that the runtime of the waste-free sampler for a given image was nearly double that of the standard sampler, on average. This is not surprising — although we optimized the waste-free sampler by parallelizing the M chains in the mutation kernel, the objects being stored and manipulated are much larger for the waste-free sampler than the standard sampler (on the order of MP for waste-free and N for standard, using the notation from Table 1). As such, there appears to be no strong statistical or computational argument for using waste-free SMC instead of standard SMC for the light source detection task considered in this report.

Contributions

- Tim: Implemented the standard and waste-free SMC samplers; designed and ran the two experiments; created Algorithm 1 and Figure 1; wrote section 1, section 2, and section 4.
- Jaylin: Summarized and prepared the results from the two experiments; created Table 1, Figure 2, Figure 3, Figure 4, and Figure 5; wrote section 3 and section 4.

References

[1] Nicolas Chopin, Omiros Papaspiliopoulos, et al. *An Introduction to Sequential Monte Carlo*. Springer, 2020.

[2] Neil J Gordon, David J Salmond, and Adrian FM Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation". In: *IEE Proceedings F (Radar and Signal Processing)*. Vol. 140. IET. 1993, pp. 107–113.

[3] Arnaud Doucet, Simon Godsill, and Christophe Andrieu. "On sequential Monte Carlo sampling methods for Bayesian filtering". In: *Statistics and computing* 10 (2000), pp. 197–208.

[4] Nicolas Chopin. "A sequential particle filter method for static models". In: *Biometrika* 89.3 (2002), pp. 539–552.

[5] Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. "Sequential Monte Carlo samplers". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 68.3 (2006), pp. 411–436.

[6] Chenguang Dai et al. "An invitation to sequential Monte Carlo samplers". In: *Journal of the American Statistical Association* 117.539 (2022), pp. 1587–1600.

[7] Alexandros Beskos et al. "On the convergence of adaptive sequential Monte Carlo methods". In: *Annals of Applied Probability* 26.2 (2016), pp. 1111–1146.

[8] Christian A. Naesseth, Fredrik Lindsten, and Thomas B. Schön. "Elements of Sequential Monte Carlo". In: *Foundations and Trends in Machine Learning* 12.3 (2019), pp. 307–392.

[9] Alexander Buchholz, Nicolas Chopin, and Pierre E Jacob. "Adaptive tuning of Hamiltonian Monte Carlo within sequential Monte Carlo". In: *Bayesian Analysis* 16.3 (2021), pp. 745–771.

[10] Hai-Dang Dau and Nicolas Chopin. "Waste-free sequential Monte Carlo". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 84.1 (2022), pp. 114–148.

[11] Ajay Jasra et al. "Interacting sequential Monte Carlo samplers for trans-dimensional simulation". In: *Computational Statistics & Data Analysis* 52.4 (2008), pp. 1765–1791.

[12] Richard M. Feder et al. "Multiband probabilistic cataloging: a joint fitting approach to point-source detection and deblending". In: *The Astronomical Journal* 159.4 (2020), p. 163.

[13] Peter Melchior et al. "The challenge of blending in large sky surveys". In: *Nature Reviews Physics* 3.10 (2021), pp. 712–718.

[14] Tim White and Jeffrey Regier. "Sequential Monte Carlo for detecting and deblending objects in astronomical images". In: *Workshop on Machine Learning and the Physical Sciences, NeurIPS 2023* (2023). URL: https://ml4physicalsciences.github.io/2023/files/NeurIPS_ML4PS_2023_172.pdf.

[15] Peter J. Green. "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination". In: *Biometrika* 82.4 (1995), pp. 711–732.

[16] Juan Kuntz, Francesca R Crucinio, and Adam M Johansen. "Product-form estimators: exploiting independence to scale up Monte Carlo". In: *Statistics and Computing* 32.1 (2022), p. 12.