University of Amsterdam

Anton Pannekoek Institute

Master Thesis

ON THE SEARCH FOR GRAVITATIONAL-WAVE PROGENITORS: REDUCING THE COMPUTATIONAL COSTS OF BINARY POPULATION SYNTHESIS USING ADAPTIVE IMPORTANCE SAMPLING

FLOOR BROEKGAARDEN
10528105

Period:
JULY 2017 – JULY 2018

Degree:
MSC. ASTRONOMY & ASTROPHYSICS

Track:
ASTRONOMY & ASTROPHYSICS

Credits:
60 ECTS

Supervisor & Examiner:
DR. SELMA E. DE MINK, PROF ILYA MANDEL, PROF STEPHEN JUSTHAM

Second Reviewer:
PROF.DR. ALEX DE KOTER

Floor Broekgaarden: *On the search for Gravitational-Wave Progenitors: Reducing the Computational costs of Binary Population Synthesis using Adaptive Importance Sampling*, Master Thesis, July 2018

Preface

ABSTRACT

Abstract

Popular Summary

CONTENTS

I RESEARCH 1

1 INTRODUCTION 3
- 1.1 Gravitational-waves 3
- 1.2 Binary stellar evolution models 3
- 1.3 Common methods 3
- 1.4 rare events simulations 3
- 1.5 Motivation 5
- 1.6 Outline 6

2 STATISTICAL METHOD 7
- 2.1 Simulating populations of binaries 7
- 2.2 Initial binary parameters 8
- 2.3 Adaptive Importance Sampling algorithm 11
- 2.4 Adptive Importance Sampling 11
- 2.5 Fiducial method 11
- 2.6 Advanced Adaptive Importance Sampling 11
- 2.7 Other implementations 11
- 2.8 Population Synthesis 11
- 2.9 Fiducial Model 11

3 RESULTS 13
- 3.1 Adaptive Importance Sampling Toy Model 13
- 3.2 Binary population Blck hole - Neutron star mergers 13
- 3.3 Exploring Adaptive importance Sampling 13

4 DISCUSSION 15

5 CONCLUSION 17

II APPENDIX 19

A A BRIEF GUIDE TO ADAPTIVE IMPORTANCE SAMPLING 21

B CODE 23

LIST OF FIGURES

LIST OF TABLES

Table 1 Table of variables used 11

ACRONYMS

Part I

RESEARCH

1 INTRODUCTION

GRAVITATIONAL-WAVES

Currently, one of the leading ideas in the field is that binary mergers originate from two massive stars that are born in a binary system, evolve over millions of years and at the end of their lives die in energetic explosions leaving behind a neutron star or black hole that spiral in and eventually merge to produce the gravitational-waves that we observe today. By comparing statistical studies of gravitational-wave observations with simulated compact object merger populations from theoretical models, we can learn about the evolution of stars in binary systems and the underlying physical processes involved. However (i) the progenitor systems of gravitational waves are extremely rare (ii) the stellar evolution simulations include many uncertainties that need to be explored and (iii) computational resources are scarce - making large inference studies of compact object mergers computationally intractable. This is currently a main challenge in this field limiting more detailed simulations.

Gravitational-wave progenitors

BINARY STELLAR EVOLUTION MODELS

detailed stellar models

binary population synthesis

COMMON METHODS

RARE EVENTS SIMULATIONS

Naive simulation becomes inefficient as the rare event probabilities gets smaller. Days or weeks are required to obtain an accurate estimate. For that reason, we need special techniques to speed up the estimation process. These special techniques for the simulation of rare events can be collected under two main categories: importance sampling and splitting . Both categories modify the simulation so that the rare event of interest occurs more frequently than in Monte Carlo simulation. It is important to note that the mathematical influence of these modifications has to be compensated to obtain the true probability. However, these two categories differ in the type of modifications.

In importance sampling the underlying probability measure is transformed to push the sample paths towards the rare event. In splitting, the underlying probability measure stays the same. Instead, a selection mechanism is used to pick the sample paths that are likely to reach rare event. Then, the chosen sample paths are split or cloned into multiple copies. This results in an artificial drift towards the rare event. A general discussion on both methods is provided below.

Importance sampling (IS) is a powerful Monte Carlo simulation variance-reduction technique that has achieved success in simulating many types of rare event problems. The generic idea of IS is to modify the probability law of the underlying system to sample the important events more frequently. This new probability measure is called the change of measure or IS distribution. Simulating the system under the IS distribution would naturally result in a biased estimator unless a correction is applied. This translation of the outcome from IS distribution to the original probability distribution is done by the likelihood ratio. The likelihood ratio is associated either with a single outcome or a sample path – sequence of outcomes. The likelihood ratio of a sample path can simply be defined as the probability of the sample path under the original measure over the probability of the same sample path under the IS measure. If the original probability laws are known, we can trace the sample path through time to calculate the probability of the sample path under the original probability measure. In this thesis, we are interested in the problems where the original probability laws are known. For more details on IS , see the recent surveys by Bucklew [21], Juneja and Shahahbuddin [68] or Rubino and Tuffin [96]. A crucial problem in IS simulation is to identify a proper choice of the change of measure. The IS distribution should be chosen such that the target event is no longer rare, and thus will be observed more frequently. However, one should be also careful about "too much occurrence of the rare event". Although the IS estimator is proved to be unbiased, it may overestimate with a high probability. Hence, if chosen incorrectly, the resulting estimator may have a greater variance than the one from Monte Carlo simulation. The variance can even become infinite. In that case, the estimator may also give biased results, even when it is unbiased in a theoretical sense. Mathematically, it is possible to pinpoint the optimal change of measure. To obtain zero-variance, every sample path generated under the IS measure should hit the rare event. This is only possible if the new distribution is chosen as the original distribution conditioned on the occurrence of the rare event. Although theoretically the optimal change of measure is known, it is practically useless since it involves a-priori knowledge of the probability of interest.

A common way to launch IS schemes for state-dependent changes of measures is to use adaptive IS techniques that attempt to learn the zero-variance change of measure by an iterative procedure. In every iteration a number of sample paths is generated, and based on the 'scores' of these paths the current change of measure is updated. For more information on adaptive IS techniques, we refer to Desai [34] and Kollman [74]. One of these adaptive methods is called the cross-entropy method which will be discussed briefly in the following section

MOTIVATION

OUTLINE

2 STATISTICAL METHOD

In this chapter we will discuss the statistical method that we developed forbinary Population synthesis. Focusing on the difference between our method and traditional sampling methods in BS.

SIMULATING POPULATIONS OF BINARIES

In binary population synthesis, one can simulate the evolution of a synthetic population of binary systems. For such simulations, bianries are randomly drawn from the distributions of initial binary parameters and evaluating it through binary evolution prescriptions. By doing so, they present a rapid code that can compute the evolution of many stars within a simulation. Since it is now known a priori which initial conditions will produce an event of interest, the full initial parameter space needs to be explored. Traditional methods in BPS tackle this problem by randomly drawing binaries $\mathbf{x_i} \sim P(\mathbf{x_i})$ from their prior distributions which are based on observations, and evaluating them with the BPS model, u, into their final state $\mathbf{x_f}$,

$$\mathbf{x_f} = u(\mathbf{x_i}). \tag{1}$$

Many parameters are used in BPS but most binaries can be described uniquely by a few important initial vari- ables: the initial mass of the primary star (i.e., the most massive star) $M_{1,i}$, the mass ratio, $q_i = M_{1,i}/M_{2,i}$, between the two stars and the initial separation, a_i, eccentricity e_i. Depending on the binary evolution we can also add the kick velocity received when the primary or secondary collapses to a compact object $\mathbf{v}_{\text{kick, SN}}$. Each initial binary sample can thus be represented by

$$\mathbf{x_i} = (M_{1,i}, a_i, q_i, e_i, \mathbf{v}_{\text{kick, SN1}}, \mathbf{v}_{\text{kick, SN2}}) \tag{2}$$

Often, BPS is used to study binaries that evolve to a certain subtype $\mathbf{X_t}$,

So in such a study the goal is to model the distribution:

$$\psi(\mathbf{x_f}) = \begin{cases} 1 & \text{if } \mathbf{x_f} \in \mathbf{X_t} \\ 0 & \text{else} \end{cases} \tag{3}$$

that equals unity if $\mathbf{x_i}$ evaluated to the target binary system $\mathbf{X_t}$ and zero if not. We will use this function throughout the paper as it is the main objective to perform the inference on. Say something about this function.

However, in cases when the target population is a rare event in the simulation, e.g. when simulating Compact object binaries since most systems will disrupt during the supernova kick, and thus $\psi(\mathbf{x_f}) = 0$ for most systems. Therefore simulating populations of such rare events becomes extremely computational expensive.

Therefore, instead, present the variance reduction method , adaptive importance sampling [1], that generates samples from a distribution function which is automatically adapted to the scientific target by focusing on areas of the parameter space found to produce events of interest. Instead of drawing random binaries from the prior distribution P we draw them from a distribution that focuses around the target distribution of interest, thereby minimizing the computational costs spend on areas that don't produce events of interest, whilst maximizing the computational costs spend on binaries that become the rare event.

The method consists of three main steps that are also shown in Figure

1. The parameter space is explored to find a small population of events of interest.
2. The set of known events is used to build an instrumental distribution, to adaptively guide future sampling of the parameter space.
3. Later simulations of population members are then drawn from this instrumental distribution. The information from these simulations can in turn be used to further improve the instrumental distribution.

By doing so, the method minimizes the computational time spend on binary systems that don't evolve to the event of interest, whilst maximizing the computational time spend on binaries of the target distribution.

INITIAL BINARY PARAMETERS

For the exploratory phase and comparison runs we choose M_1, q, a, e, v_k similar to the distributions used in common binary population synthesis models (e.g. Belczynski, Pols)

Initial mass

The distribution of the initial primary mass M1 follows a power law distribution also known as the initial mass func- tion (IMF) (Kroupa 2001):

$$p(M_{1,i}) = C_M M_{1,i}^{-\alpha} \text{for} M_{1,i} \in [M_{1,i,min}, M_{1,i,max}] \tag{4}$$

where C_M is the normalization constant given by

$$C_M = \frac{\alpha + 1}{M_{1,i,max^{\alpha+1}} - M_{1,i,min}^{\alpha+1}}. \tag{5}$$

We choose $\alpha = 2.35$ in agreement with (Salpeter or Kroupa), and $M_1 \in [5, 100] M_\odot$ to align with Vigna-Gomez (2017).

initial mass ratio

The mass ratio q_i is suggested from observations to have a flat distribution (Mazeh et al., 1992; Goldberg & Mazeh, 1994; Tout, 1991), given by

$$p(q) = \frac{1}{q_{max} - q_{min}}, \text{for } q \in [q_{min}, q_{max}], \tag{6}$$

where $[q_{min}, q_{max}] = (0, 1]$ by definition of q. Nevertheless, it is also suggested that there is some dependency of the mass ratio on the period of the system (e.g. Moe & Di Stefano 2016). But this is beyond the scope of this thesis.

initial separation

$$p(a_i) = C_a \frac{1}{a} \text{for } a \in [a_{min}, a_{max}] \tag{7}$$

where C_a is the normalization constant

$$C_a = \frac{1}{\log a_{max} - \log a_{min}}. \tag{8}$$

We choose $[a_{min}, a_{max}] = [0.1, 1000]$ AU consistent with (Vigna-Gomez 2018).

eccentricity

We assume that all of our binaries initially are orbiting in a circular orbit, $e = 0$ consistent with (Alejandro et al)

Supernova

We differentiate between three supernova scenarios: core col- lapse supernovae (CCSN), ultra-stripped supernova (USSN) and electron-capture supernova (ECSN). For the CCSN treatment, we use the rapid explosion scenario, as presented in Fryer et al. (2012), to determine the compact object remnant mass according to the total and carbon-oxygen (CO) core mass of the progenitor, with a maximum allowed NS mass of mNS,max = 2.0 M?. In this scenario, the collapse does not allow for accretion onto the proto-NS, and is able to reproduce the proposed mass gap between neutron stars and black holes (Ozel et al. 2010; Farr et al. 2011). There is no consensus yet whether the mass gap is due to observational selection effects or if it is intrinsic to the explosion mechanism (Kreidberg et al. 2012; Wyrzykowski et al. 2016). Another explosion scenario comes from USSN (Tauris et al. 2013, 2015). A star becomes stripped when it loses its hydrogen envelope during its evolution; if, during later stages, it manages to lose its helium envelope, it becomes ultra-stripped. In COMPAS, any star which engages in a stable case BB mass transfer episode with a NS as an accre- tor, is considered to be ultra-stripped.We define case BB as a mass transfer episode which involves a Helium donor star which has stopped burning helium in the core (naked helium star Hertzprung Gap, HeHG). Ultra-stripped stars are left with an ONeMg core with a thin carbon and helium layer (Tauris et al. 2013). The compact object remnant mass of an USSN is determined in the same way as for CCSN. A single star with 8 mZAMS / M? 10 (binary stars spread the initial mass range) may collapse in an ECSN (Nomoto 1984). We assume the baryonic mass of the de- generate ONeMg core leading to an ECSN is 1.3

Table 1: Table of variables used

variable	description
u	BPS model
$\mathbf{x_i}$	initial state of a binary system
$\mathbf{x_f}$	final state of a binary system
$\mathbf{X_t}$	target subpopulation of binaries of interest
M_1	mass of the primary star
a	initial separation of the binary
q	initial mass ratio $q = M_2/M_1$ of the binary
e	initial eccentricity
v	
v_k	kick velocity amplitude
θ_k	
ϕ_k	

ADAPTIVE IMPORTANCE SAMPLING ALGORITHM

ADPTIVE IMPORTANCE SAMPLING

FIDUCIAL METHOD

Exploratory phase

Improve distribution

Run simulation (step 3)

ADVANCED ADAPTIVE IMPORTANCE SAMPLING

OTHER IMPLEMENTATIONS

POPULATION SYNTHESIS

For the population synthesis model we use COMPAS (Alejandro et al in prep and Stevenson et al. 2017)

based on (...)

FIDUCIAL MODEL

Our fiducial model

SN

mass transfer and stabillity

CE treatment

fall back

kicks

3 RESULTS

ADAPTIVE IMPORTANCE SAMPLING TOY MODEL

BINARY POPULATION BLCK HOLE - NEUTRON STAR MERGERS

EXPLORING ADAPTIVE IMPORTANCE SAMPLING

4 DISCUSSION

5 CONCLUSION

Part II

APPENDIX

a

A BRIEF GUIDE TO ADAPTIVE IMPORTANCE SAMPLING

b

CODE