STROOPWAFEL: a Dutch cookie and an Adaptive Importance Sampling algorithm

Floor Broekgaarden

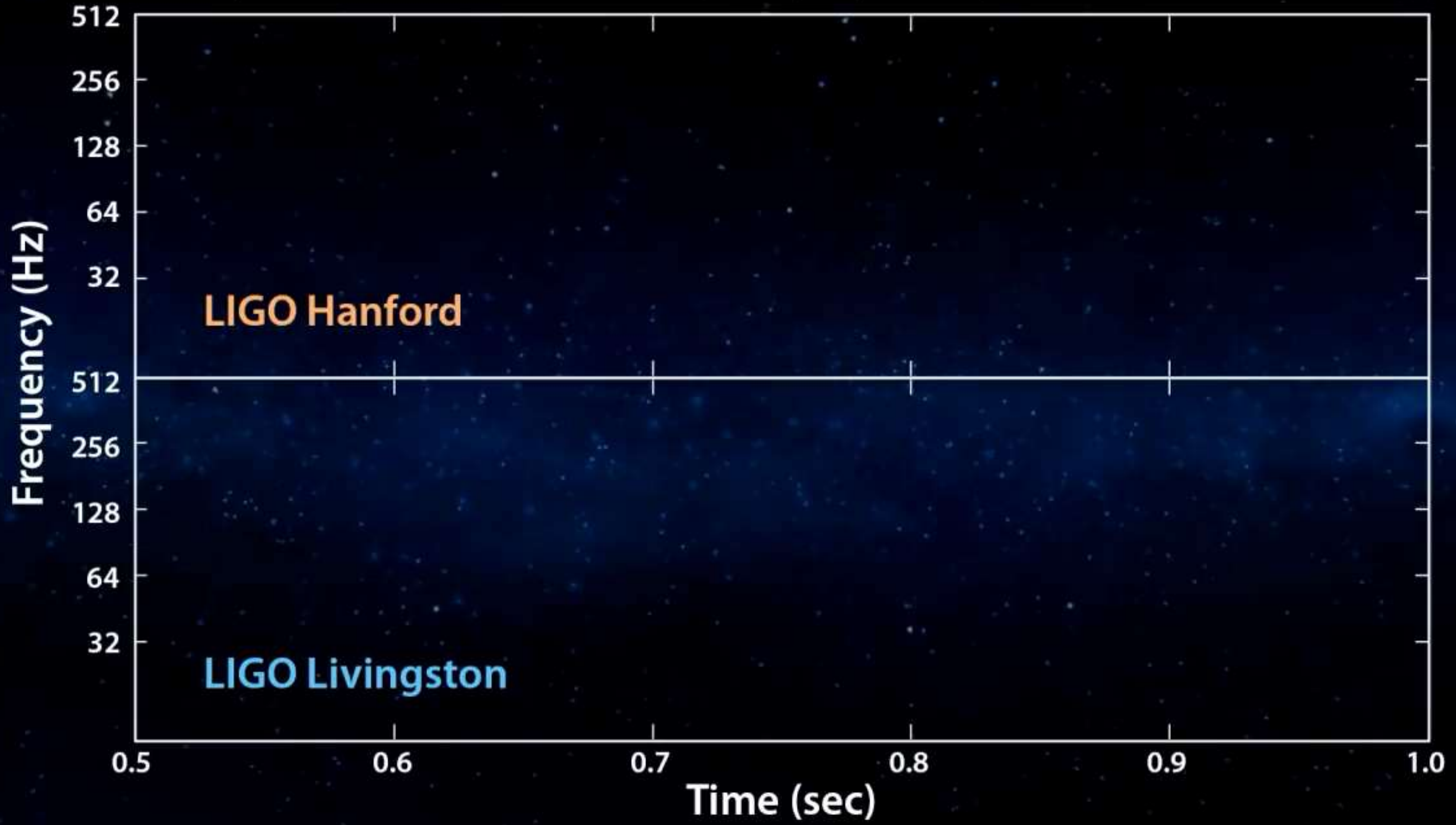
Center for Astrophysics | Harvard & Smithsonian

In collaboration with:

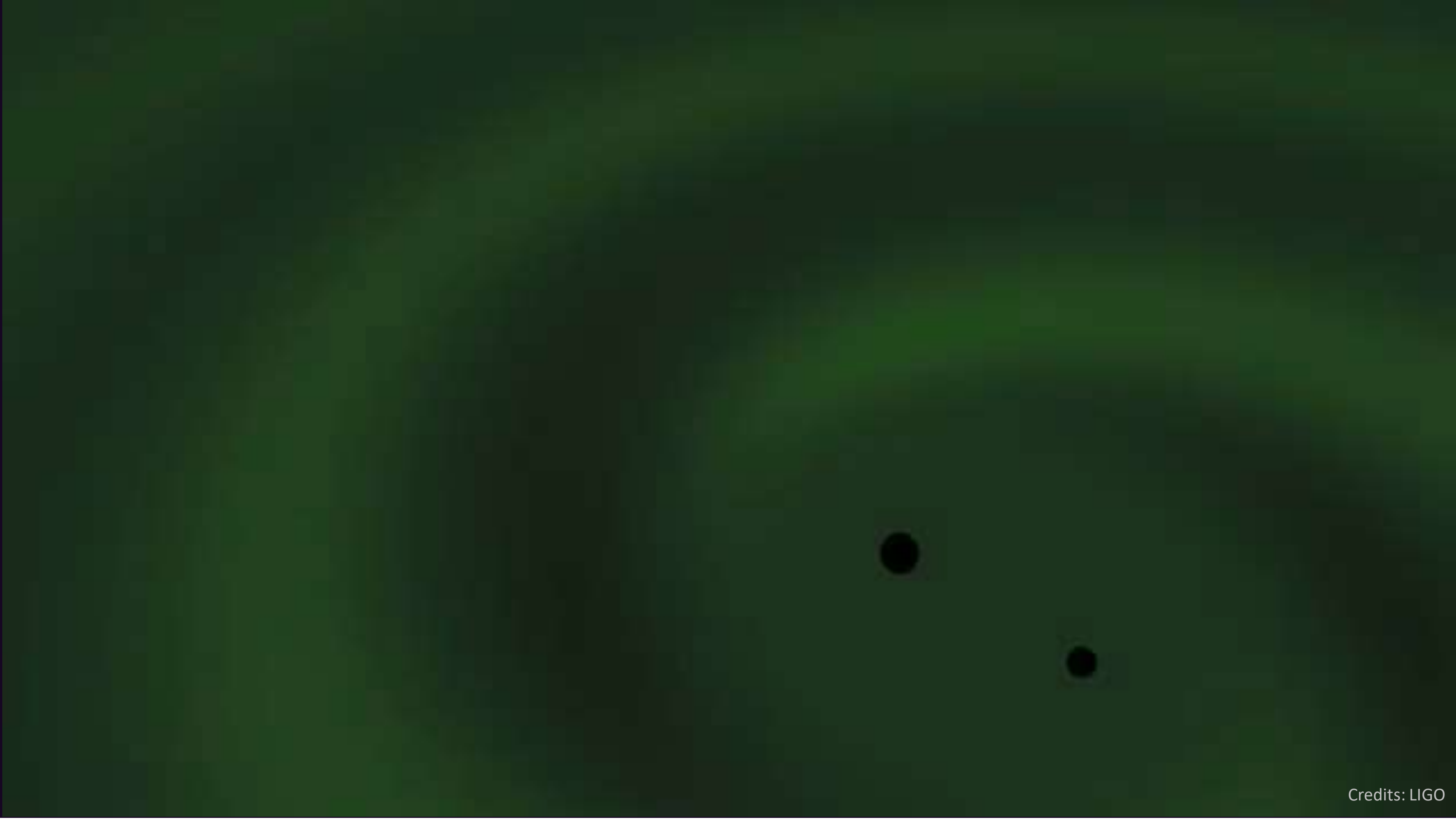
Jonathan Gair, Floris Kummer, Lokesh Khandelwal, Stephen Justham, Ilya Mandel, Selma de Mink, Luyau Lin, Coen Neijssel, Alejandro Vigna-Gomez, Simon Stevenson, Tom Wagg, Lieke van Son, Michelle Wassink, Poojan Agrawal, Serena Vinciguerra, Reinhold Willcox, Jeff Riley, Mike Lau and many others…

ICERM meeting | 20 November 2020

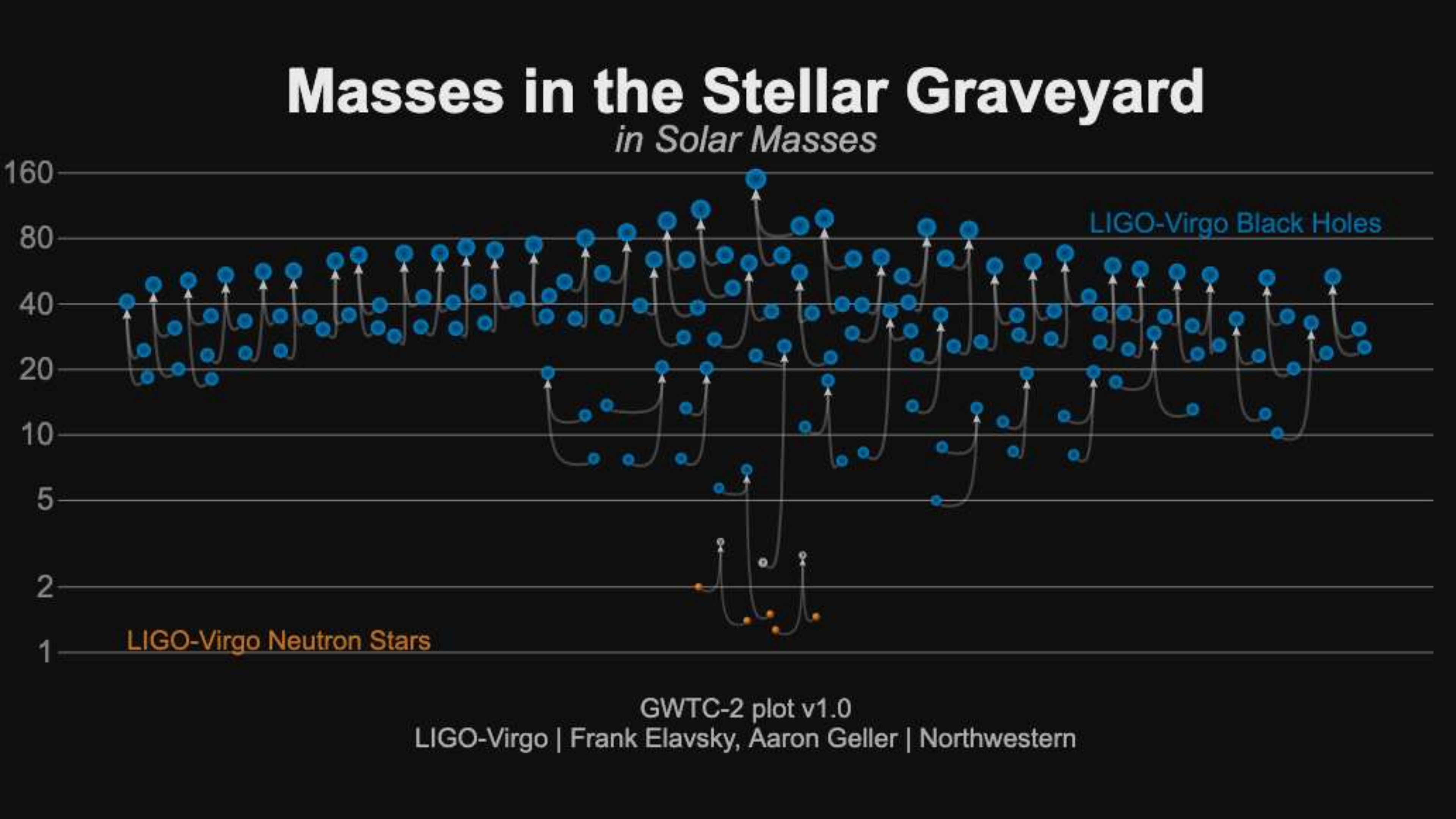
Statistical Methods for the Detection, Classification, and Inference of Relativistic Objects



Credits: LIGO

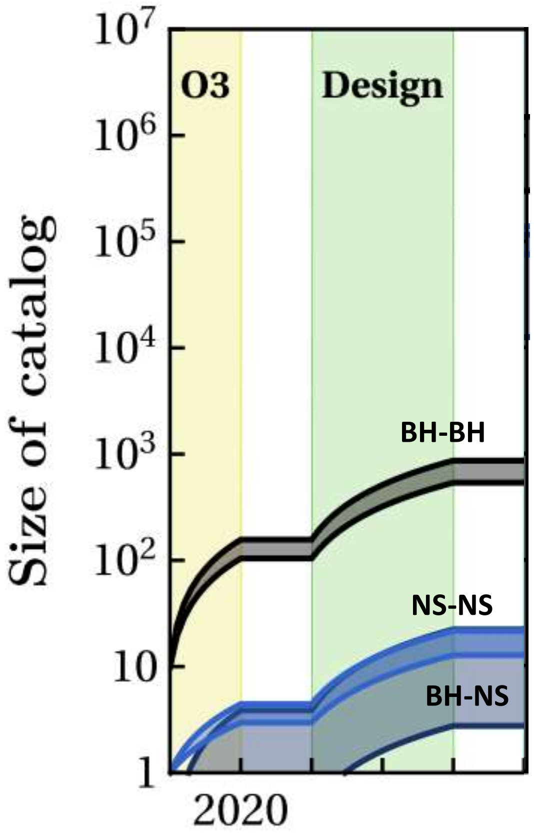


Credits: LIGO



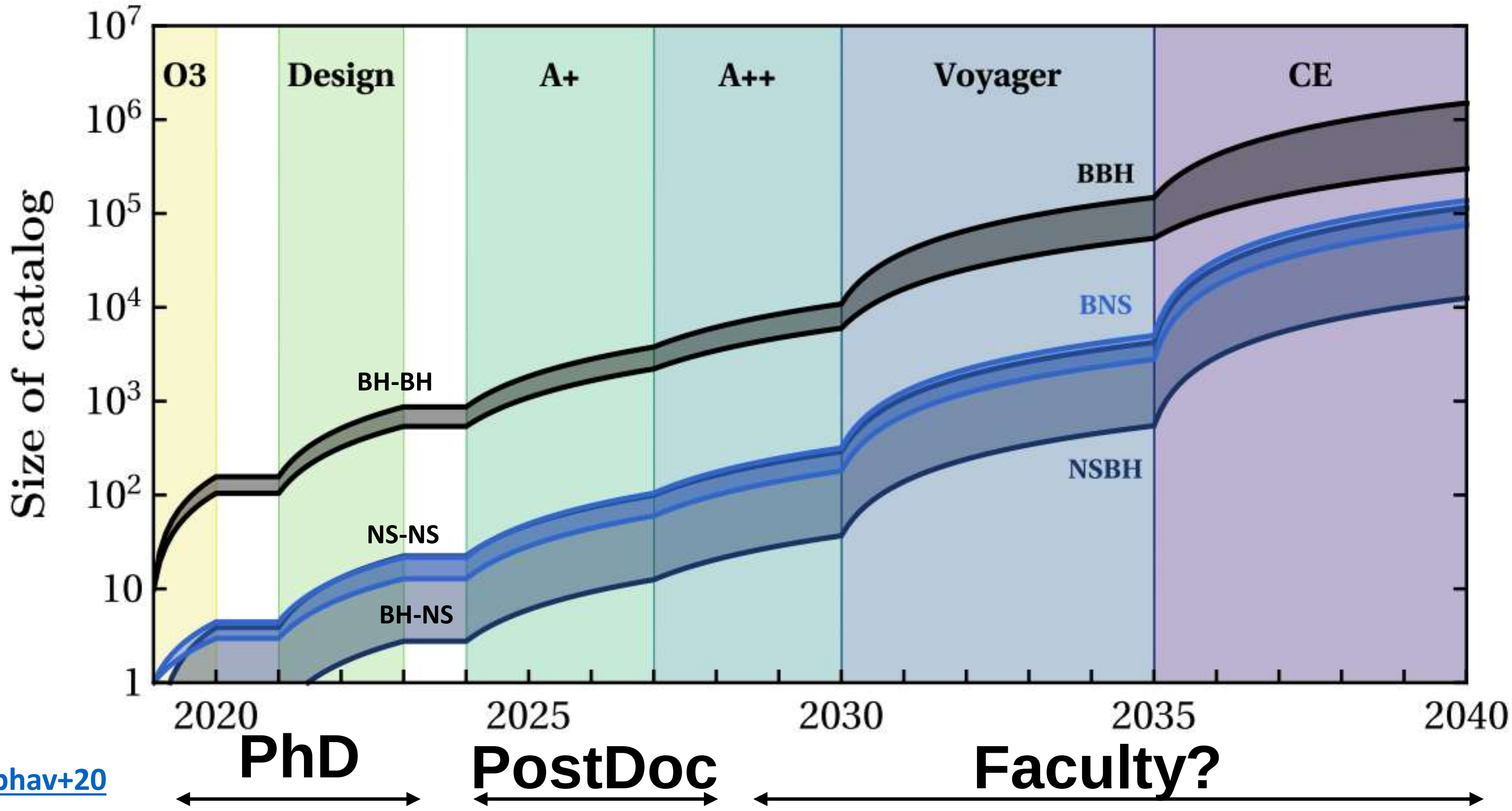
Masses in the Stellar Graveyard
in Solar Masses
160
80
40
20
10
5
2
1
LIGO-Virgo Black Holes
LIGO-Virgo Neutron Stars
GWTC-2 plot v1.0
LIGO-Virgo | Frank Elavsky, Aaron Geller | Northwestern

Building up a population of observations



Baibhav+20

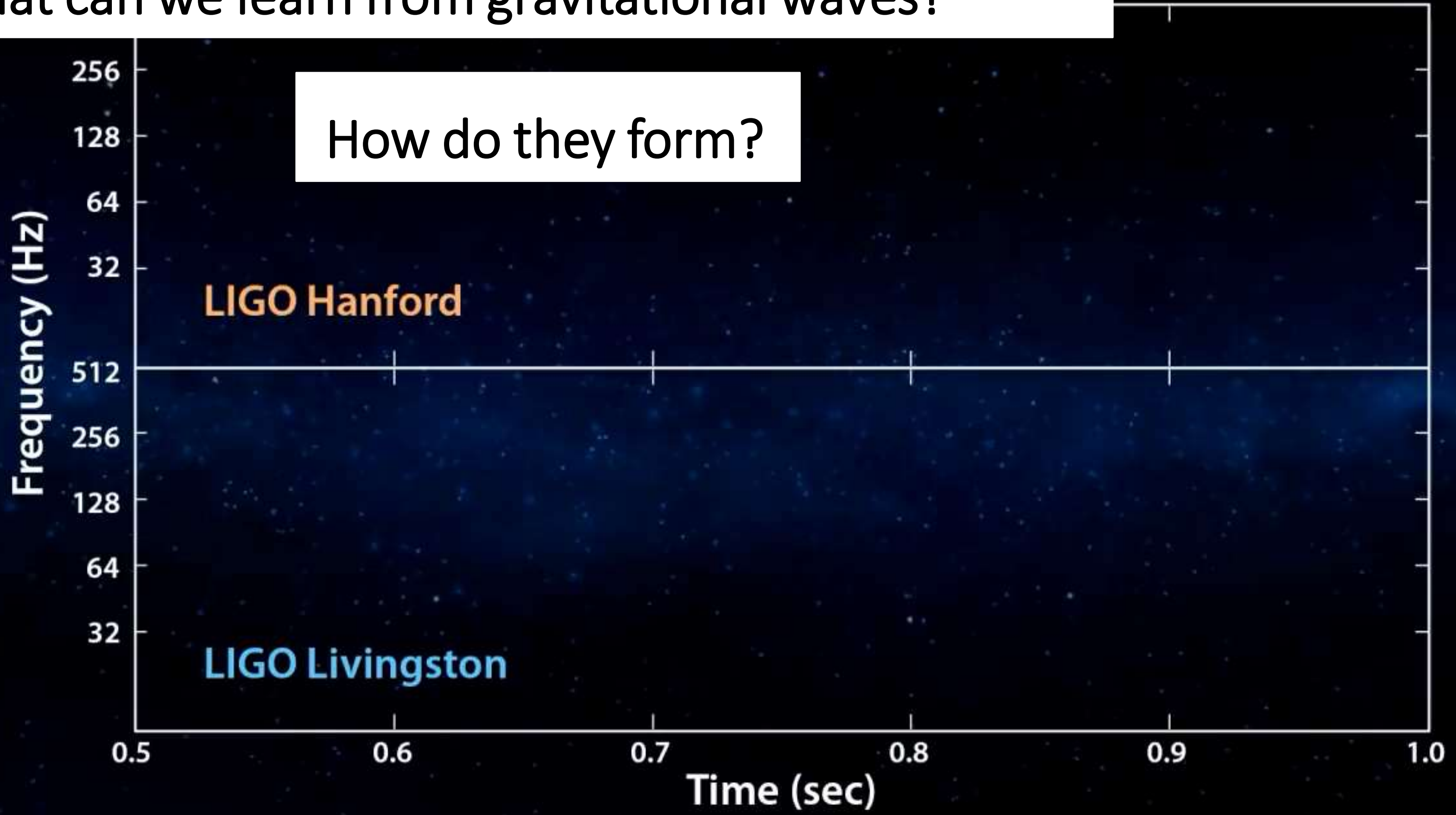
Building up a population of observations



Baibhav+20

What can we learn from gravitational waves?

How do they form?



Credits: LIGO

How do they form?

$z = 0$

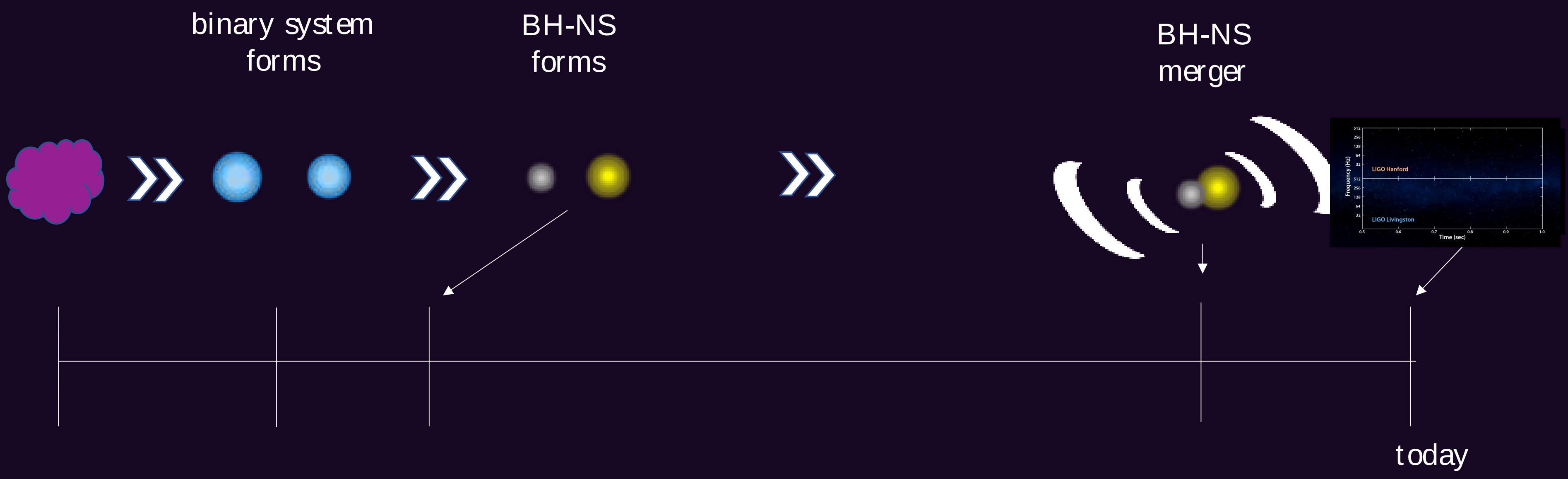


Image not to scale!

e.g. Paczynski+76, Smarr & Blandford+76

Rapid binary population synthesis

Stevenson+17, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

Small images: Tauris+17

$$\mathbf{x}_\mathrm{i} = (m_{1,\mathrm{i}}, m_{2,\mathrm{i}}, a_\mathrm{i}, \ldots)$$

Initial conditions: mass 1, mass 2, separation, …

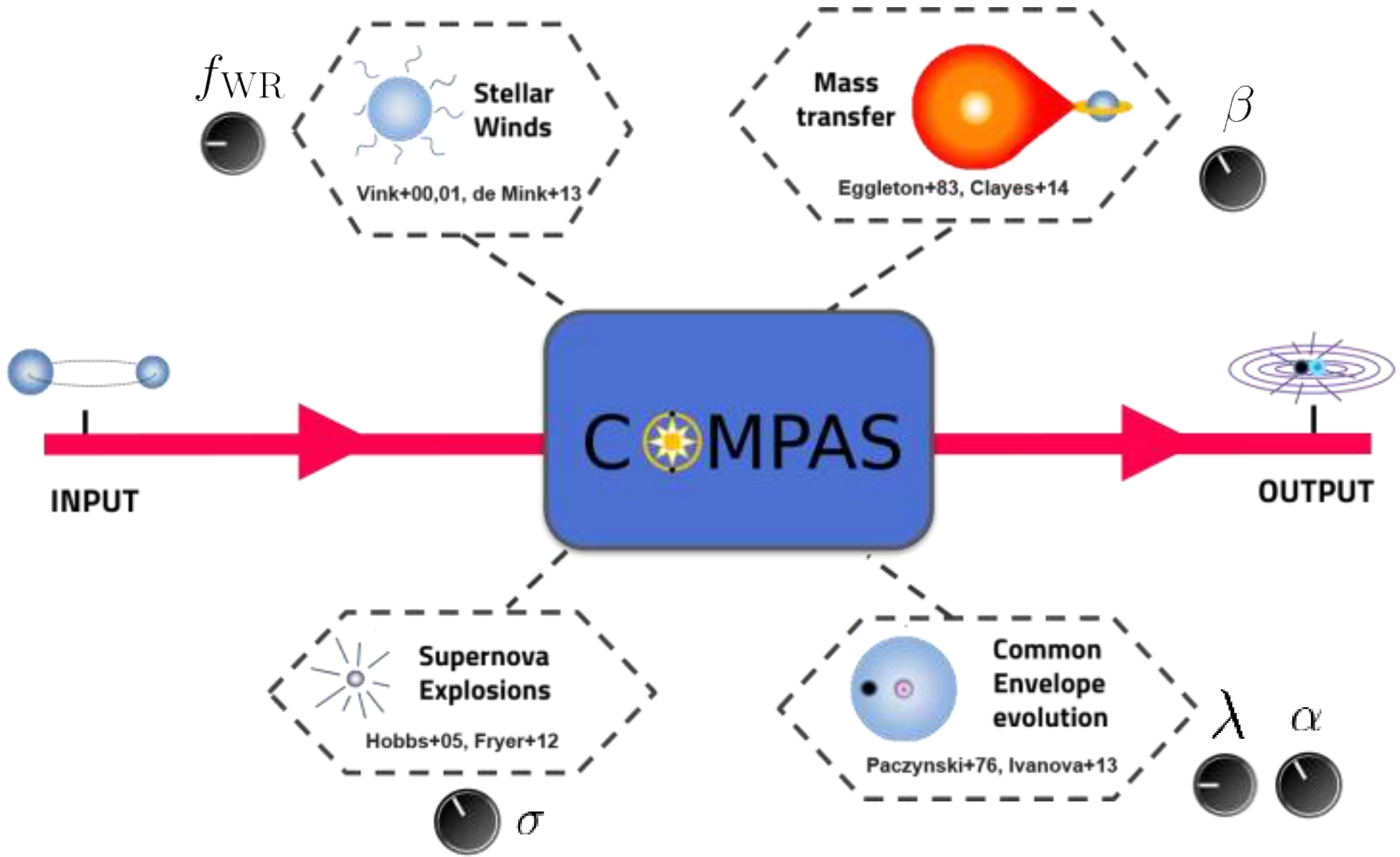
$$\mathbf{u}(\mathbf{x}_\mathrm{i}) = (m_{1,\mathrm{f}}, m_{2,\mathrm{f}}, \ldots)$$

final conditions: mass 1, mass 2, …

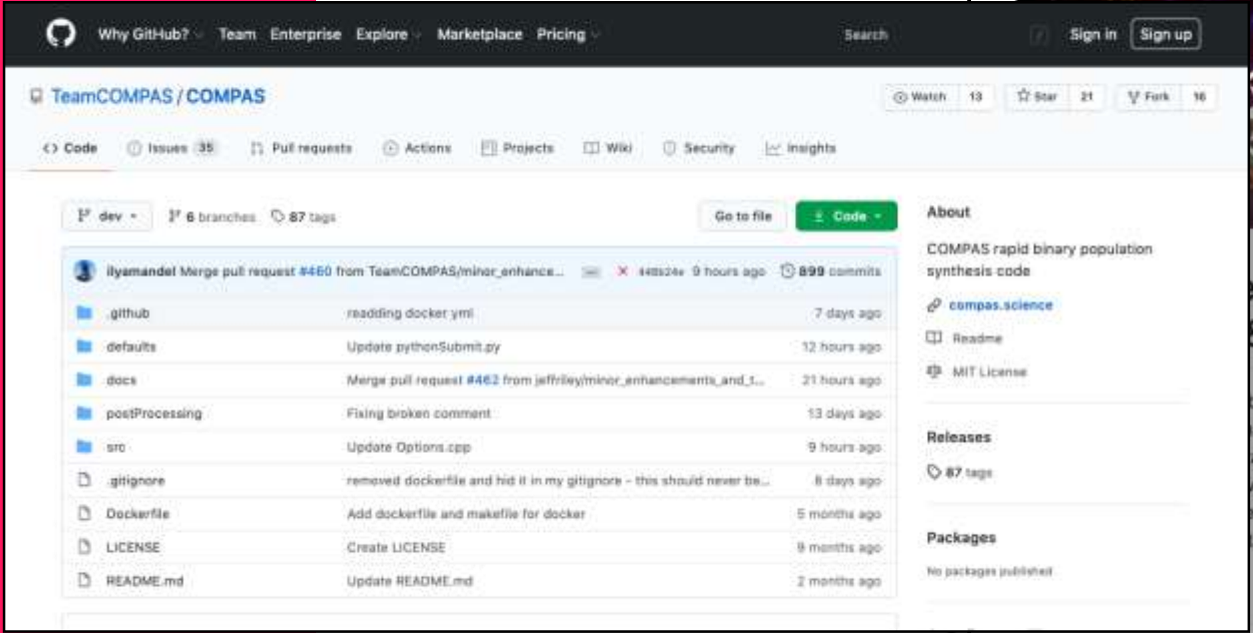
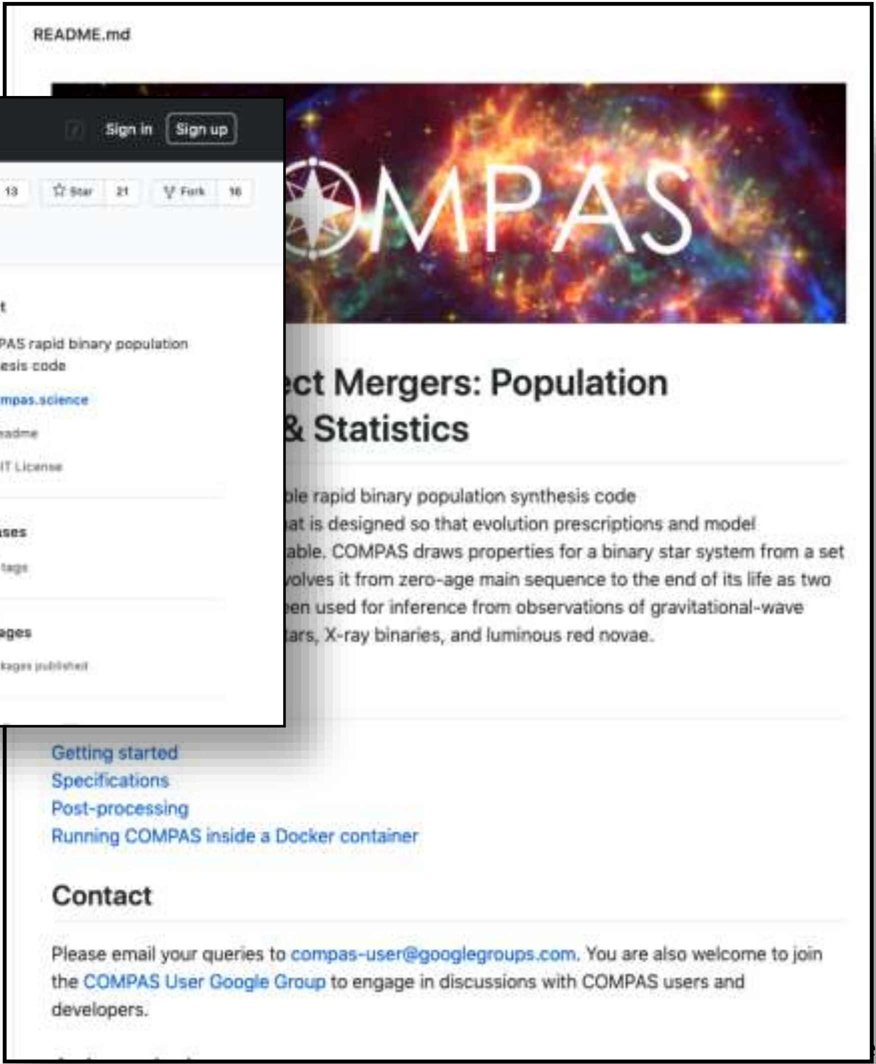
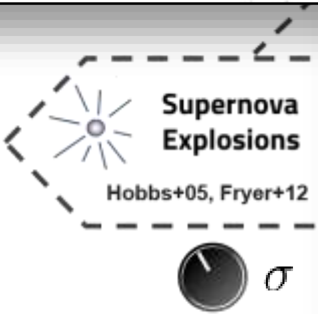
Rapid binary population synthesis

Stevenson+17, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

Small images: Tauris+17



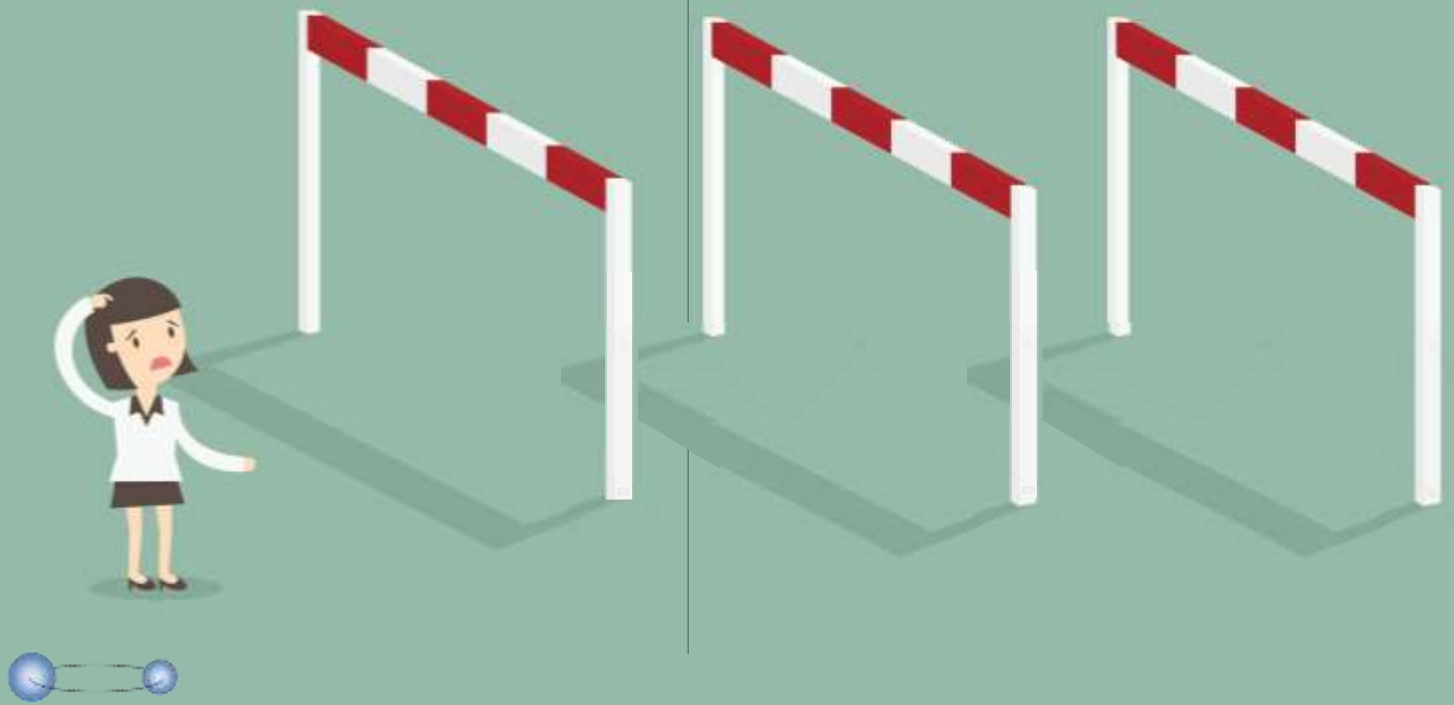
Floor Broekgaarden


Why GitHub? Team Enterprise Explore Marketplace Pricing
Search
Sign in
Sign up
TeamCOMPAS / COMPAS
Watch 13
Star 21
Fork 16
Code
Issues 35
Pull requests
Actions
Projects
Wiki
Security
Insights
dev
6 branches
87 tags
Go to file
Code
ilyamandel Merge pull request #460 from TeamCOMPAS/minor_enhance...
9 hours ago
899 commits
.github
readding docker yml
7 days ago
defaults
Update pythonSubmit.py
12 hours ago
docs
Merge pull request #462 from jeffriley/minor_enhancements_and_f...
21 hours ago
postProcessing
Fixing broken comment
13 days ago
src
Update Options.cpp
9 hours ago
.gitignore
removed dockerfile and hid it in my gitignore - this should never be...
8 days ago
Dockerfile
Add dockerfile and makefile for docker
5 months ago
LICENSE
Create LICENSE
9 months ago
README.md
Update README.md
2 months ago
About
COMPAS rapid binary population synthesis code
compas.science
Readme
MIT License
Releases
87 tags
Packages
No packages published

README.md
COMPAS
ct Mergers: Population
& Statistics
ble rapid binary population synthesis code
at is designed so that evolution prescriptions and model
able. COMPAS draws properties for a binary star system from a set
volves it from zero-age main sequence to the end of its life as two
en used for inference from observations of gravitational-wave
ars, X-ray binaries, and luminous red novae.
Getting started
Specifications
Post-processing
Running COMPAS inside a Docker container
Contact
Please email your queries to compas-user@googlegroups.com. You are also welcome to join the COMPAS User Google Group to engage in discussions with COMPAS users and developers.

Supernova
Explosions
Hobbs+05, Fryer+12
σ

cartoon from DooderVector/everypixel

3 computational challenges

1 Simulating populations



Cost:

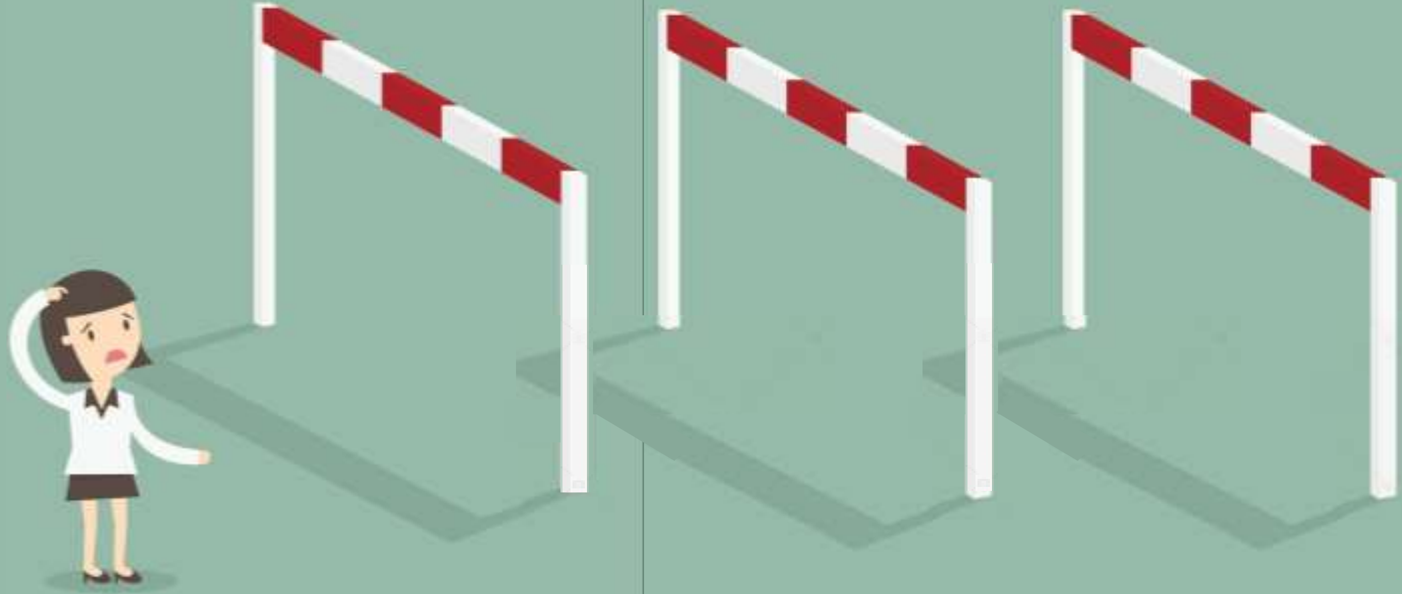
~1 sec

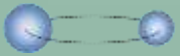
Floor Broekgaarden

3 computational challenges

1 Simulating populations

2 Testing physics

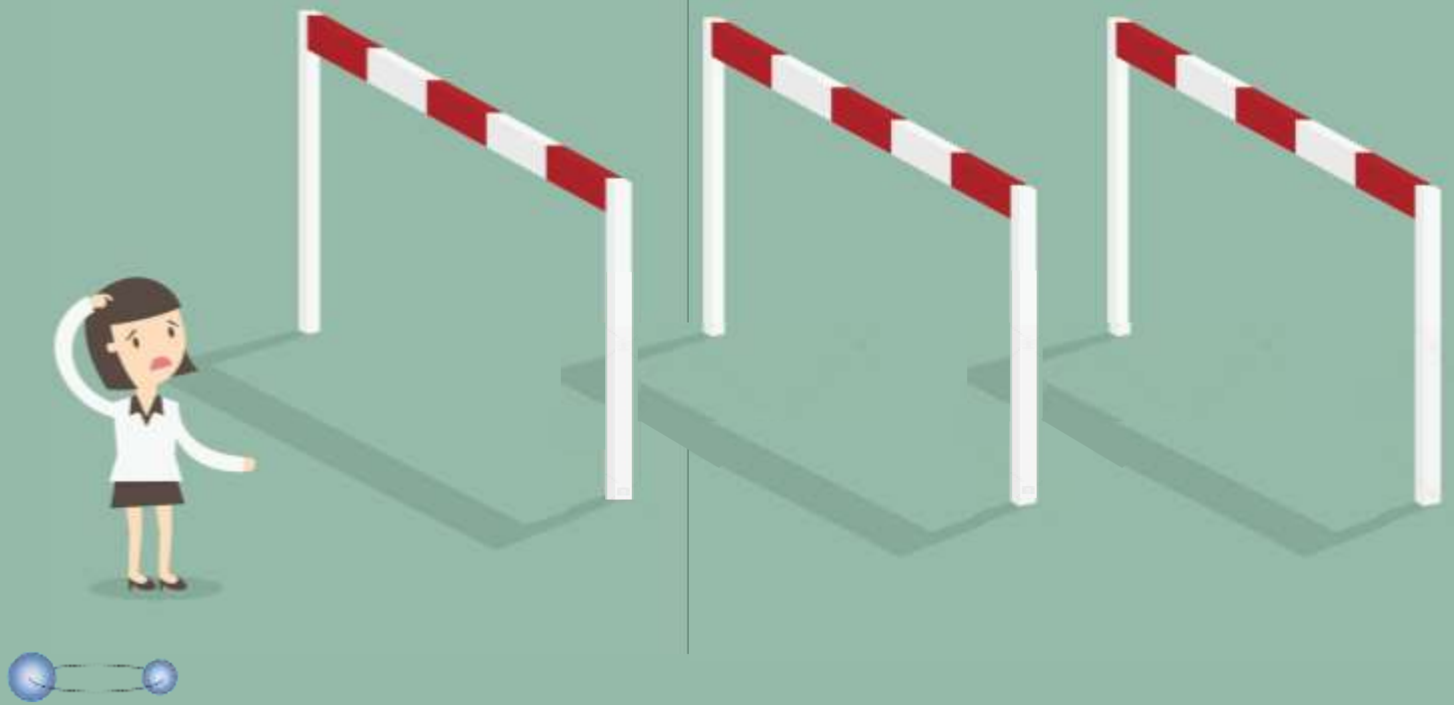




Cost: ~ 1 sec $\times\ 10^6$

Floor Broekgaarden

3 computational challenges



1 Simulating populations

2 Testing physics

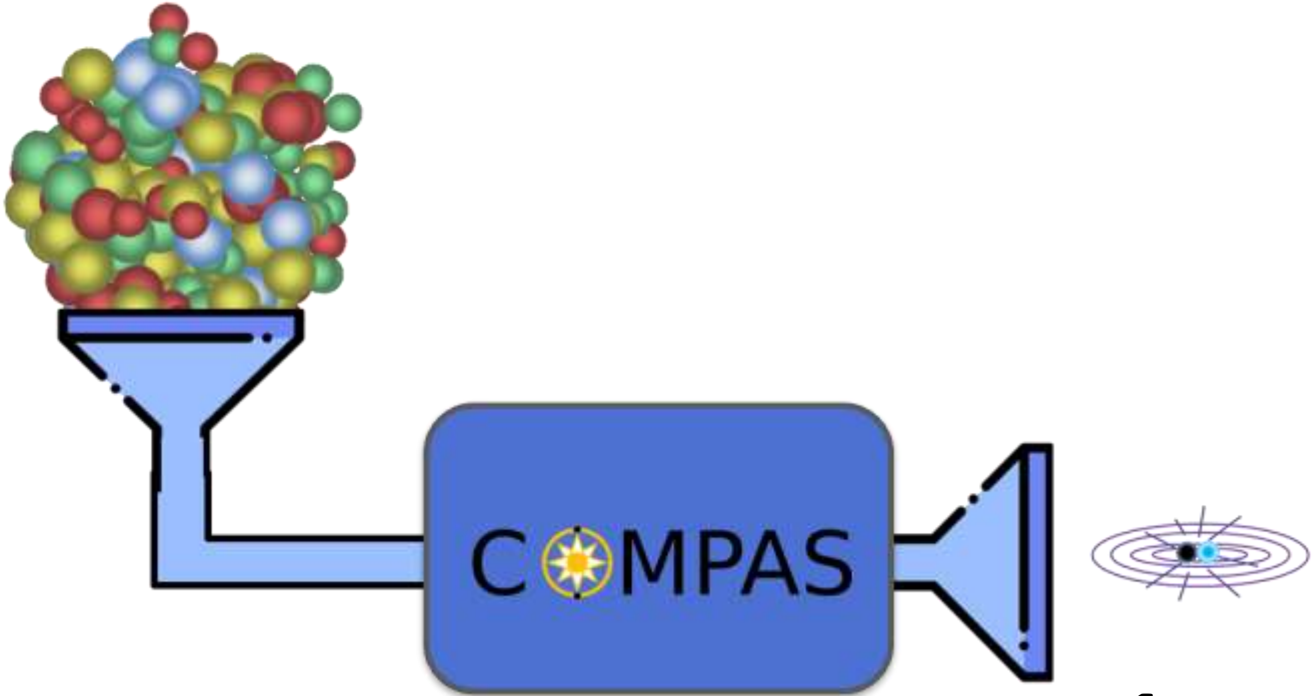
3 Rare events

Cost: ~ 1 sec $\times 10^6$ $\times 10 - 10^6$

Floor Broekgaarden

3 Rare events:

1000 binaries



only ~1

BH-NS merger!

E.g. Belczynski+02

3 computational challenges

1 Simulating populations

2 Testing physics

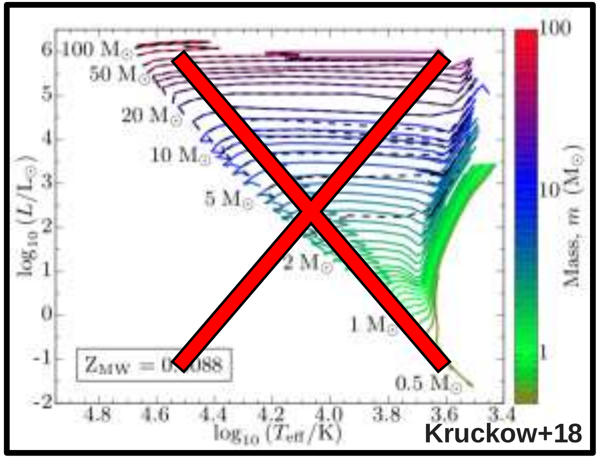
3 Rare events

Cost: ~ 1 sec $\times 10^6$ $\times 10 - 10^6$ $\times 10 - 10^3$

$\sim 3 - 3 \times 10^6$ years computing time

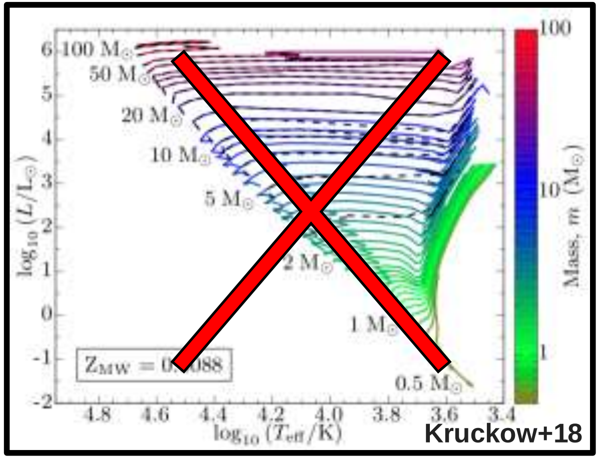
Floor Broekgaarden

Current binary population synthesis models pay a high price…

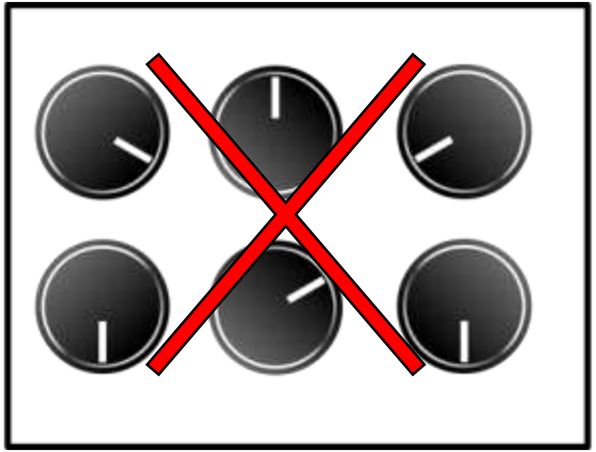


Do not include detailed prescriptions

Current binary population synthesis models pay a high price…

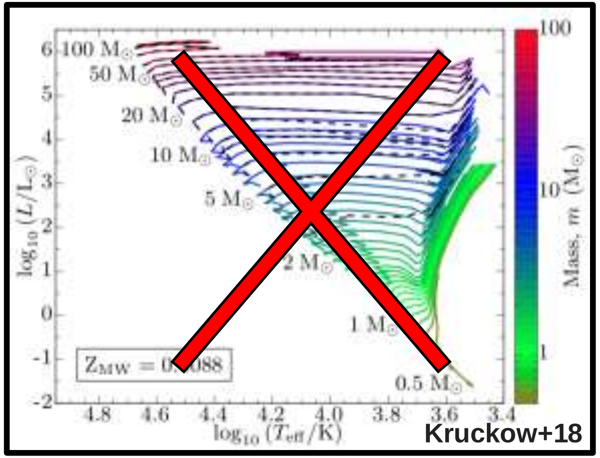


Do not include detailed prescriptions

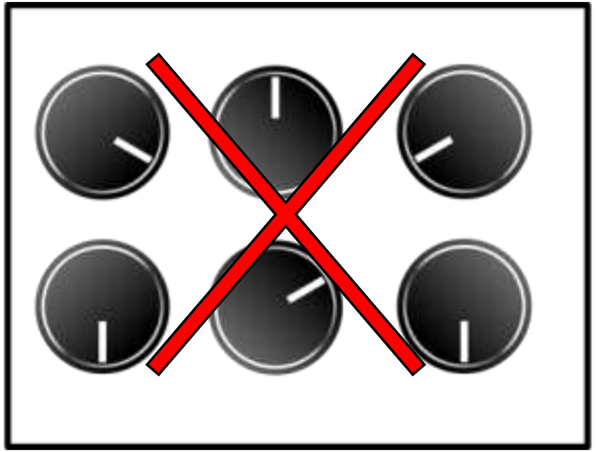


Perform only a small parameter study

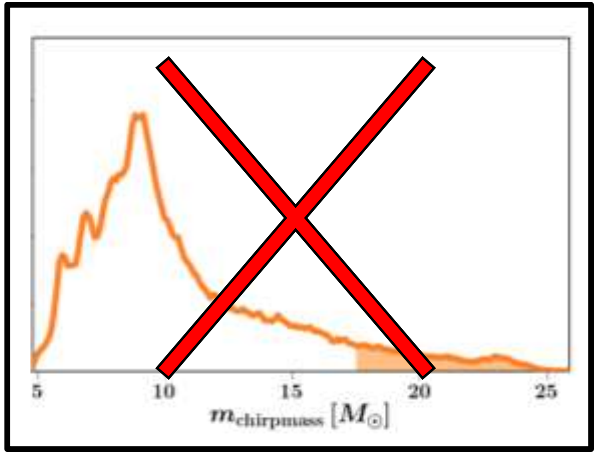
Current binary population synthesis models pay a high price…



Do not include detailed prescriptions



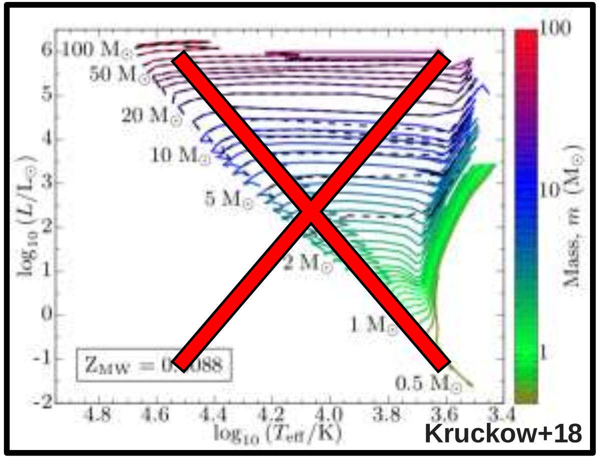
Perform only a small parameter study



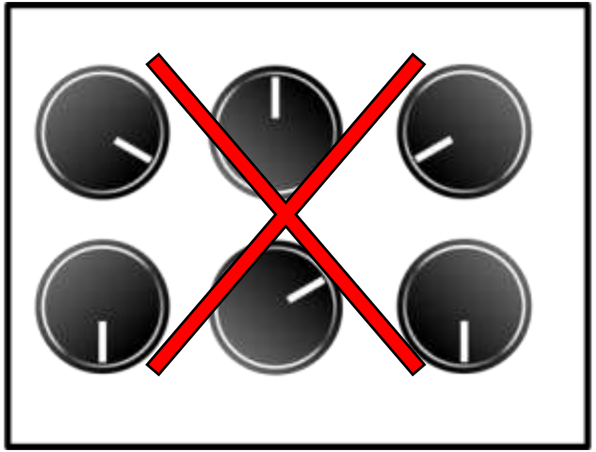
Do not explore tails of distributions

Floor Broekgaarden

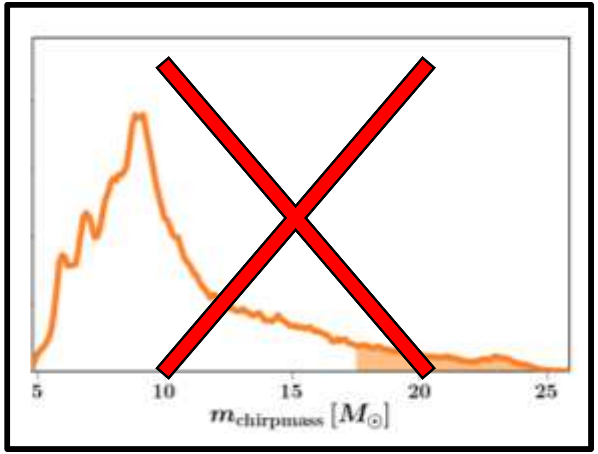
C… opulation synthesis models pay a high price…



Do not include detailed prescriptions



Perform only a small parameter study



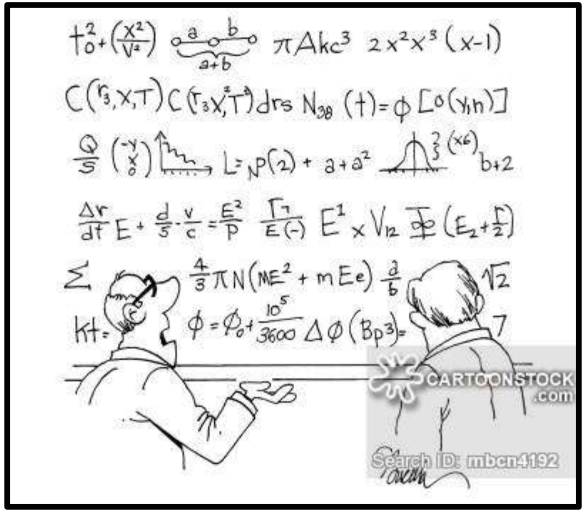
Do not explore tails of distributions

Floor Broekgaarden

Previous work

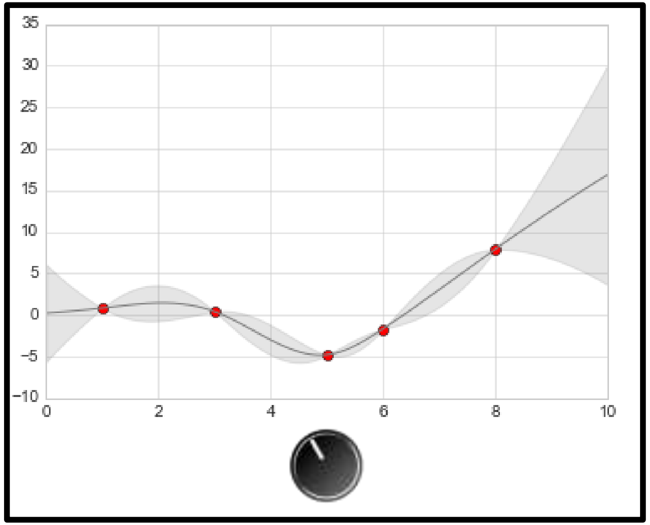
Analytical formalisms

Kolb 1993, Politano 1996, Kalogera 1996
Kalogera&Webbink 1998, Kalogera+2000



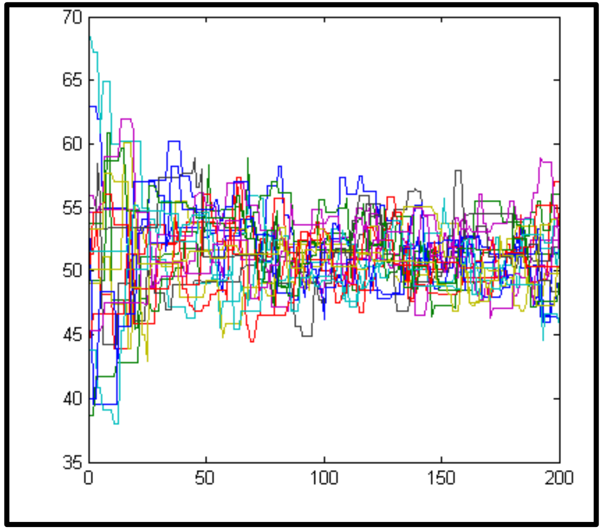
Emulators (GPR)

Barret+2017, Taylor & Gerosa 2018,
Wong & Gerosa 2019



Markov Chain Monte Carlo

Andrews+2017; (dart_board)



We present:

(drum roll)

New sampling algorithm:

Simulating The Rare Outcomes Of Populations With AIS For Efficient Learning

STROOPWAFEL

Adaptive Importance Sampling (AIS)

Marin+06, Douc+07, Owen+09, Martino+15
AIS: Torrie & Valleau 1977, Hesterberg 1995, Cappe+2004, Pennanen & Koivu 2006, Cornuet+2012, Ortiz & Pack Kaelbling (2013), …

Floor Broekgaarden

New sampling algorithm:

Simulating The Rare Out

Wi

Adaptive In

Monthly Notices
of the Royal Astronomical Society

Issues ▼ Advance articles Submit ▼ Purchase Alerts About ▼

All Monthly Notices of the ▼

MONTHLY NOTICES

Volume 490, Issue 4
December 2019

Article Contents

STROOPWAFEL: simulating rare outcomes from astrophysical populations, with application to gravitational-wave sources*

Floor S Broekgaarden ✉, Stephen Justham, Selma E de Mink, Jonathan Gair, Ilya Mandel, Simon Stevenson, Jim W Barrett, Alejandro Vigna-Gómez, Coenraad J Neijssel

Monthly Notices of the Royal Astronomical Society, Volume 490, Issue 4, December 2019, Pages 5228–5248, https://doi.org/10.1093/mnras/stz2558
Published: 27 September 2019 **Article history** ▾

Cite Permissions Share ▾

ABSTRACT

Gravitational-wave observations of double compact object (DCO) mergers are providing new insights into the physics of massive stars and the evolution of binary systems. Making the most of expected near-future observations for understanding stellar physics will rely on comparisons with binary population

), ...

New sampling algorithm:

Simulating The Rare Outcomes Of Populations With AIS For Efficient Learning

Adaptive Importance Sampling (AIS)

Marin+06, Douc+07, Owen+09, Martino+15
AIS: Torrie & Valleau 1977, Hesterberg 1995, Cappe+2004, Pennanen & Koivu 2006, Cornuet+2012, Ortiz & Pack Kaelbling (2013), …

Floor Broekgaarden

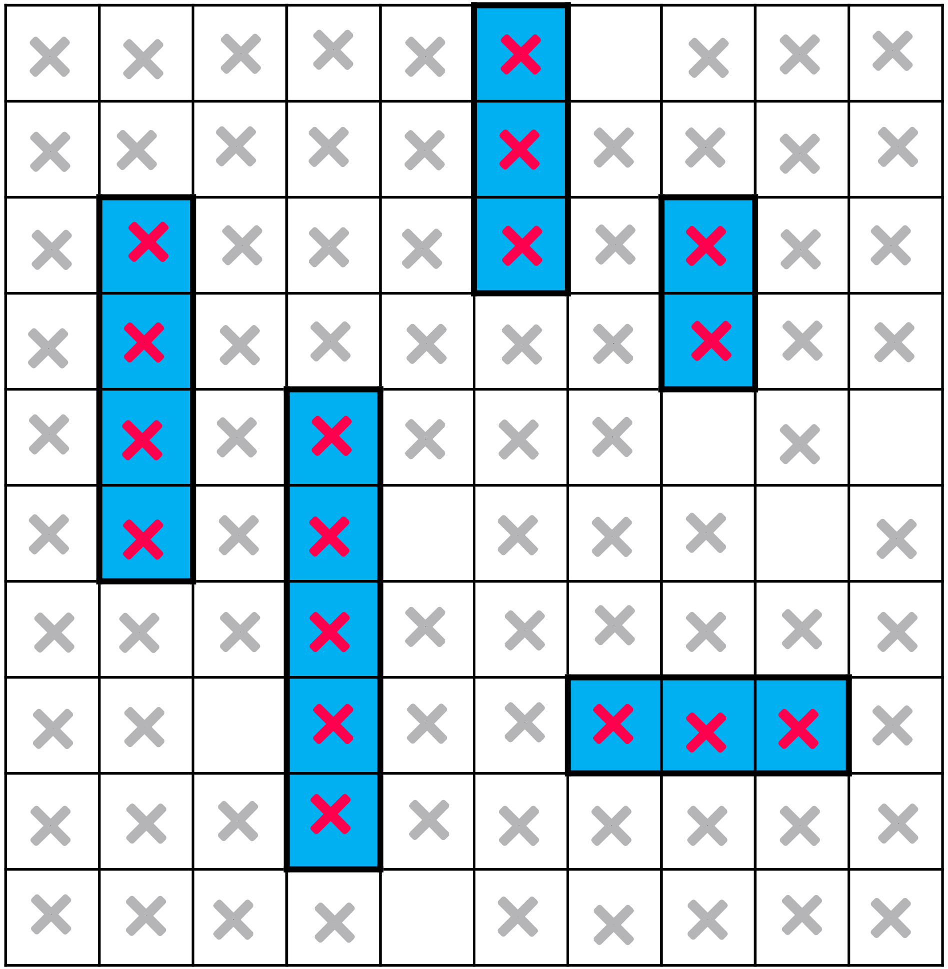
Traditional models use *"random shooting"*:

requires

< 96 >

shots

to complete the game



Floor Broekgaarden

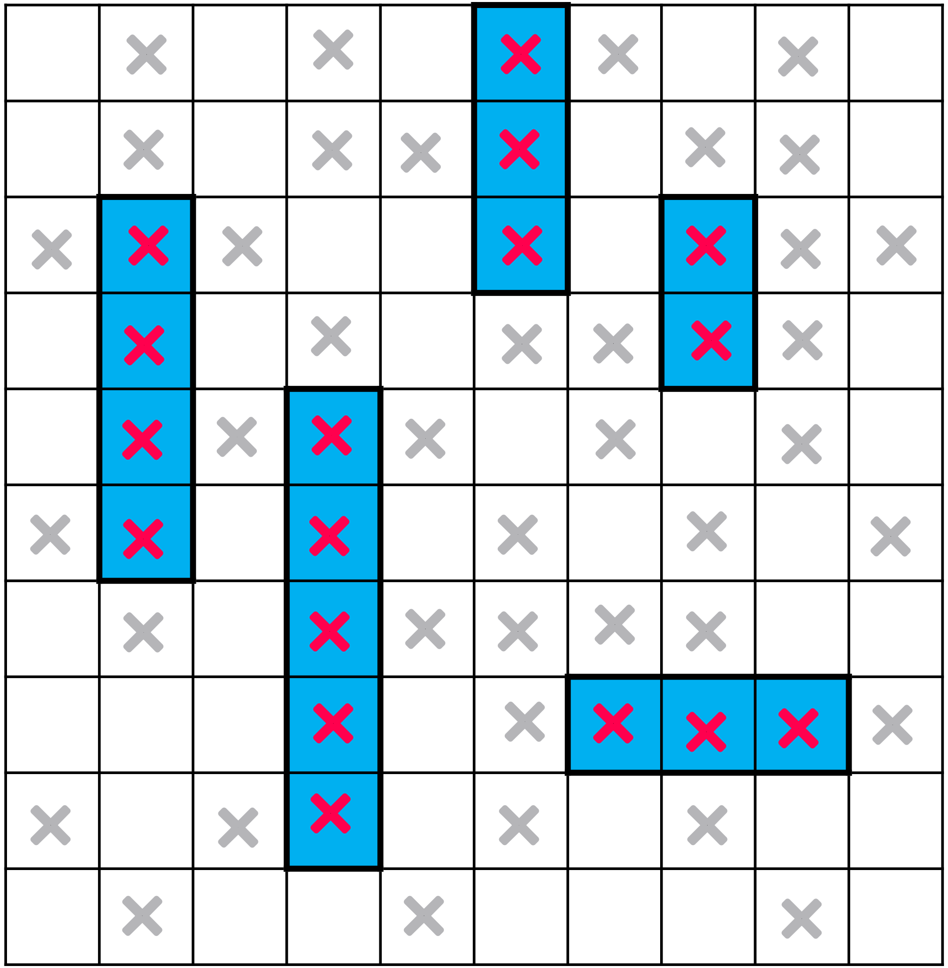
STROOPWAFEL
uses *“explore/refine”*:

requires

<??>

shots

to complete the game



Floor Broekgaarden

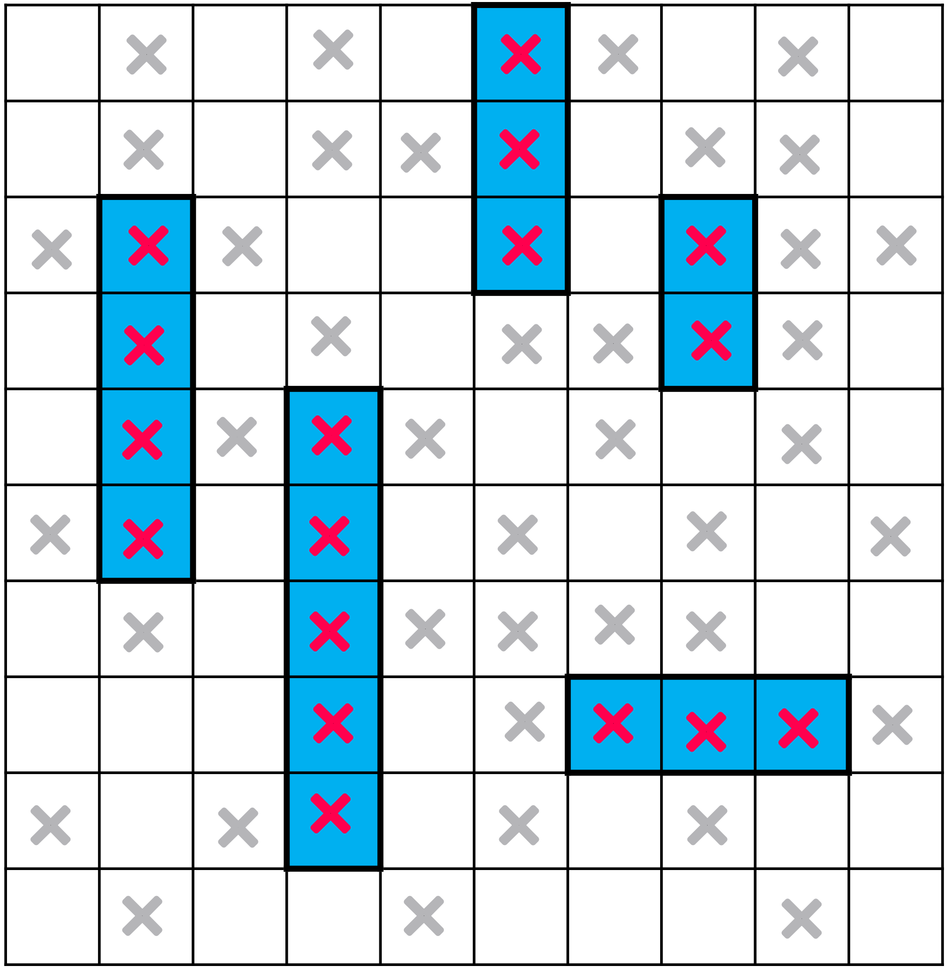
STROOPWAFEL
uses *“explore/refine”*:

requires

< 65 >

shots

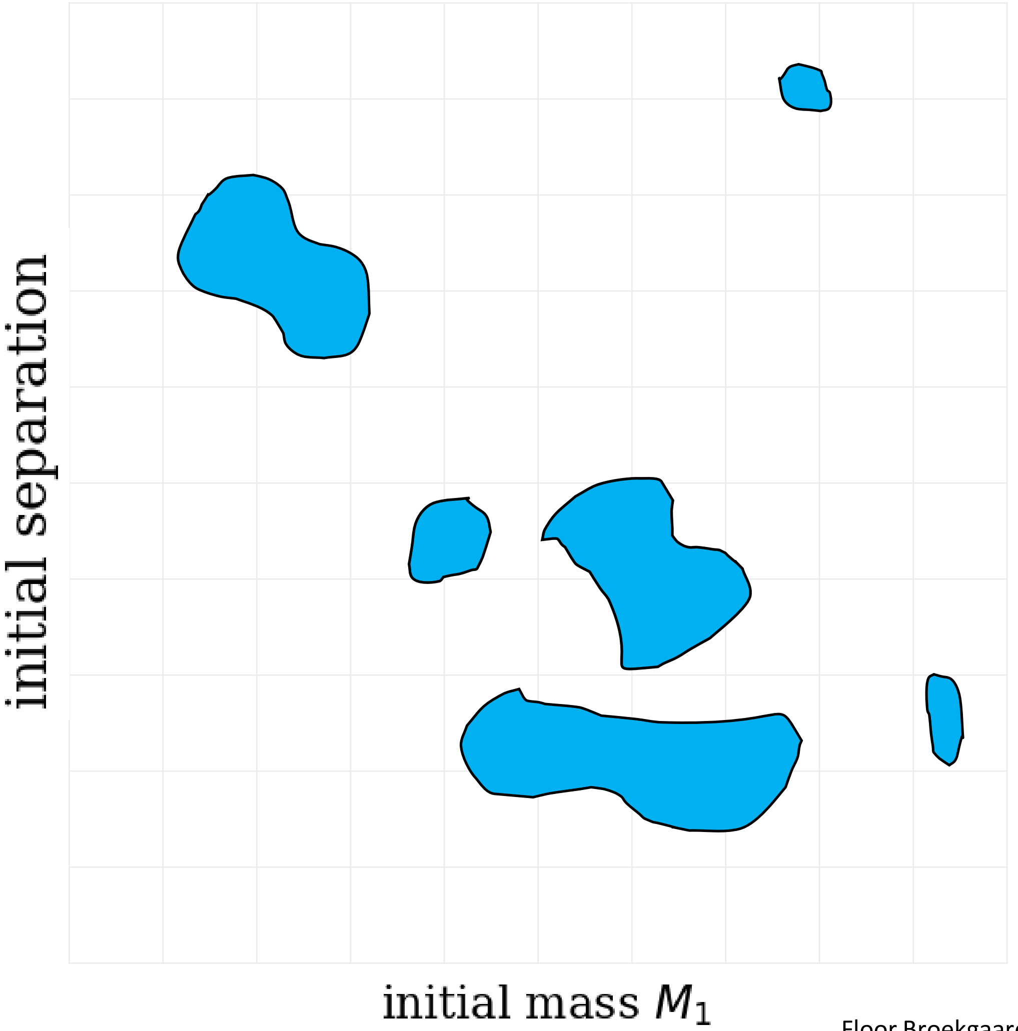
to complete the game



Floor Broekgaarden

Binary population synthesis:

- High-dimensional space,
- Unknown “islands” that form, e.g., BH-BH mergers
- unknown “rate” $\sim \frac{1}{1000}$

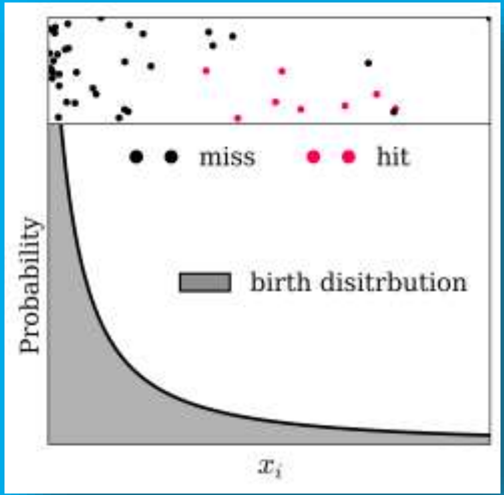


Floor Broekgaarden

prior

$$\pi(\boldsymbol{x}_i) = \pi(m_{1,i}, m_{2,i}, a_i, \cdots)$$

1) Exploring phase



Floor Broekgaarden

prior

$$\pi(\boldsymbol{x}_i) = \pi(m_{1,i}, m_{2,i}, a_i, \cdots)$$

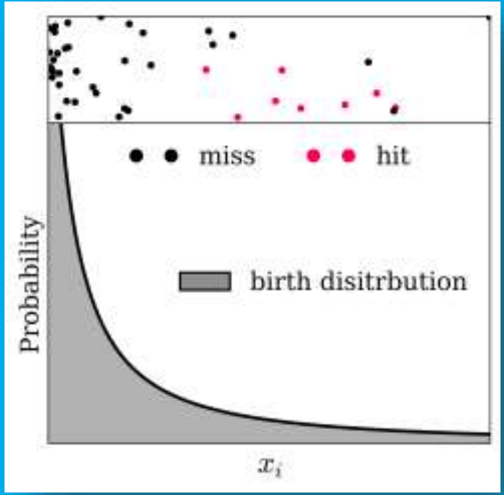
Gaussian
around hits

$$q(\boldsymbol{x}) = \frac{1}{N_{\mathrm{T,expl}}} \sum_{k=1}^{N_{\mathrm{T,expl}}} q_k(\boldsymbol{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

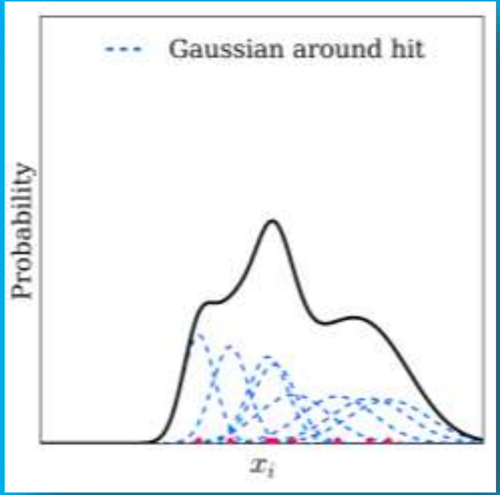
$$\sigma_{j,k} = \kappa \frac{1}{\pi_j(x_k) N_{\mathrm{expl}}^{1/d}},$$

Gaussian width

1) Exploring phase



2) Create adapted distribution



prior

$$\pi(\boldsymbol{x}_i) = \pi(m_{1,i}, m_{2,i}, a_i, \cdots)$$

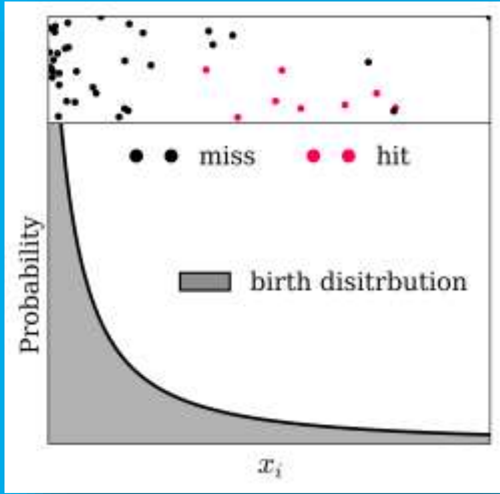
Gaussian around hits

$$q(\boldsymbol{x}) = \frac{1}{N_{\mathrm{T,expl}}} \sum_{k=1}^{N_{\mathrm{T,expl}}} q_k(\boldsymbol{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

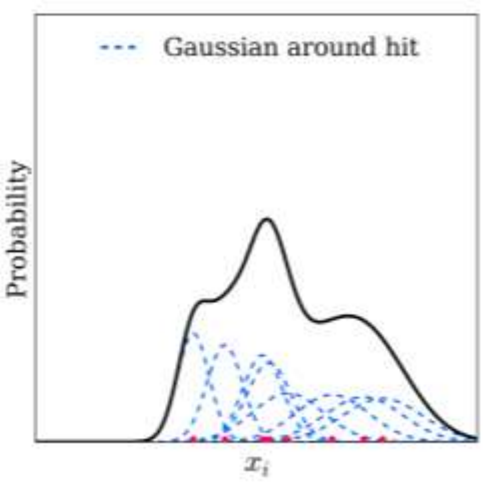
Gaussian width

$$\sigma_{j,k} = \kappa \frac{1}{\pi_j(x_k) N_{\mathrm{expl}}^{1/d}},$$

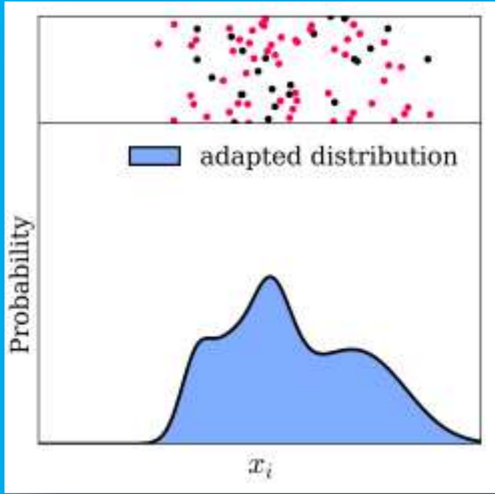
1) Exploring phase



2) Create adapted distribution



3) Refinement phase

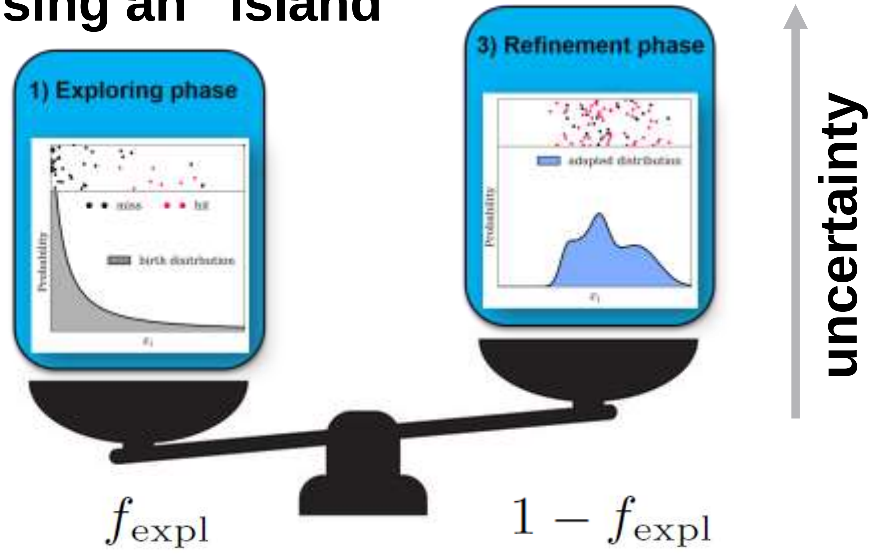


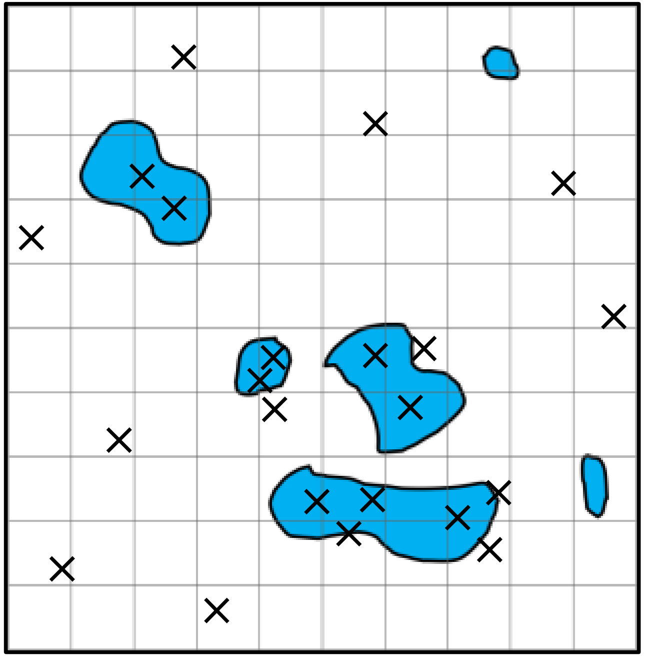
Floor Broekgaarden

When to switch from exploring to refinement?

Getting more "hits"

Missing an "island"



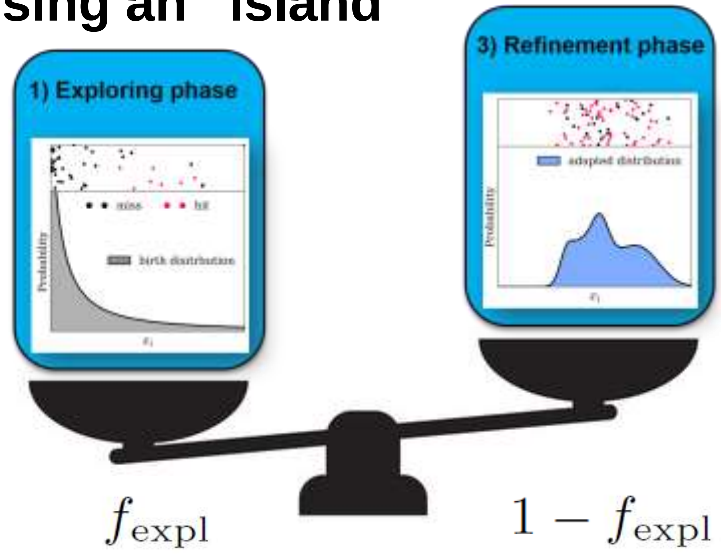


Floor Broekgaarden

When to switch from exploring to refinement?

Getting more “hits”

Missing an “island”



Alternative to e.g. Effective Sample Size (ESS) Hesterberg 1995; Liu 2008

Uncertainty from refining

$$f_{\mathrm{expl}} = 1 - \frac{z_1(\sqrt{1-z_1} - \sqrt{z_2})}{\sqrt{1-z_1}(\sqrt{z_2(1-z_1)} + z_1)}$$

uncertainty from missing an island

Ingredients

- 6 different target populations

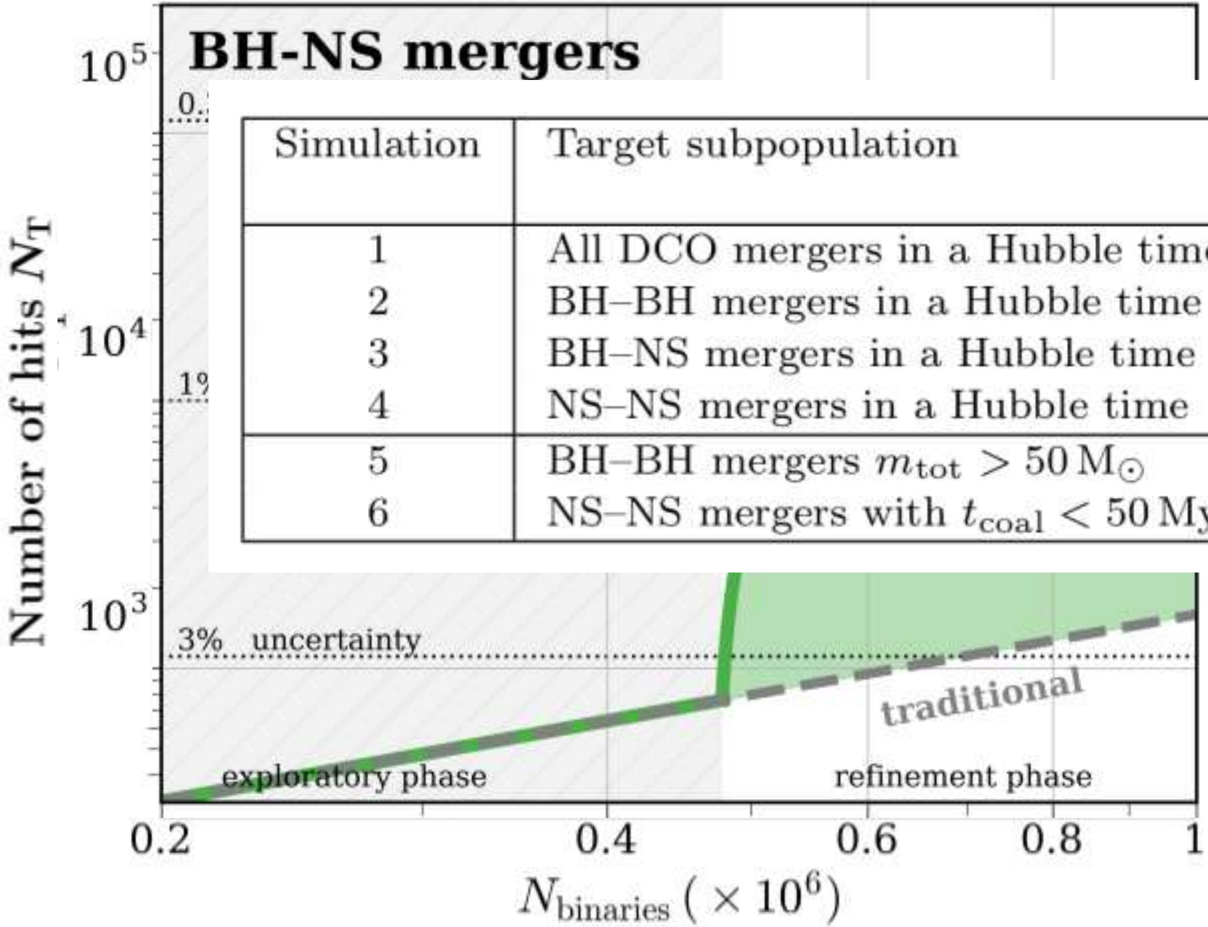
- 3 parameters:
  m1, m2 & separation

- traditional vs STROOPWAFEL sampling

Simulation	Target subpopulation
1	All DCO mergers in a Hubble time
2	BH–BH mergers in a Hubble time
3	BH–NS mergers in a Hubble time
4	NS–NS mergers in a Hubble time
5	BH–BH mergers $m_{\rm tot} > 50\,{\rm M}_{\odot}$
6	NS–NS mergers with $t_{\rm coal} < 50\,{\rm Myrs}$

Floor Broekgaarden

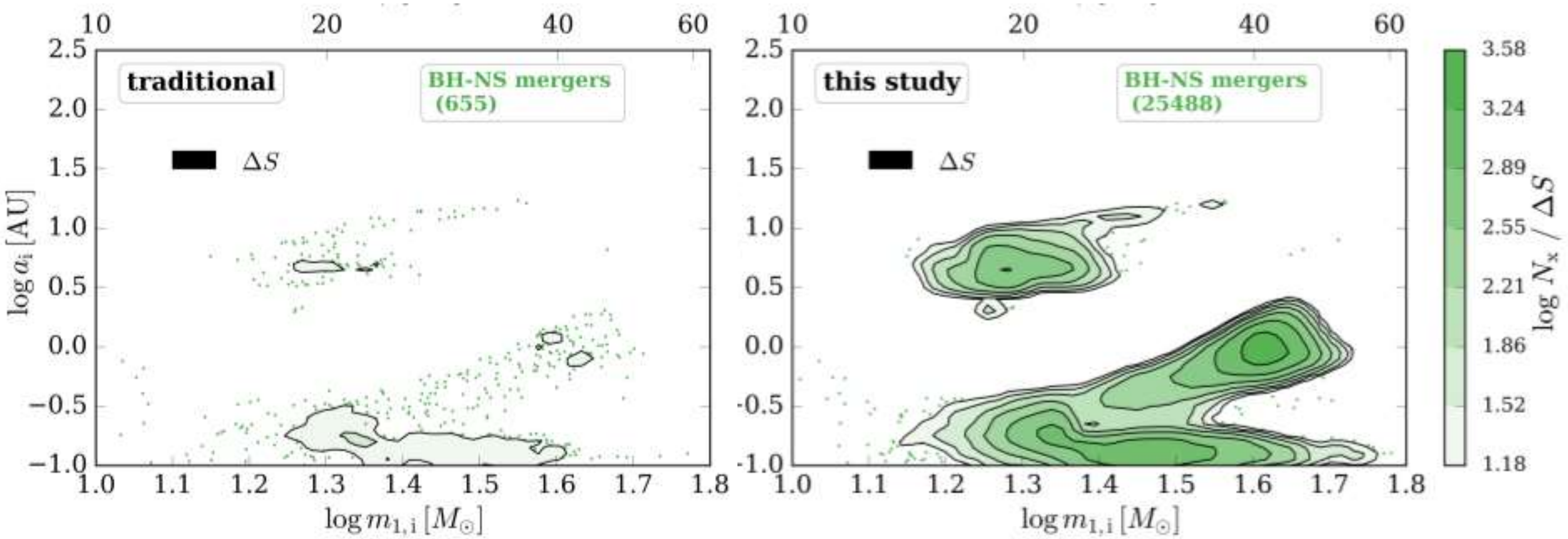
Higher efficiency



Simulation	Target subpopulation	gain
1	All DCO mergers in a Hubble time	35×
2	BH–BH mergers in a Hubble time	53×
3	BH–NS mergers in a Hubble time	39×
4	NS–NS mergers in a Hubble time	45×
5	BH–BH mergers $m_{\mathrm{tot}} > 50\,\mathrm{M}_\odot$	202×
6	NS–NS mergers with $t_{\mathrm{coal}} < 50\,\mathrm{Myrs}$	24×

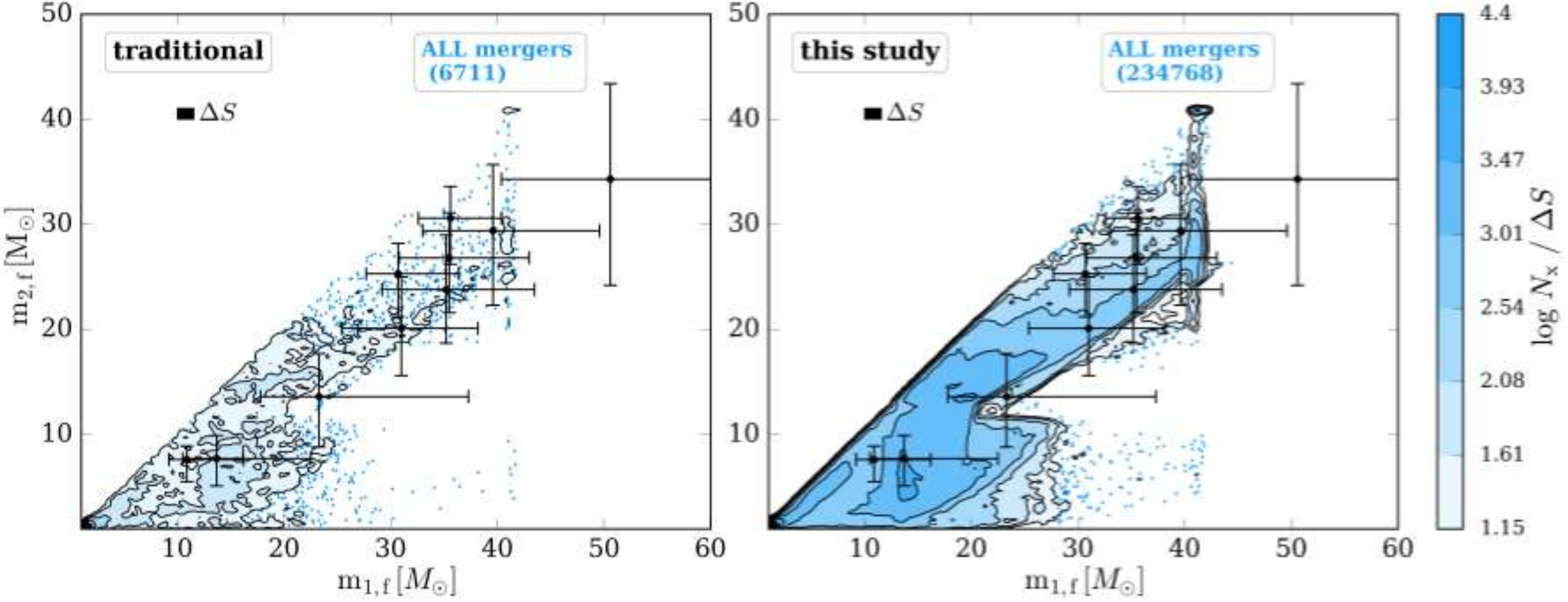
Floor Broekgaarden

Higher resolution on input parameters



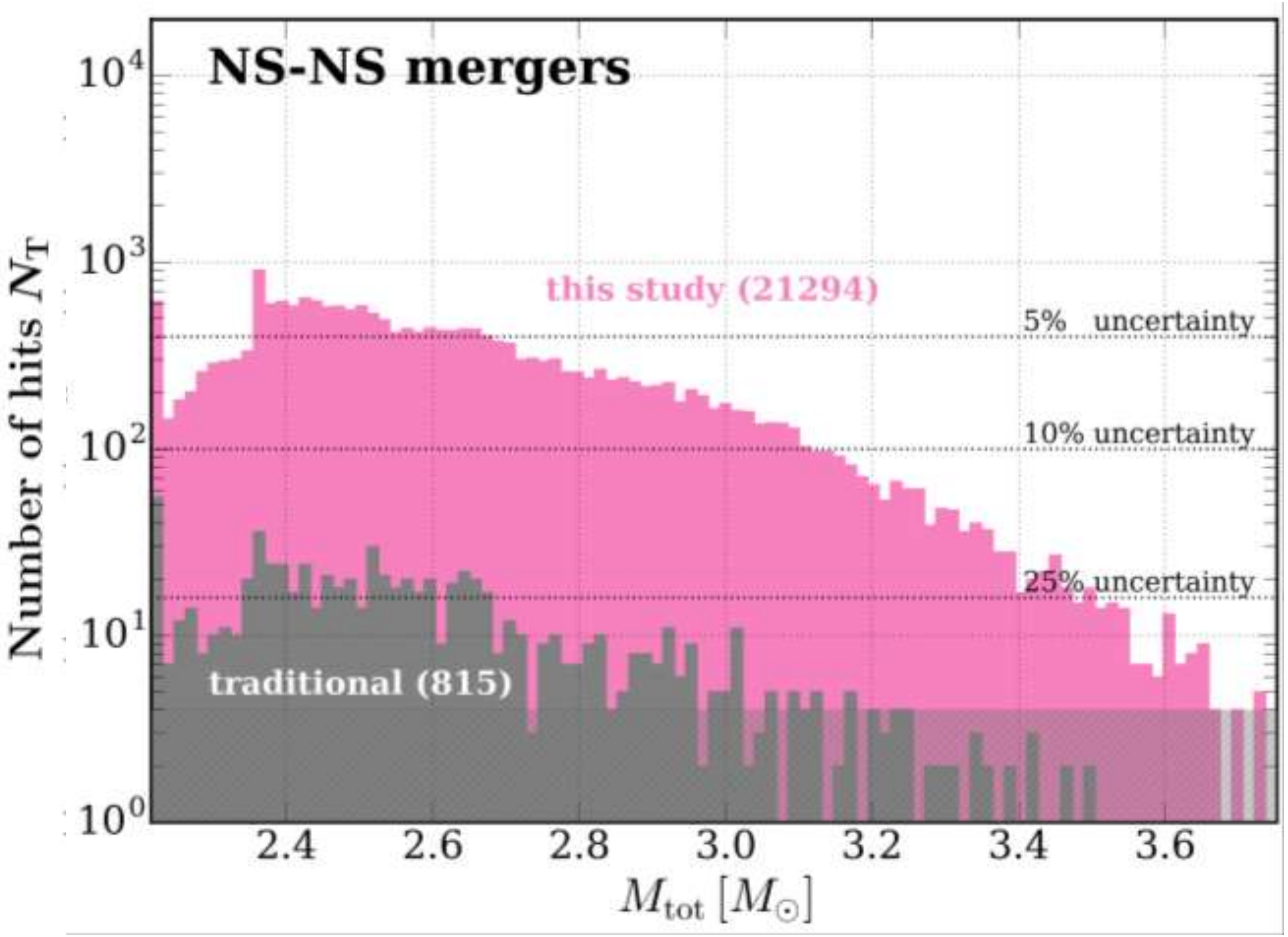
Floor Broekgaarden

Higher resolution output



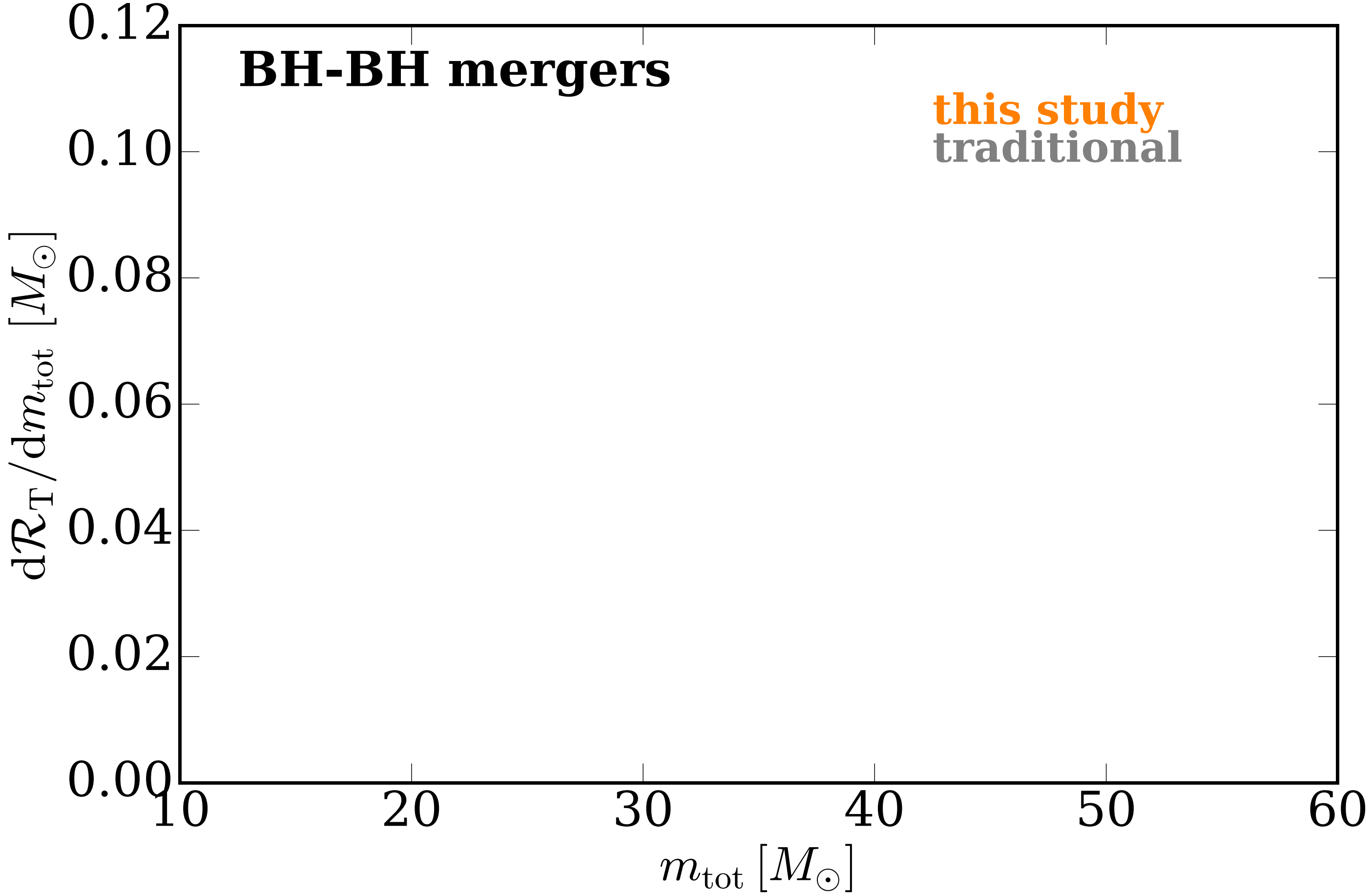
Floor Broekgaarden

Resolve tails of distributions



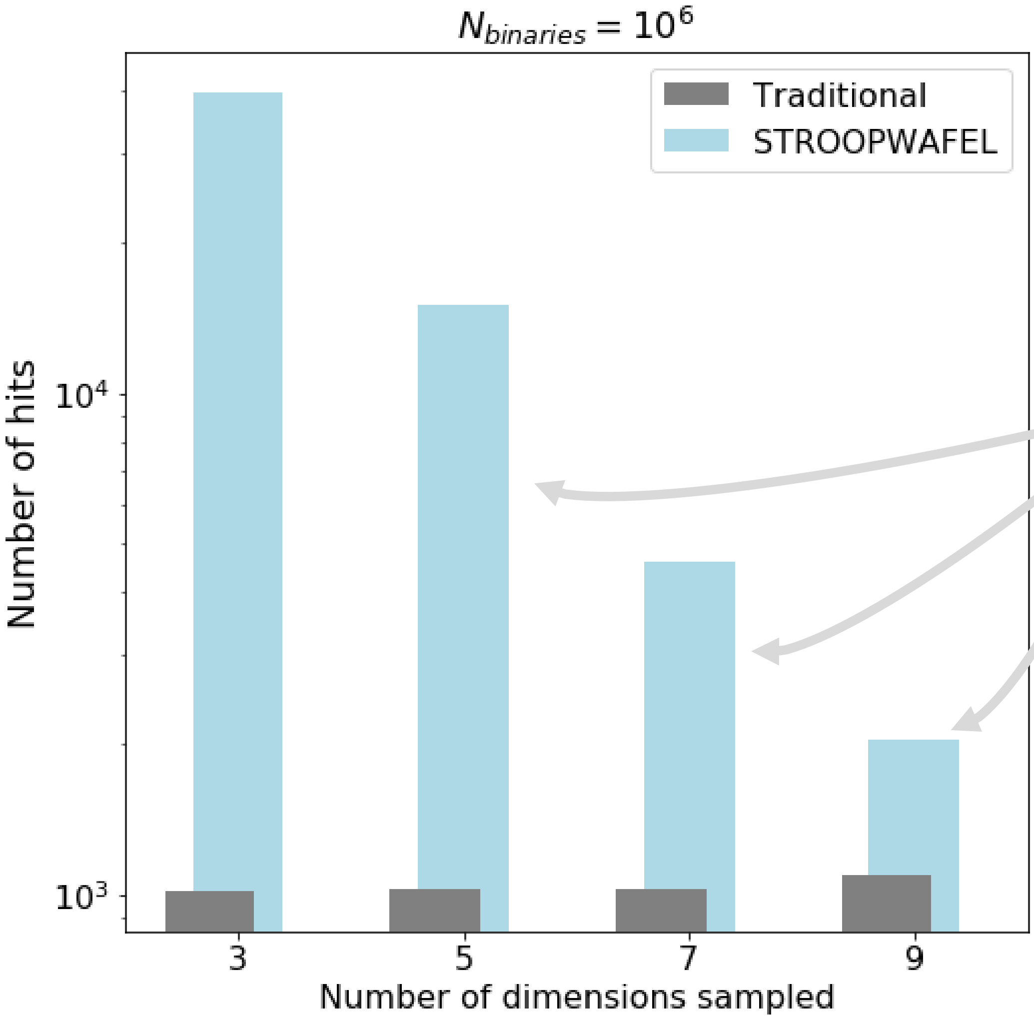
Floor Broekgaarden

Better distribution functions



Floor Broekgaarden

STROOPWAFEL paper accepted & public release of code & data:

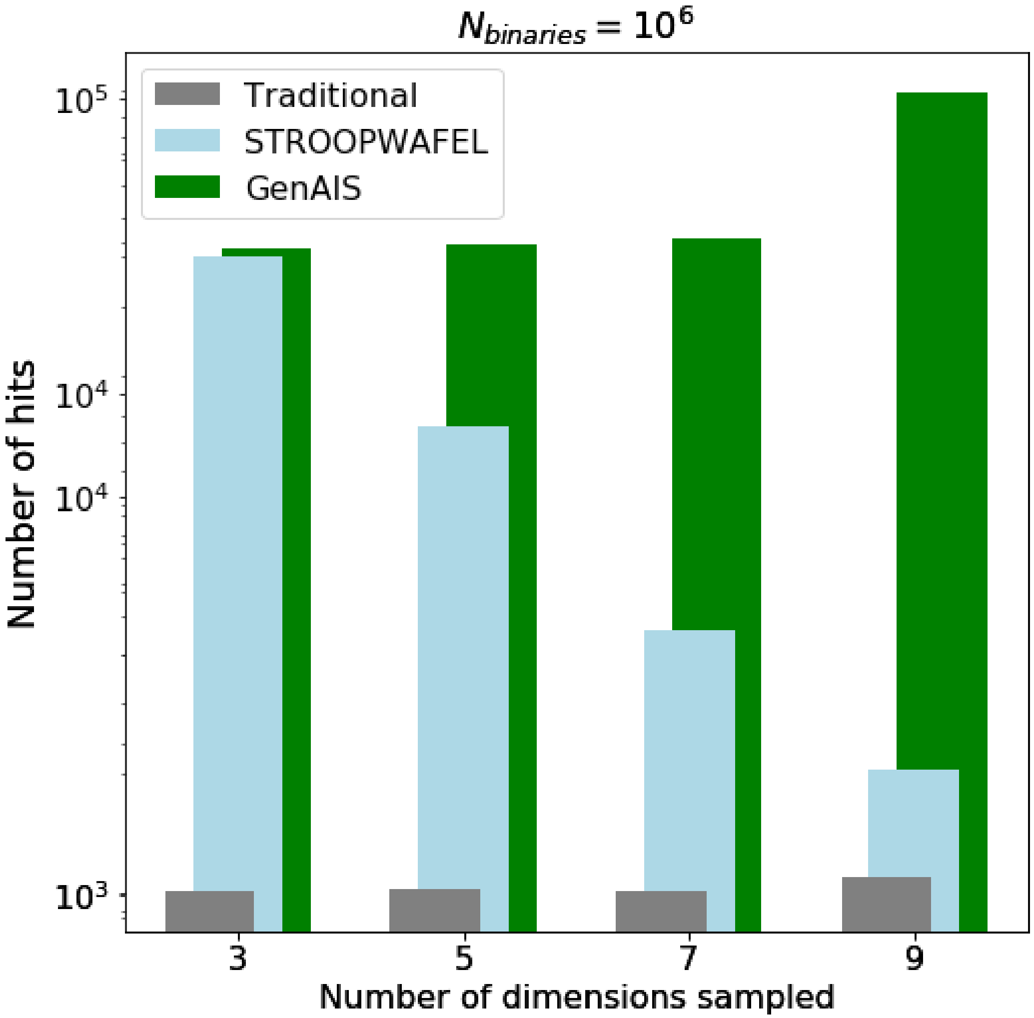


Higher dimensions?:

From Khandelwal in prep.

Curse of dimensionality suffered by STROOPWAFEL

Lokesh Khandelwal



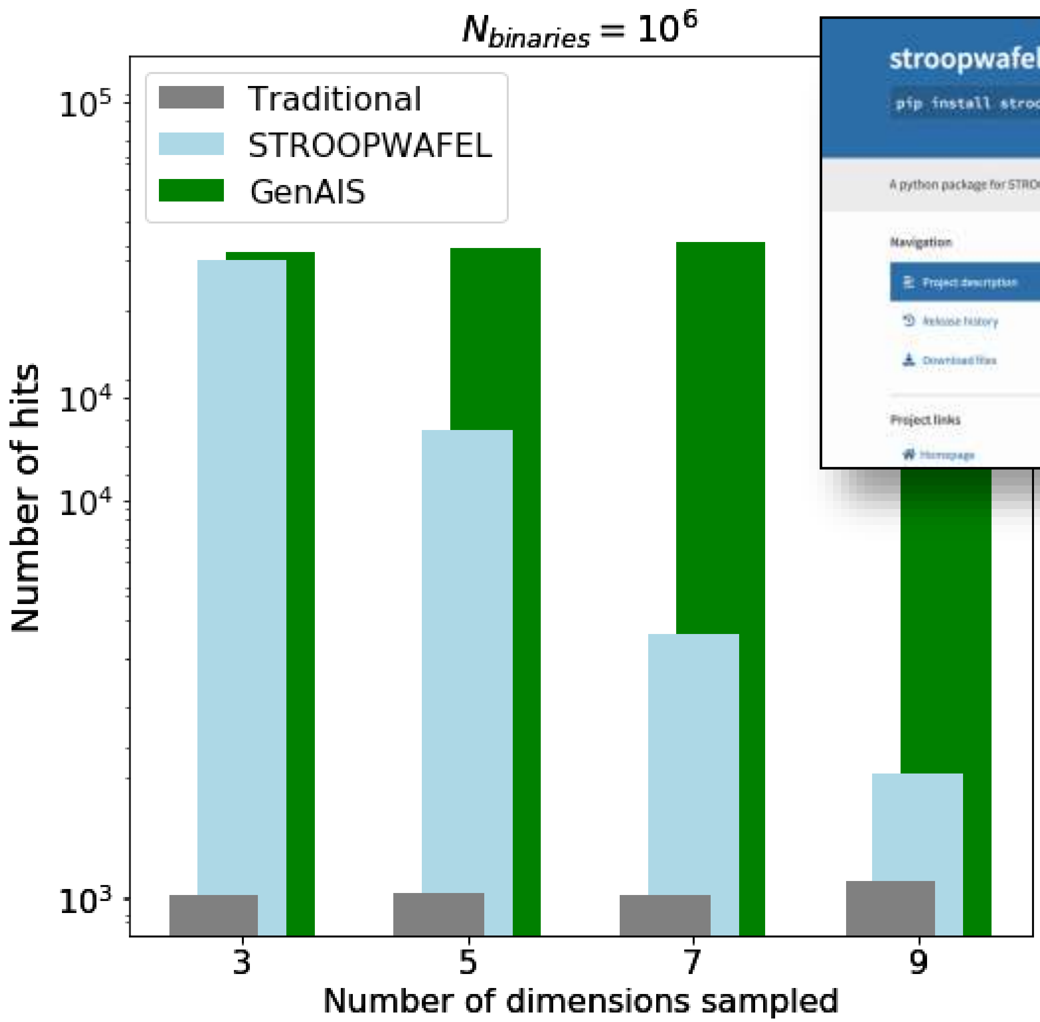
Higher dimensions?:

From Khandelwal in prep.

Curse of dimensionality suffered by STROOPWAFEL
improved using GenAIS (generational Adaptive Importance Sampling)
Based on Wraith+2009

Lokesh Khandelwal

Floor Broekgaarden



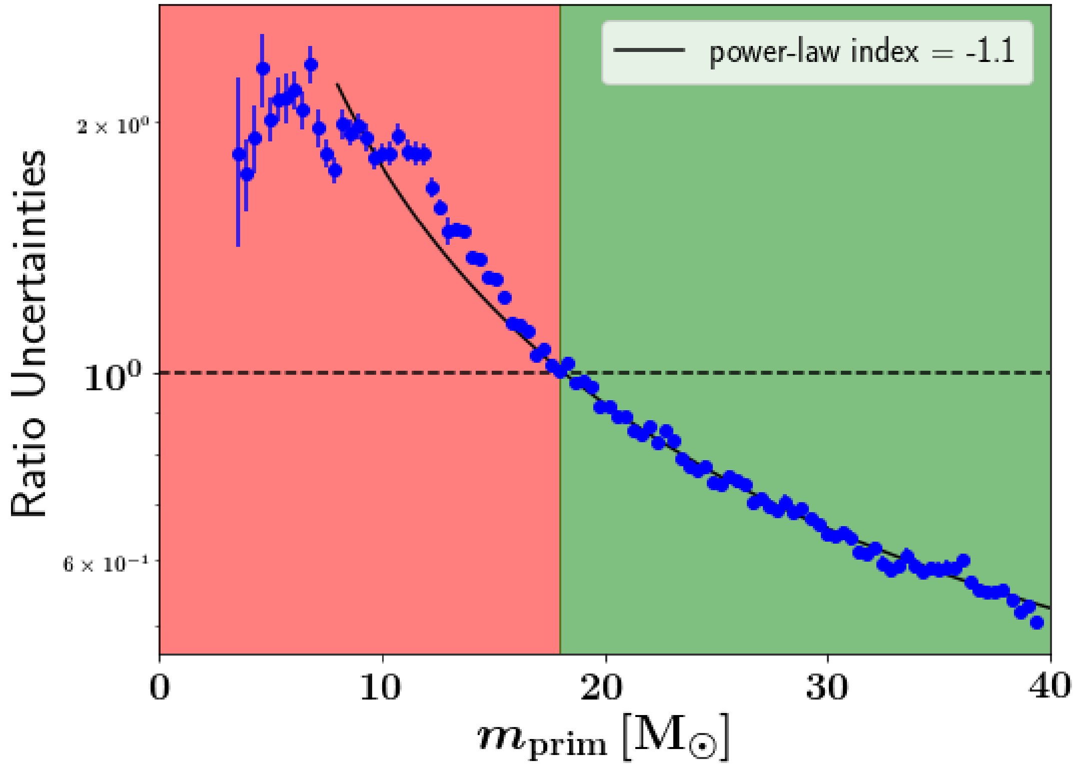
Based on Wraith+2009

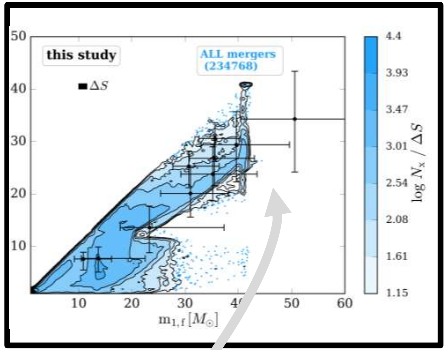
Lokesh Khandelwal

Floor Broekgaarden

Observational selection biases?

Kummer et al. in prep.



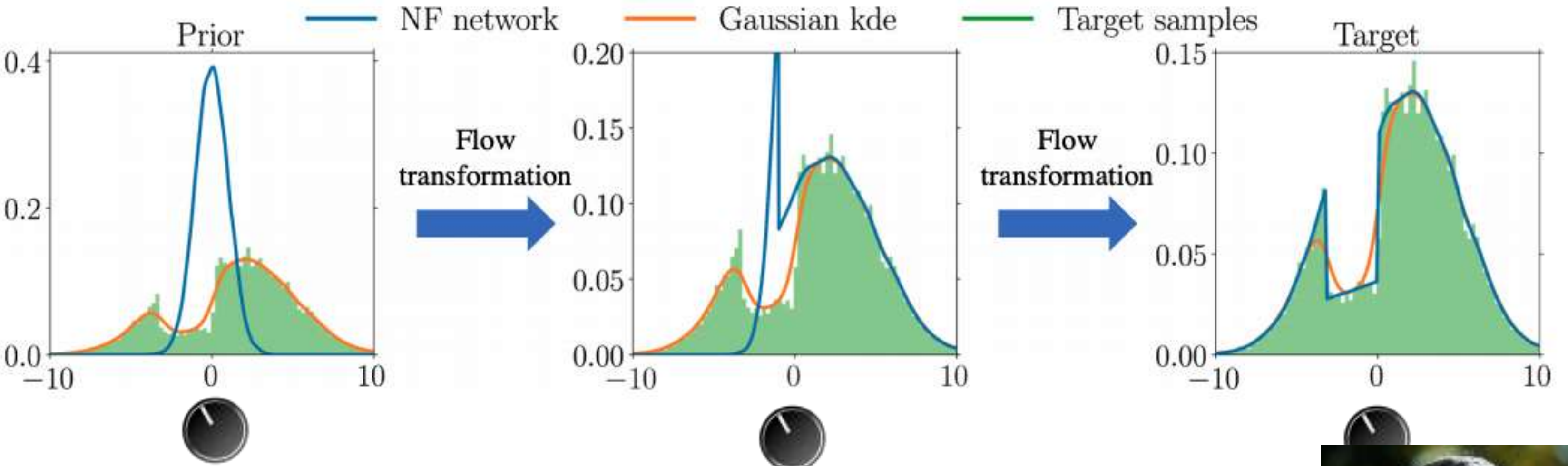


We want to improve resolution for the massive BH-BHs

Floris Kummer

Hierarchical Bayesian modeling with normalizing flows

ArXiv: 2002.09491
and 1909.06373

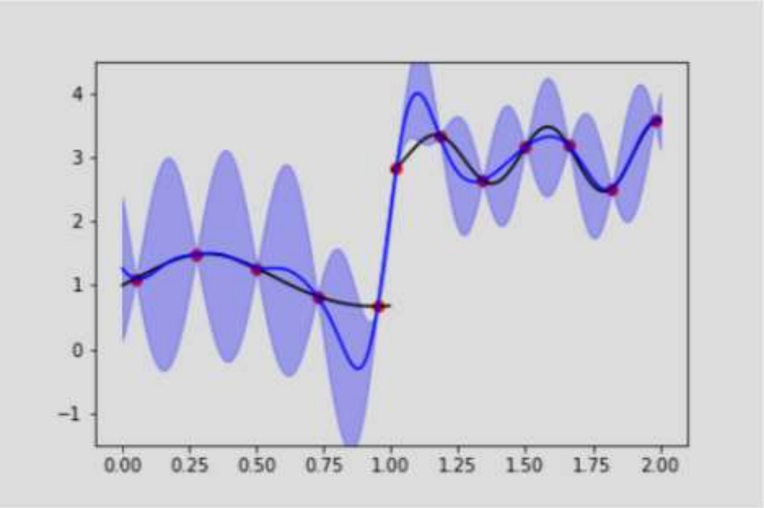


Kaze Wong
finishing PhD

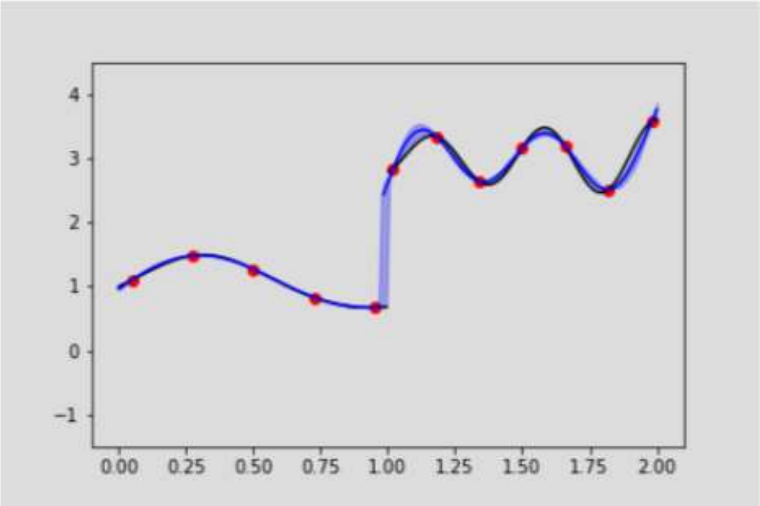
See also Kaze et al. recent work on using this for GWTC-2 in:
ArXiv: 2011.03564 and 2011.01865

Future work: Gaussian process classifiers & emulation + local response surface model

Traditional Gaussian Process Regression



Luyao Lin (in prep.)

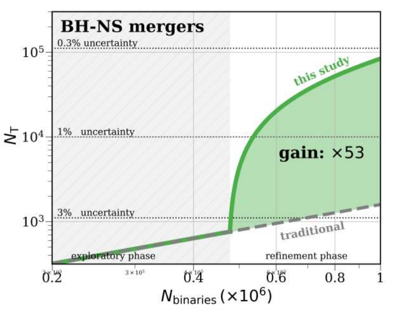


Luyao Lin

Derek Bingham

Floor Broekgaarden

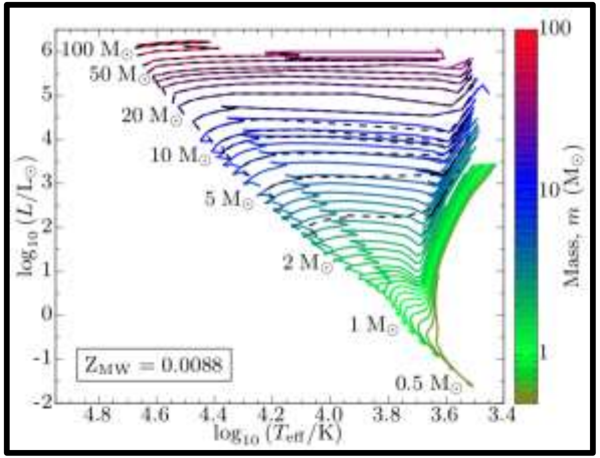
Conclusions

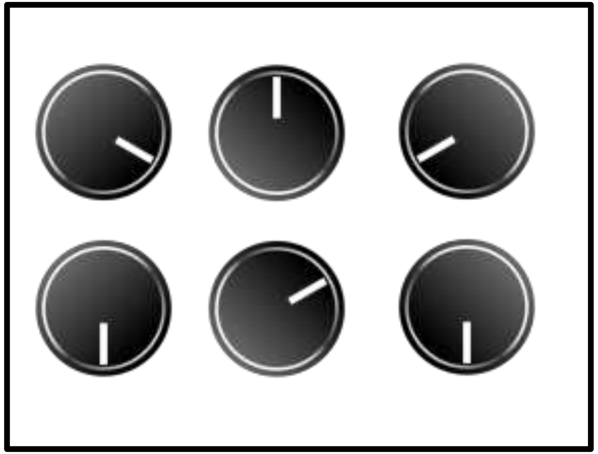


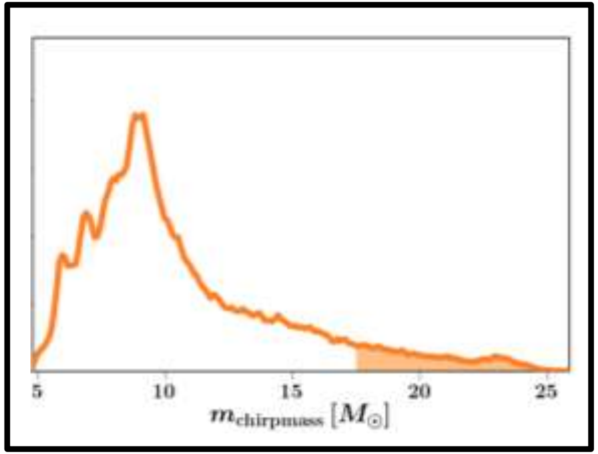
- STROOPWAFEL is an sampling algorithm that obtains $\times\, 30 - 200$ more hits or equivalently: speed up

Broekgaarden et al., 2019
Data, code and all scripts availa
 on Zenodo/github

STROOPWAFEL helps next generation simulations:







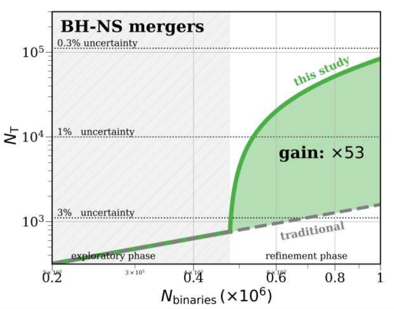
Include detailed prescriptions

Perform larger parameter study

Explore tails of distributions

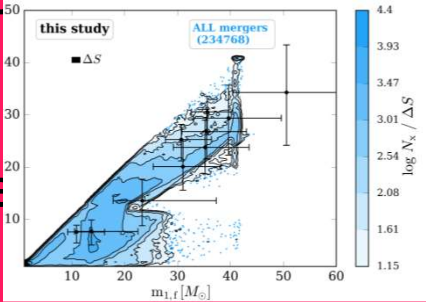
Floor Broekgaarden

Conclusions



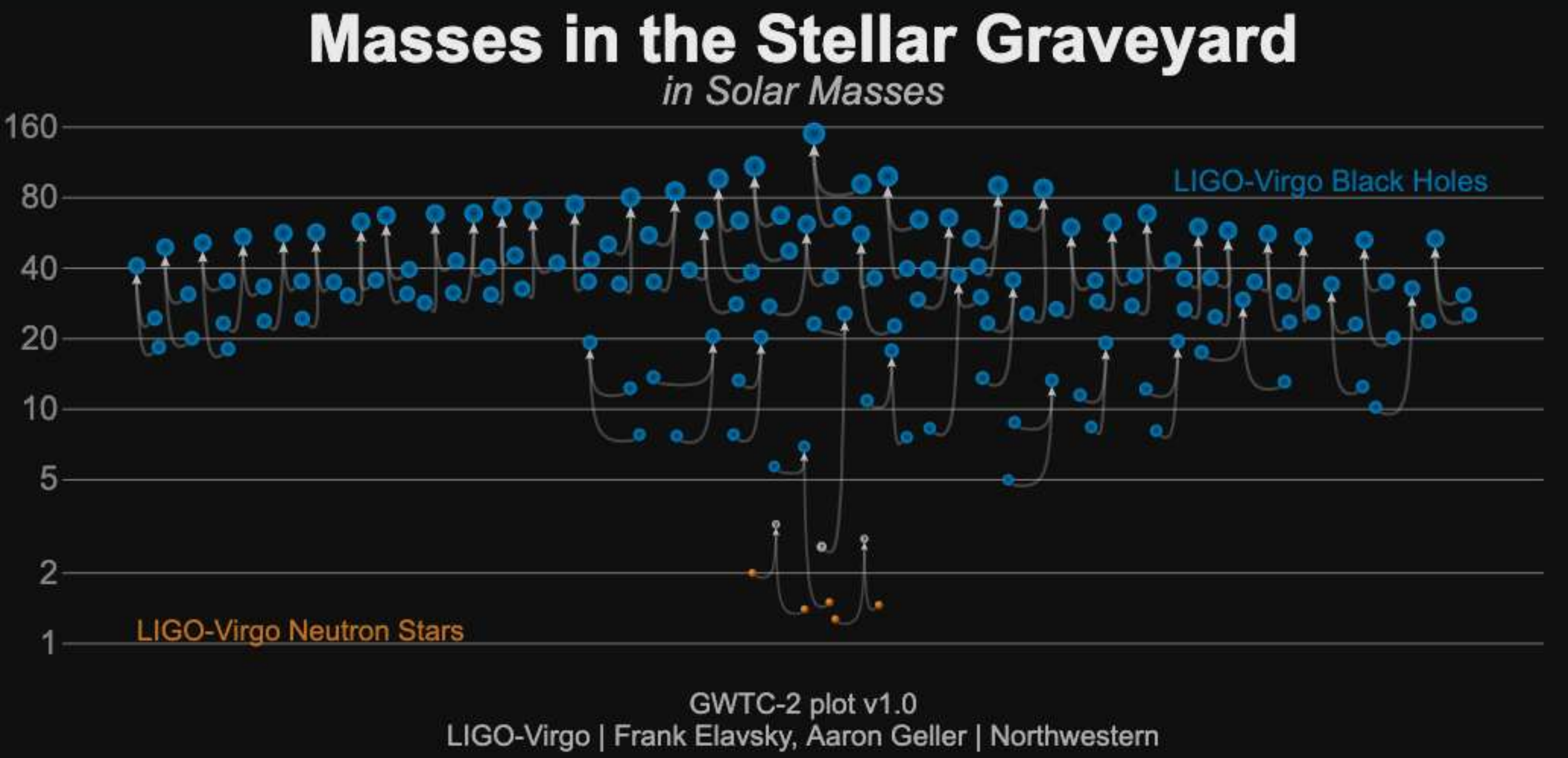
- STROOPWAFEL is an sampling algorithm that obtains $\times 30 - 200$ more hits or equivalently: speed up

- STROOPWAFEL improves resolution of population synthesis studies of GW events



- Future improvements are being actively developed + STROOPWAFEL is being implemented in other codes

Broekgaarden et al., 2019
Data, code and all scripts available on Zenodo/github



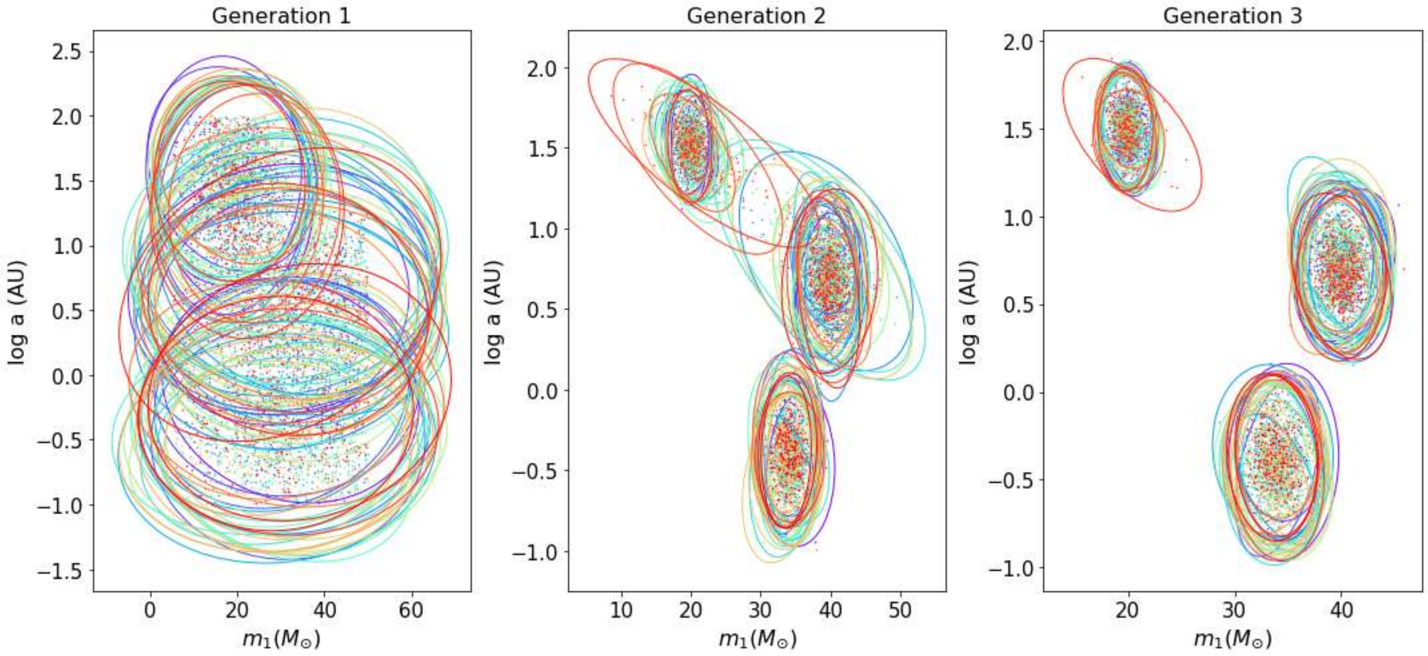
Floor Broekgaarden

Thank you!

Image source: xkcd

Back-up slides

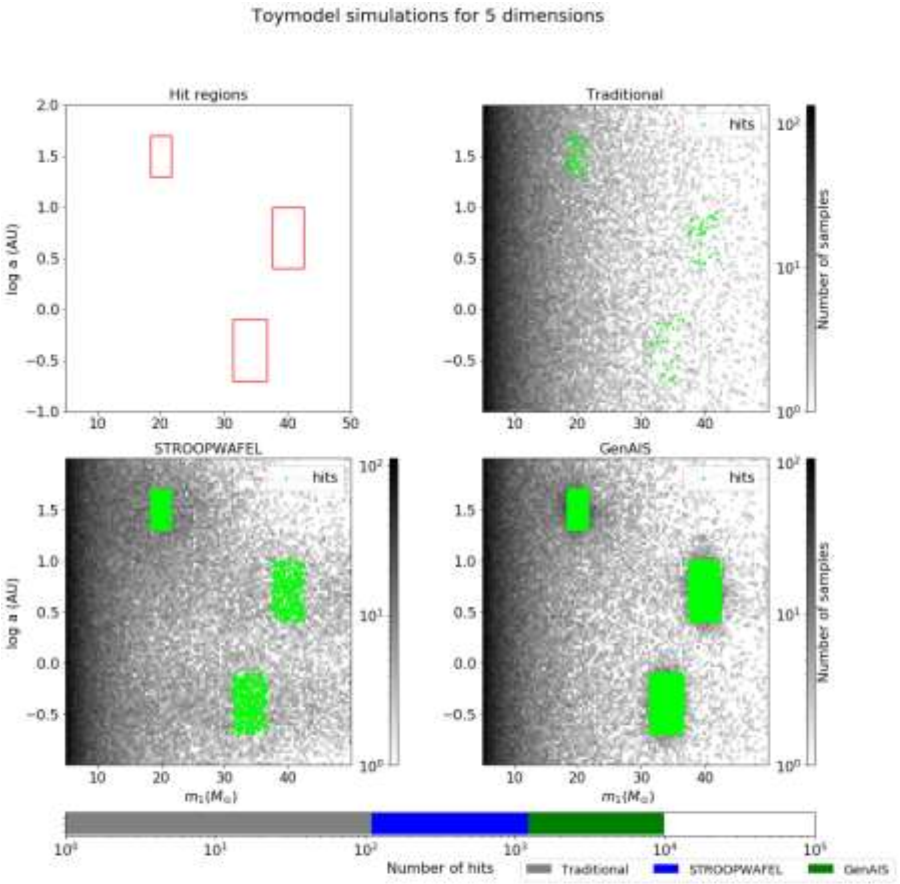
Future Prospects: more dimensions



Floris Kummer

Lokesh Khandelwal

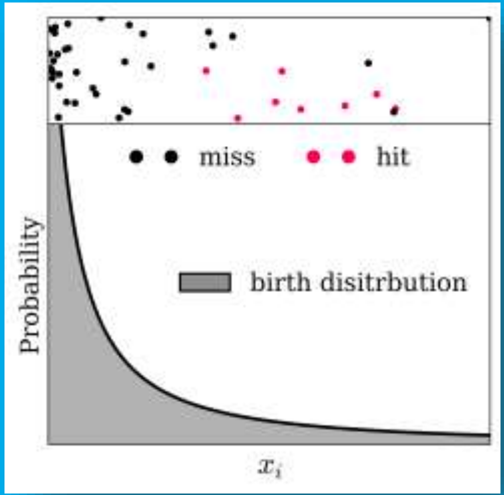
Future Prospects: more dimensions



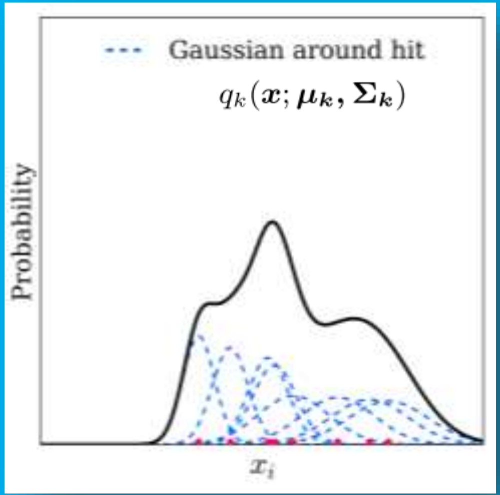
Floris Kummer

Lokesh Khandelwal

1) Exploring phase



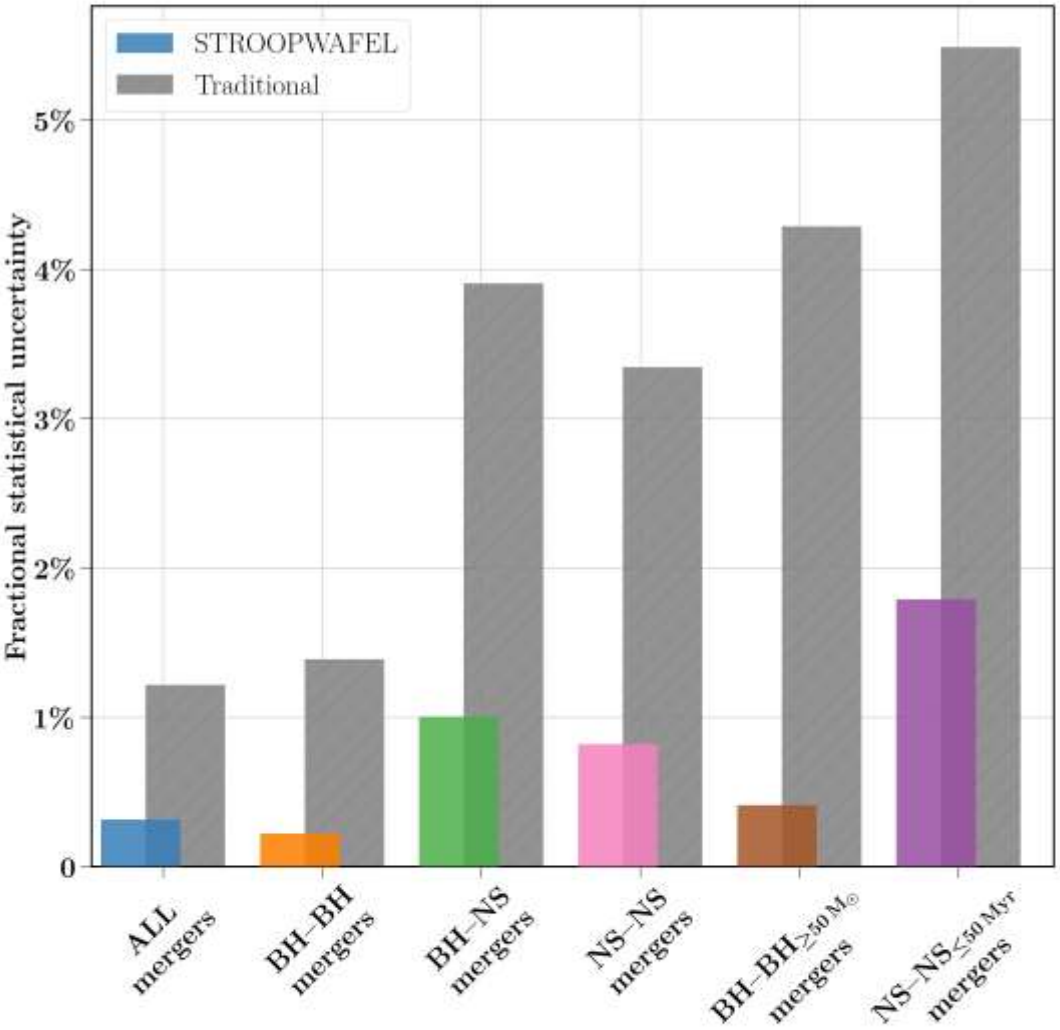
2) Create adapted distribution



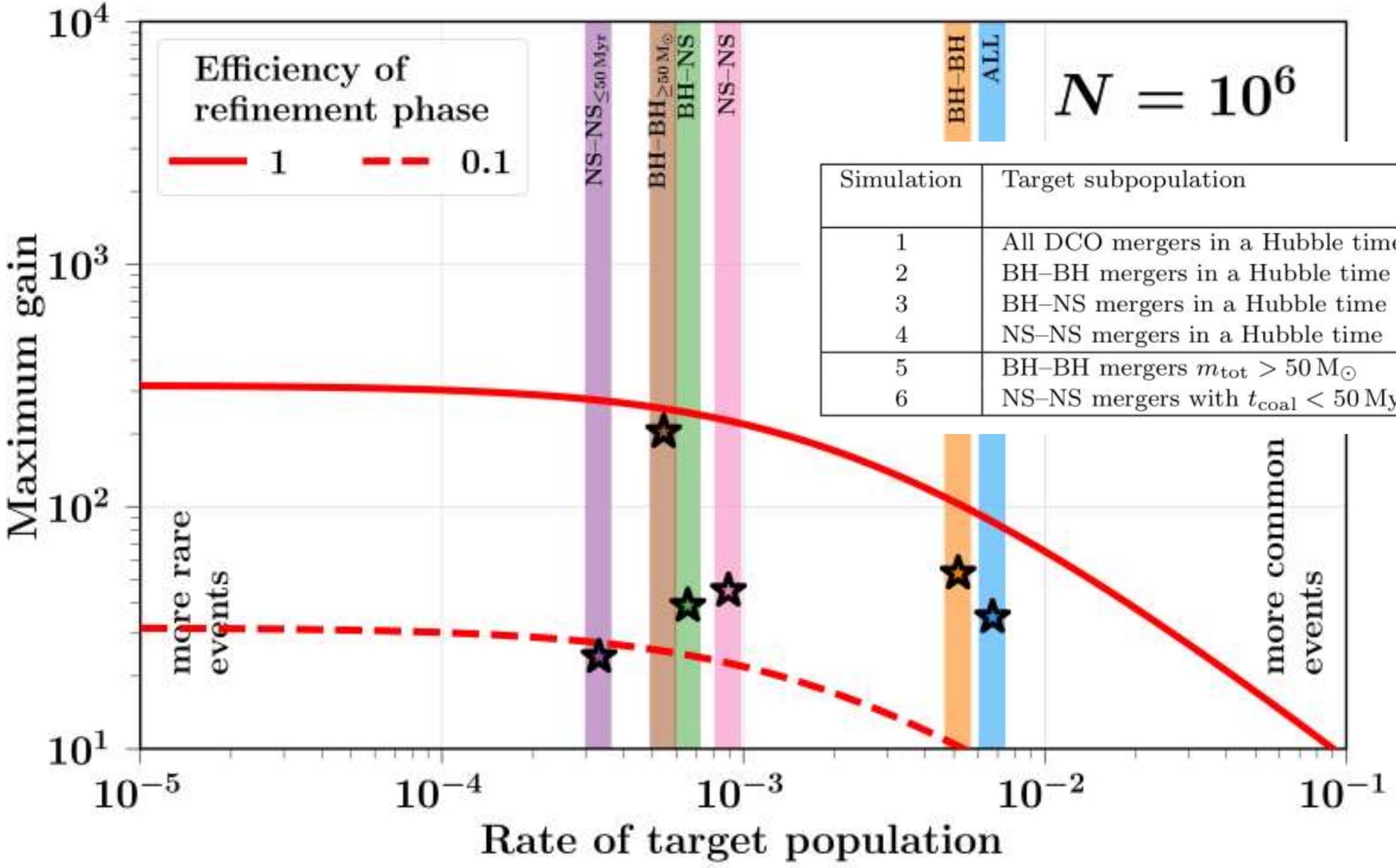
$$\boldsymbol{\Sigma_k} = \begin{bmatrix} \sigma_{1,k}^2 & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & \sigma_{d,k}^2 \end{bmatrix},$$

$$\sigma_{j,k} = \kappa \frac{1}{\pi_j(x_k) N_{\mathrm{expl}}^{1/d}},$$

Floor Broekgaarden

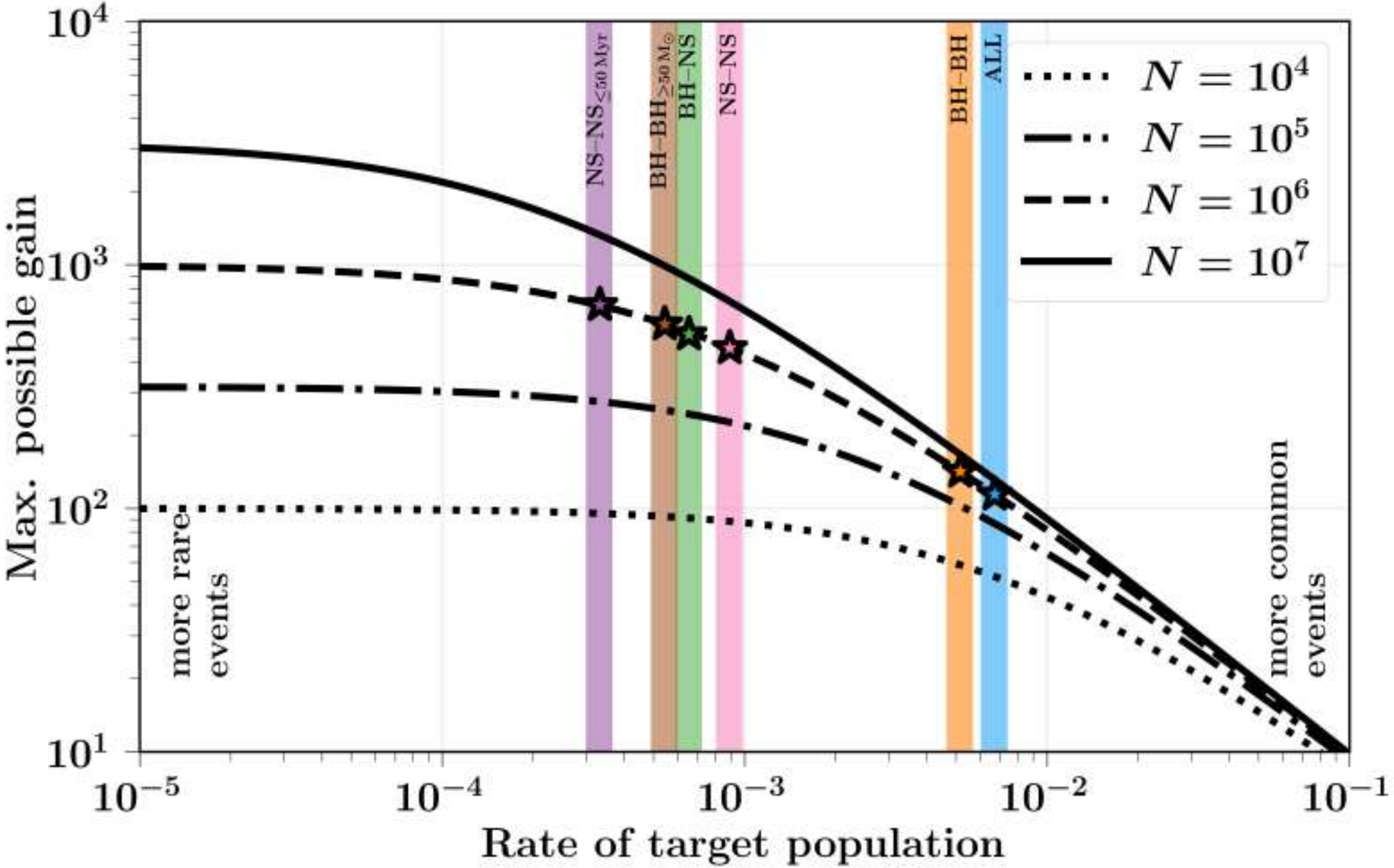


Floor Broekgaarden

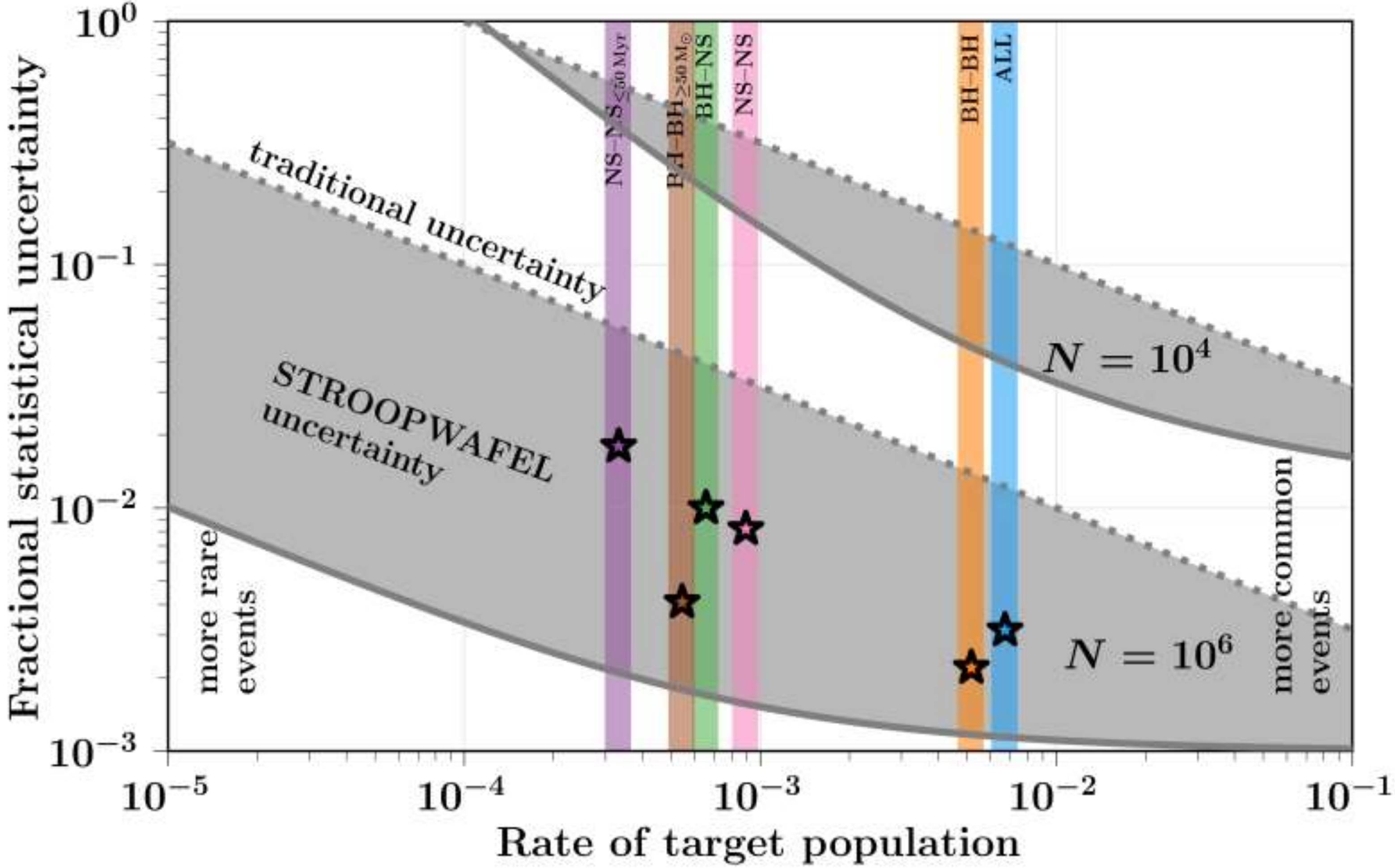


Simulation	Target subpopulation	gain
1	All DCO mergers in a Hubble time	35×
2	BH–BH mergers in a Hubble time	53×
3	BH–NS mergers in a Hubble time	39×
4	NS–NS mergers in a Hubble time	45×
5	BH–BH mergers $m_{\mathrm{tot}} > 50\,\mathrm{M}_\odot$	202×
6	NS–NS mergers with $t_{\mathrm{coal}} < 50\,\mathrm{Myrs}$	24×

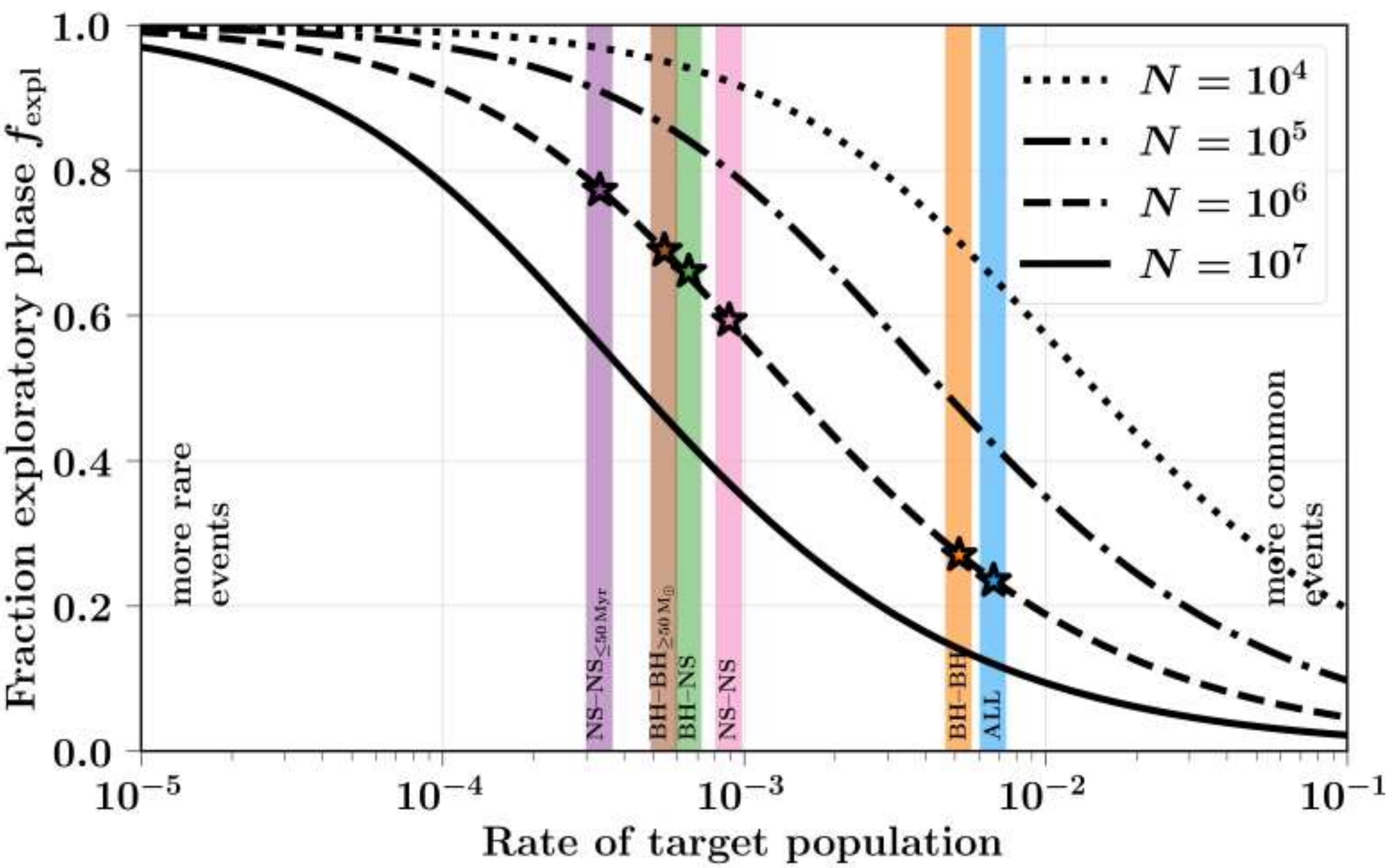
Floor Broekgaarden



Floor Broekgaarden

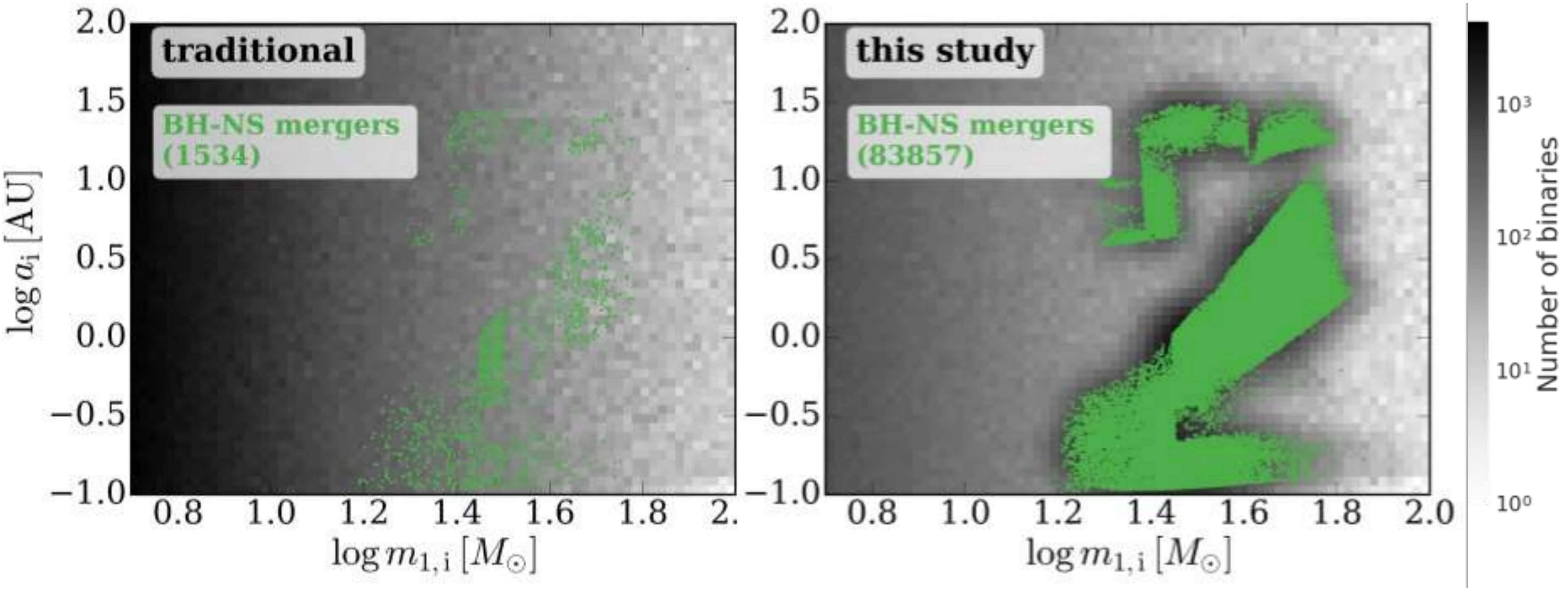


Floor Broekgaarden



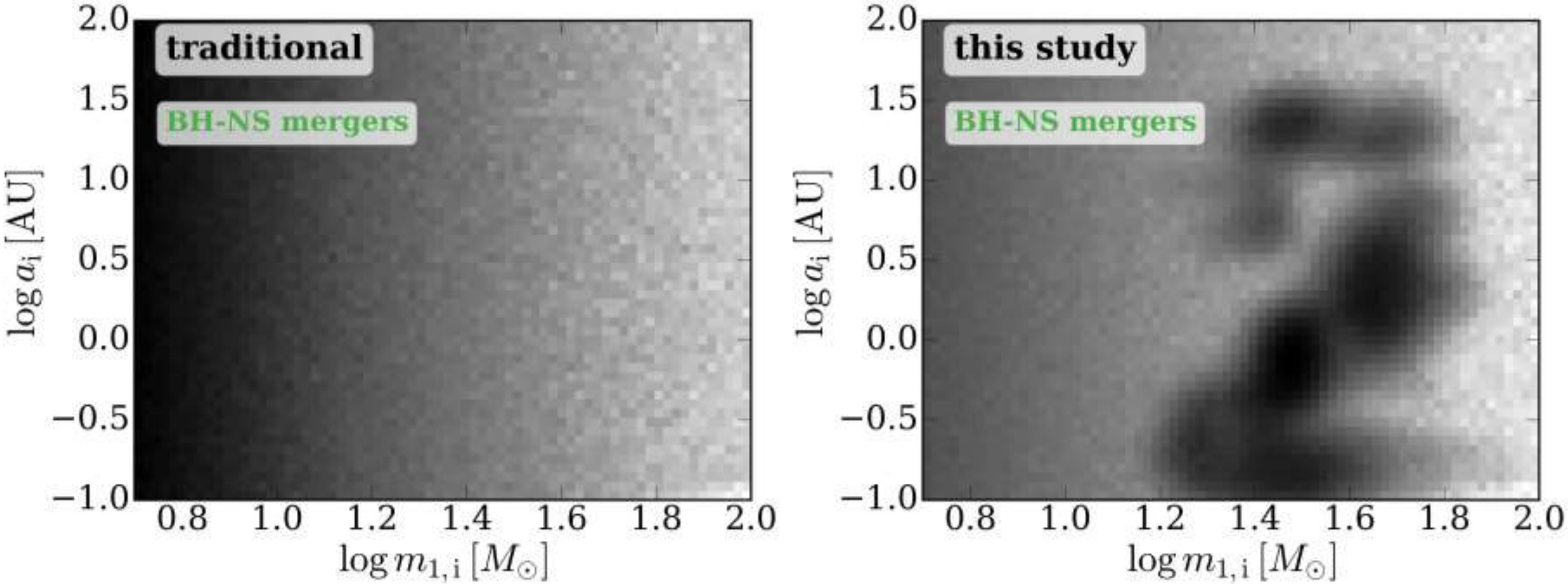
Floor Broekgaarden

Better resolution progenitors

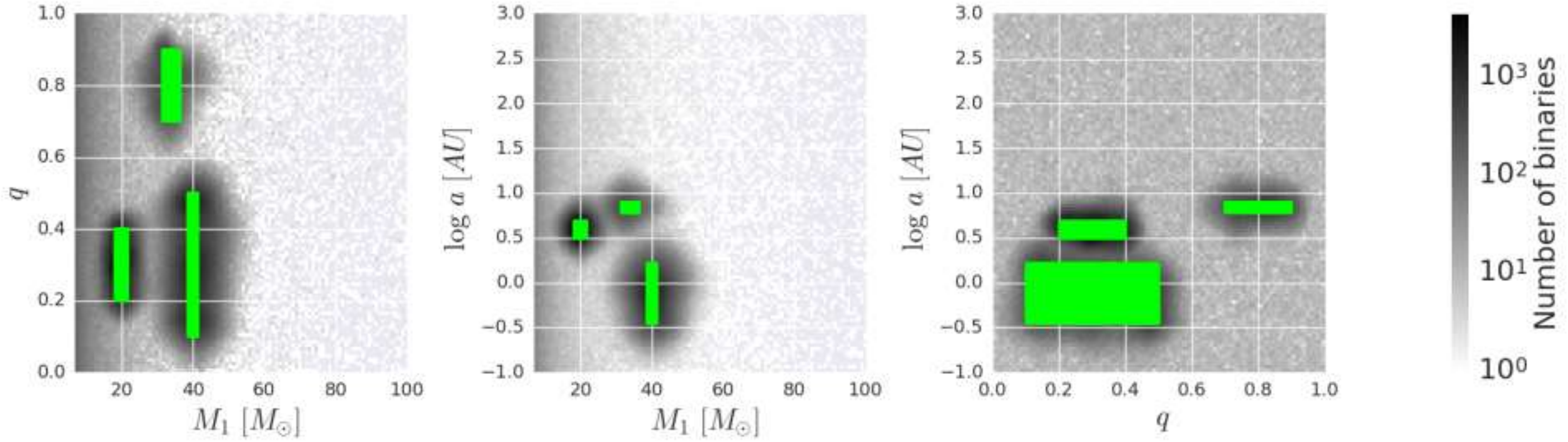


Floor Broekgaarden

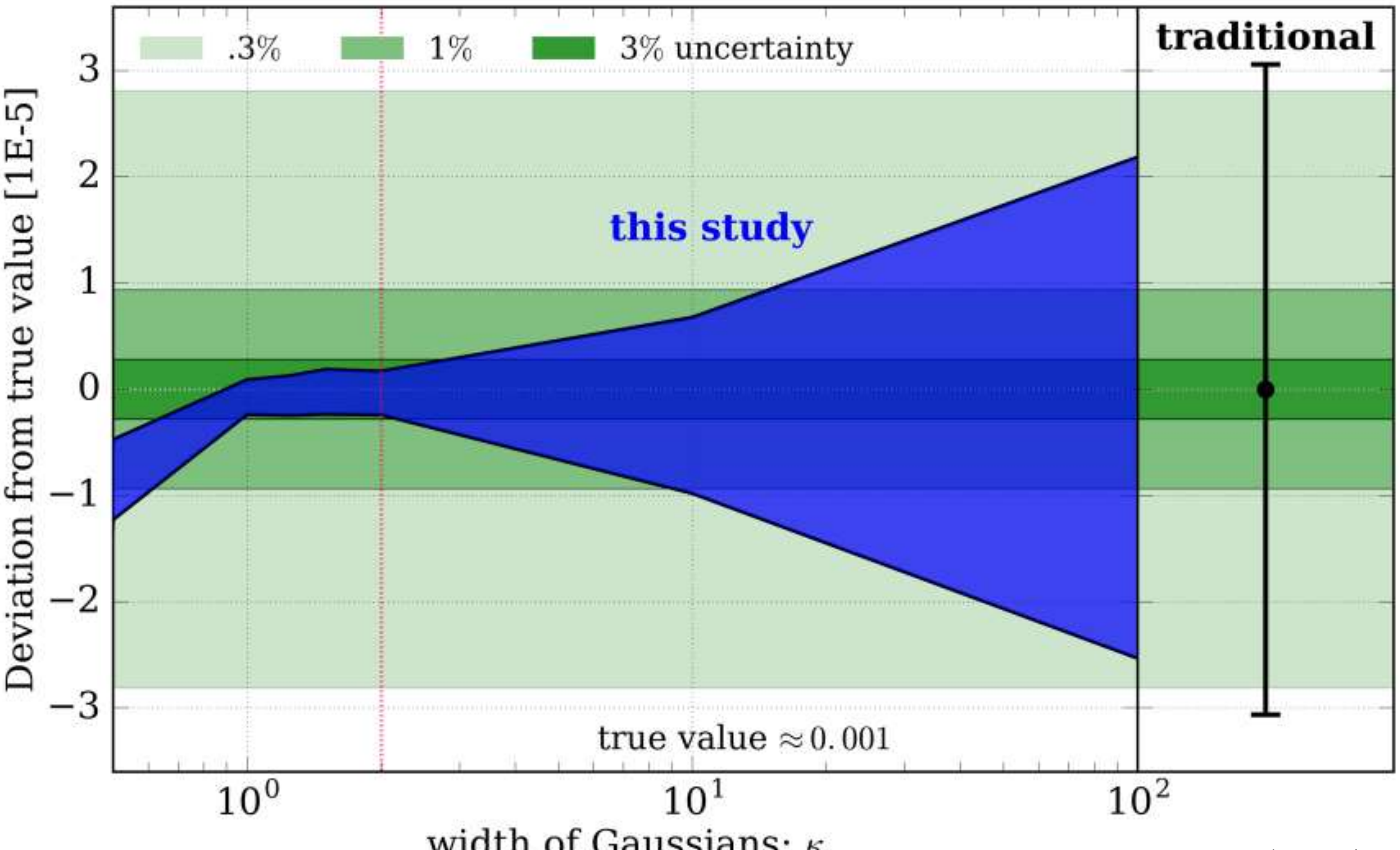
Refinement phase



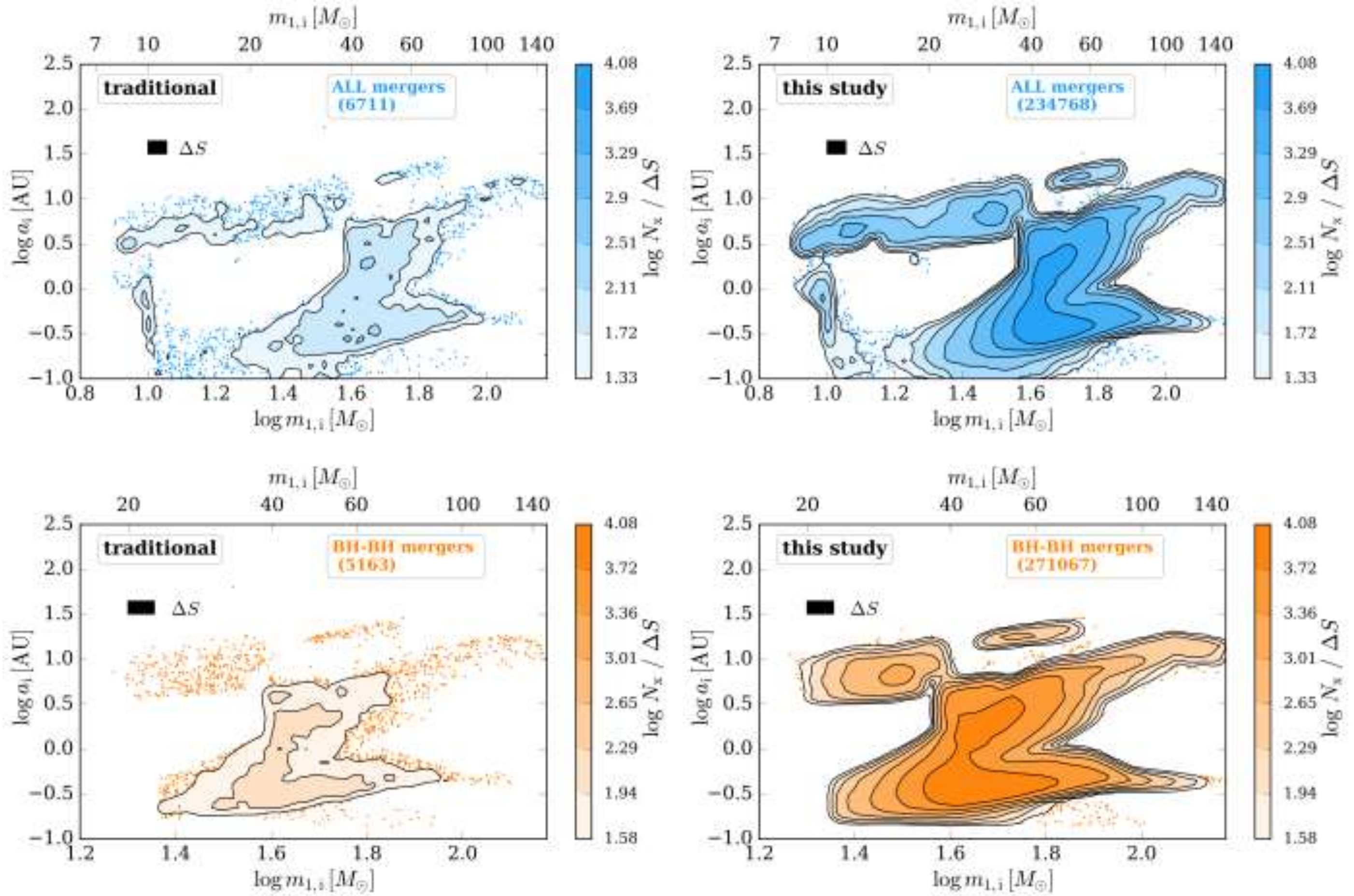
Floor Broekgaarden



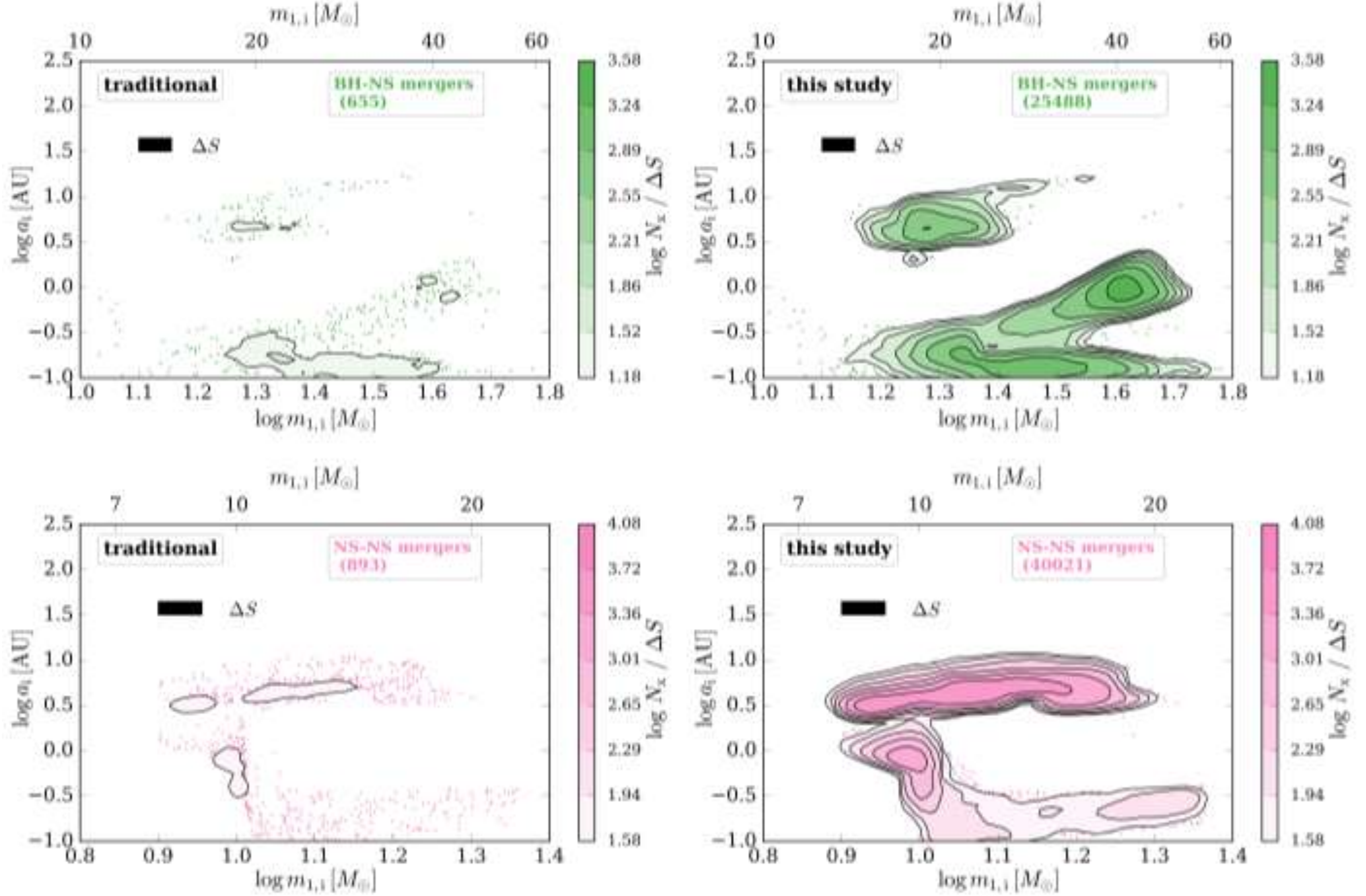
Floor Broekgaarden



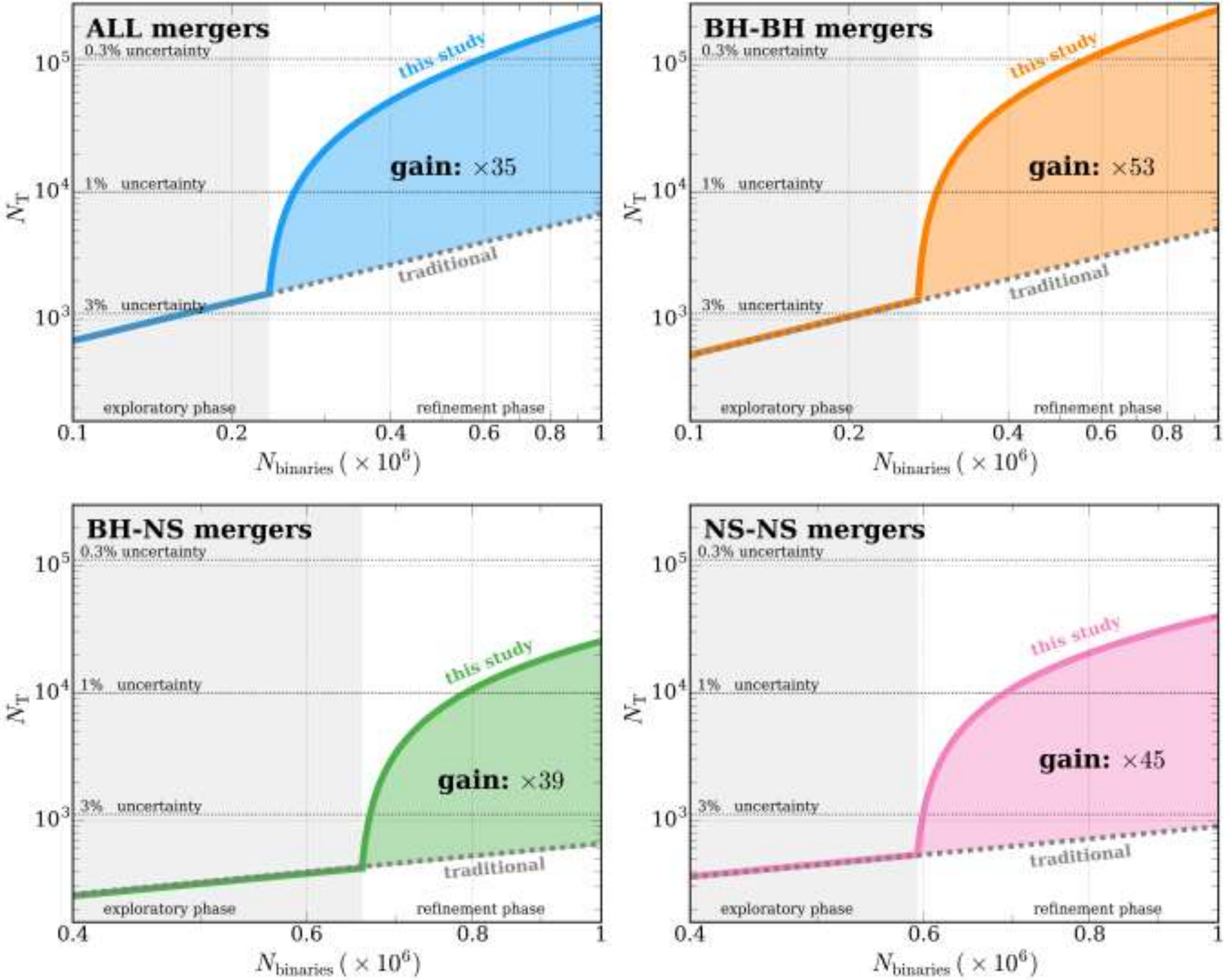
Floor Broekgaarden



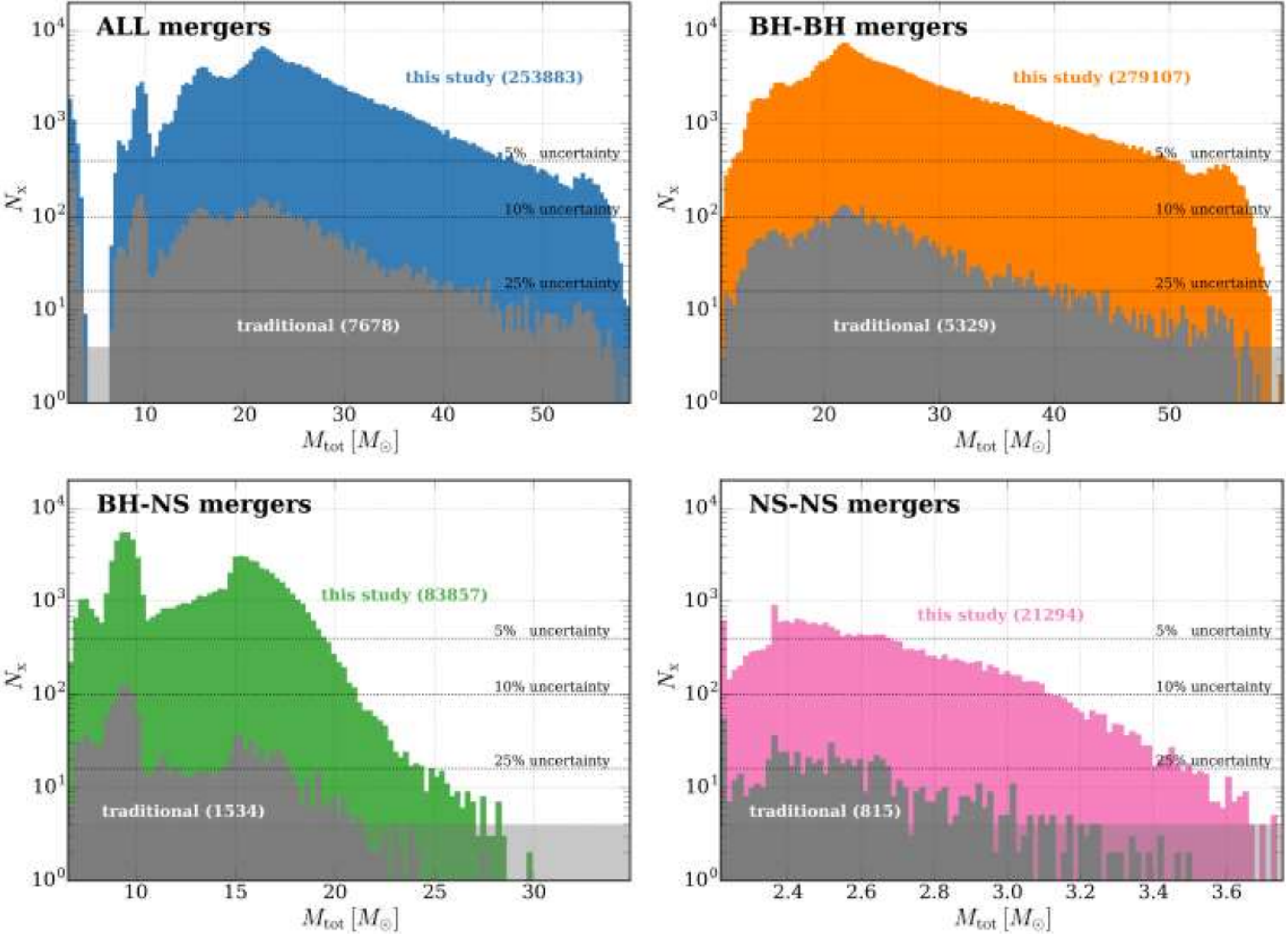
Floor Broekgaarden



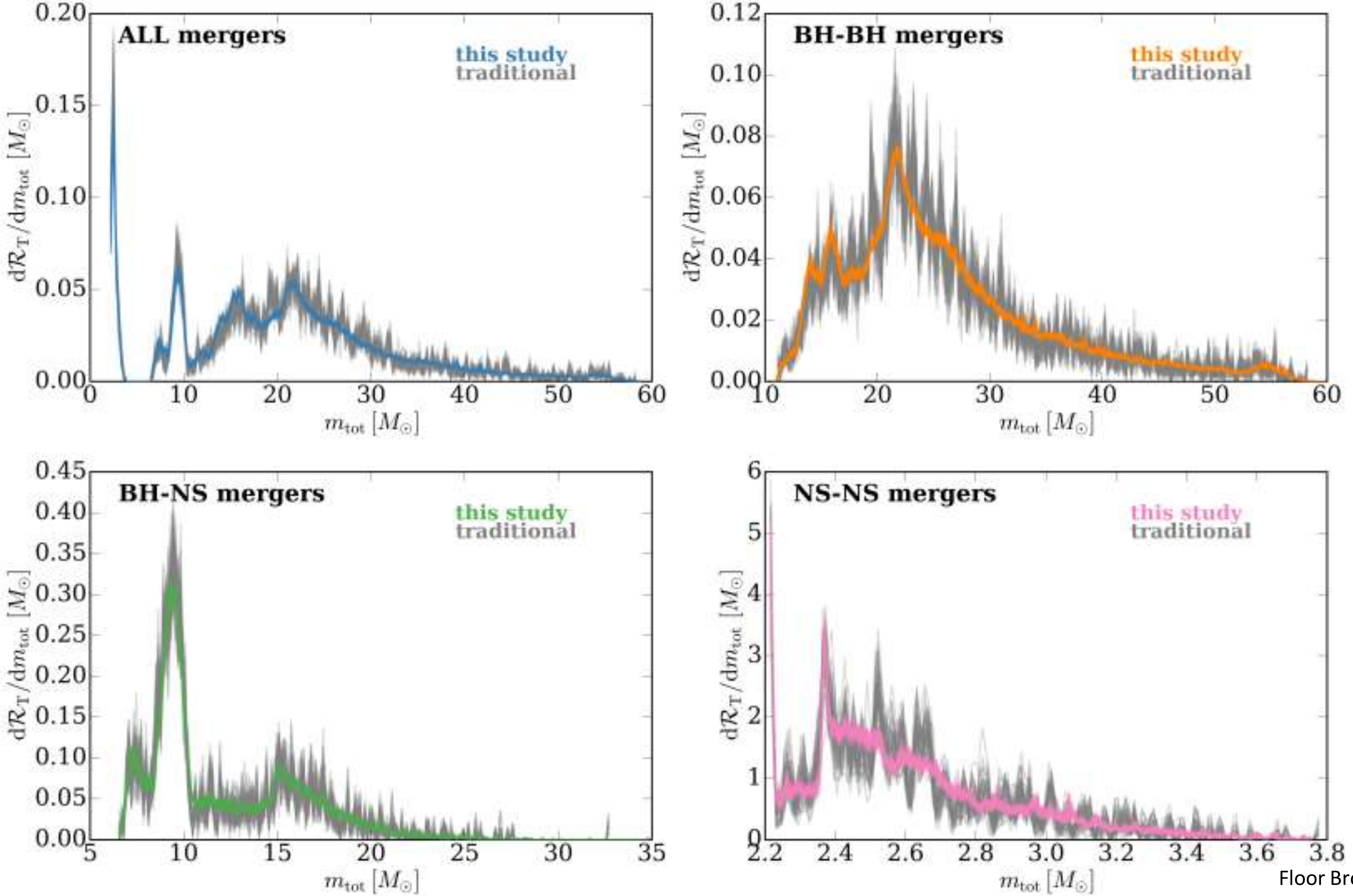
Floor Broekgaarden



ALL mergers
BH-BH mergers
BH-NS mergers
NS-NS mergers
0.3% uncertainty
1% uncertainty
3% uncertainty
this study
traditional
gain: ×35
gain: ×53
gain: ×39
gain: ×45
exploratory phase
refinement phase
N_{T}
$N_{\mathrm{binaries}}\ (\times 10^6)$

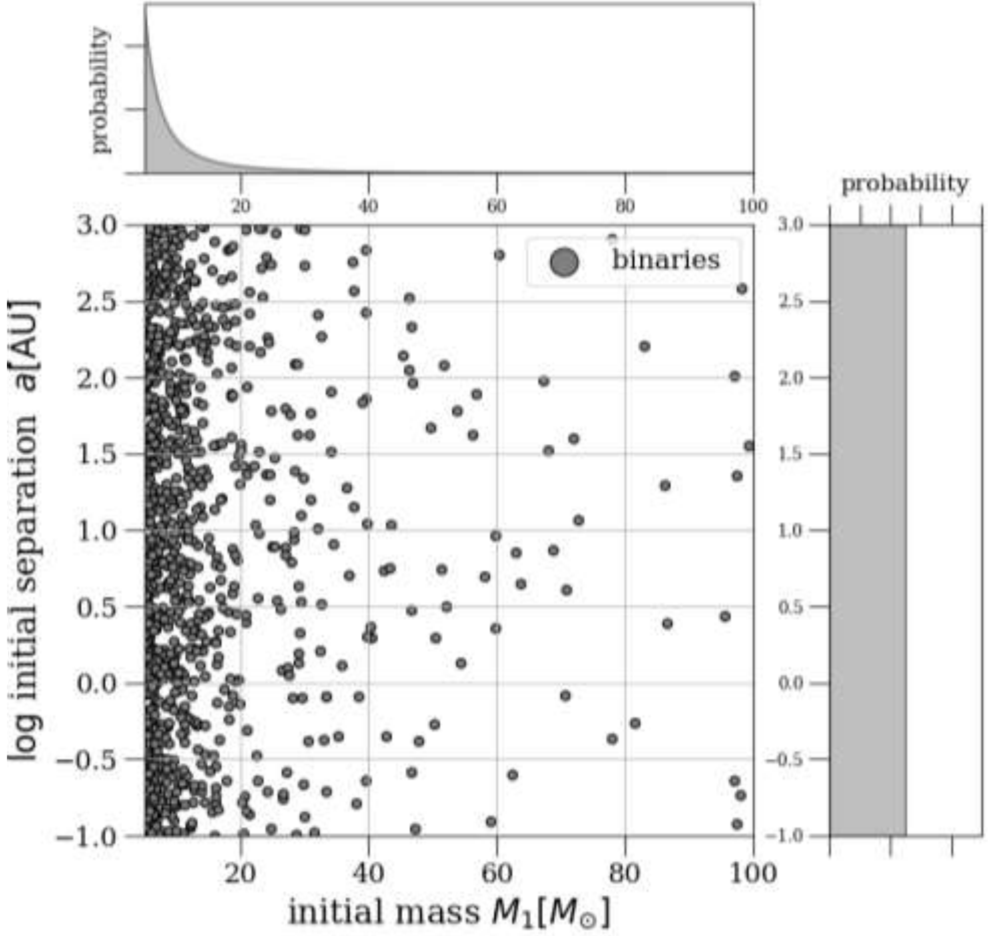


Floor Broekgaarden



Floor Broekgaarden

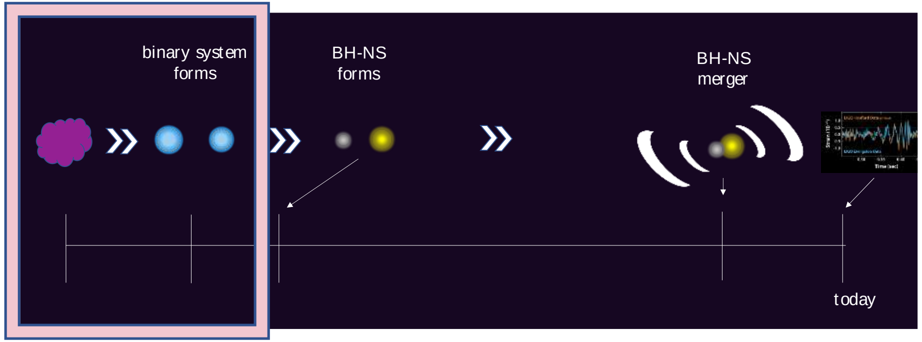
1 Populations:



E.g. Belczynski+02

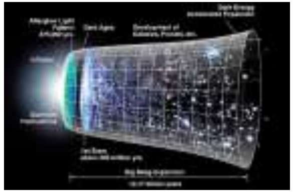
Two processes impact the predicted gravitational wave observations:

1. Star formation history over cosmic time?



Describes "model for the Universe":

$$\mathrm{MSSFR}(z_{\mathrm{form}}) = \underbrace{\frac{d^2 M_{\mathrm{SFR}}}{dt_s dV_c}(z_{\mathrm{form}})} \times \underbrace{\frac{dM_{\mathrm{SFR}}}{dZ}(z_{\mathrm{form}})},$$

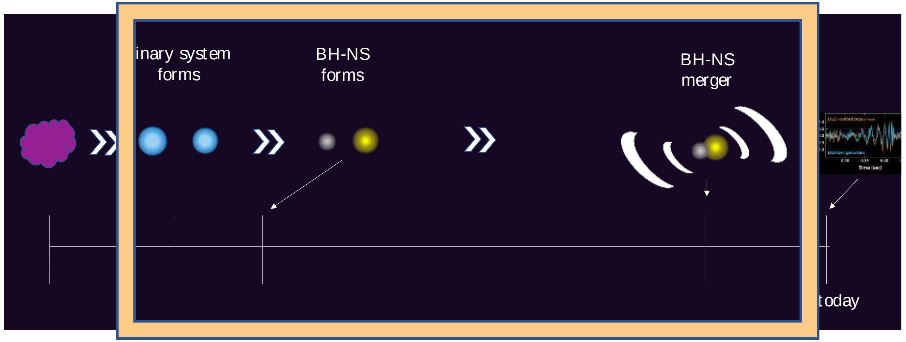


Star formation rate **SFR(z)**

Chemical evolution **Z(z)**

Two processes impact the predicted gravitational wave observations:

1. Star formation history over cosmic time?



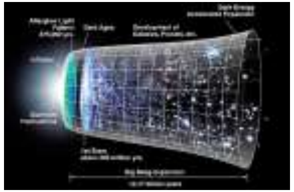
2. (Binary) stellar evolution?

Describes "model for the Universe":

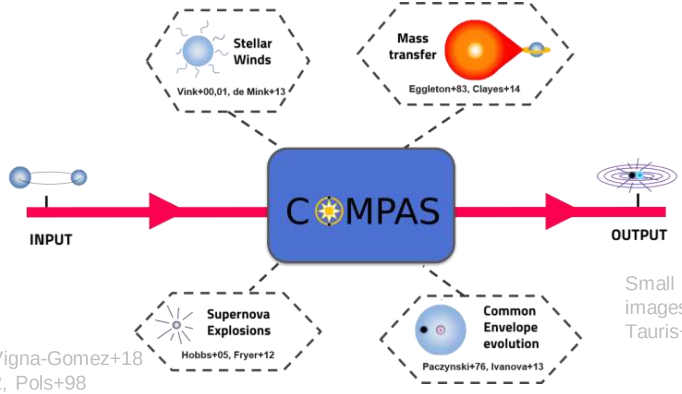
$$\mathrm{MSSFR}(z_{\mathrm{form}}) = \underbrace{\frac{d^2 M_{\mathrm{SFR}}}{dt_s dV_c}(z_{\mathrm{form}})}_{} \times \underbrace{\frac{dM_{\mathrm{SFR}}}{dZ}(z_{\mathrm{form}})}_{},$$

Star formation rate **SFR(z)**

Chemical evolution **Z(z)**



Describes "isolated binary evolution":



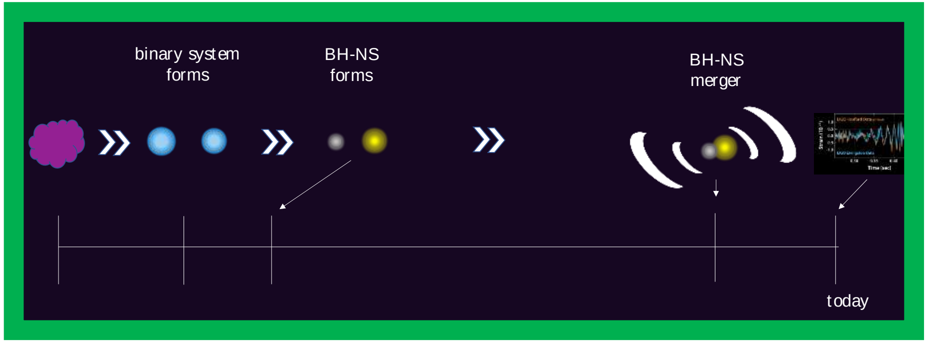
Small images: Tauris+17

cf Chruslinska+19, Belczynski+19 Neijssel+19, Tang+19...

Stevenson+15,17,19, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

We need to understand the combined effect!

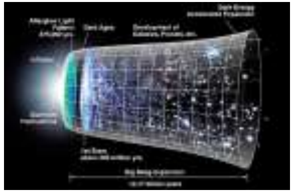
1. Star formation history over cosmic time?



2. (Binary) stellar evolution?

Describes “model for the Universe”:

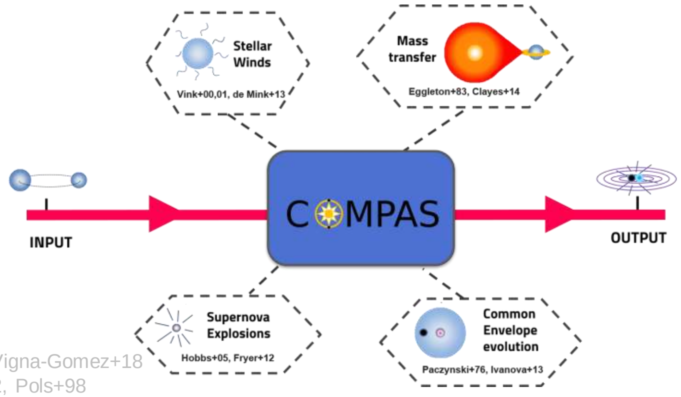
$$\mathrm{MSSFR}(z_{\mathrm{form}}) = \underbrace{\frac{d^2 M_{\mathrm{SFR}}}{dt_s dV_c}(z_{\mathrm{form}})} \times \underbrace{\frac{dM_{\mathrm{SFR}}}{dZ}(z_{\mathrm{form}})},$$



Star formation rate
SFR(z)

Chemical evolution **Z(z)**

Describes “isolated binary evolution”:



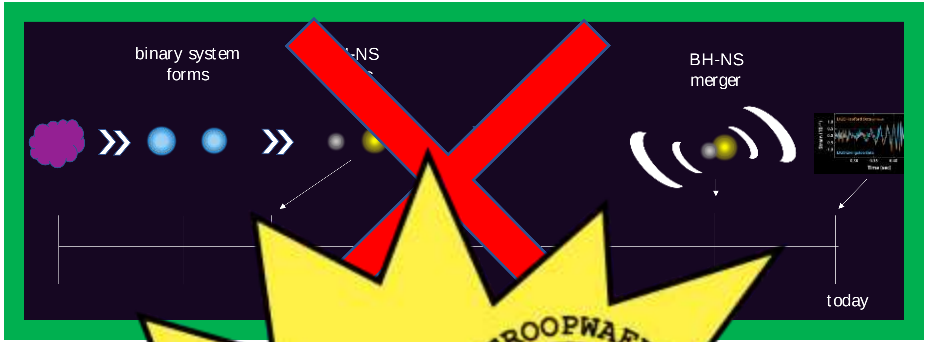
cf Chruslinska+19, Belczynski+19 Neijssel+19, Tang+19...

Stevenson+15,17,19, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

We need to understand the combined effect!

1. Star formation history over cosmic time?

2. (Binary) stellar evolution?



Describes

MSSFR(z...

Star fo...

Requires simulating ma...
metallicities...

New algorithm that speeds up simulations

100x

Broekgaarden+19

And billions of binaries...
= many TBs of data &
>20 000 000 CPU hours...

OUTPUT

Supernova Explosions
Hobbs+05, Fryer+12

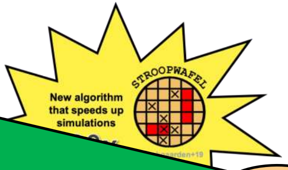
Common Envelope evolution
Paczynski+76, Ivanova+13

cf Chruslinska+19, Belczynski+19 Neijssel+19, Tang+19...

Stevenson+15,17,19, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

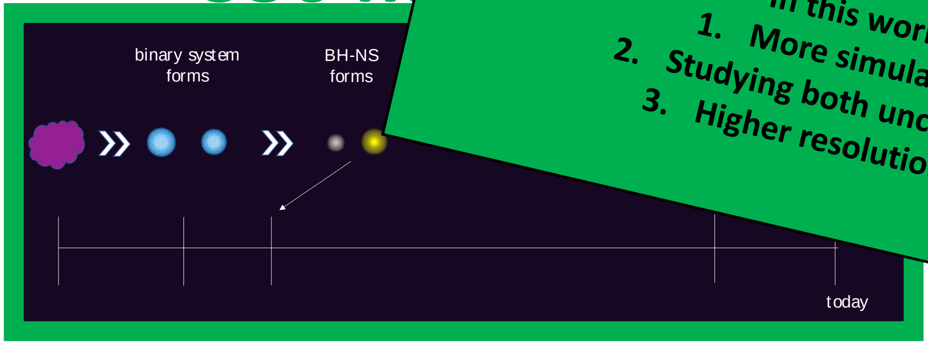
STROOPWAFEL all... to study the impact from b...

336 m...



1. Star formation history over cosmic time?

...) stellar

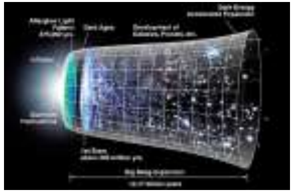


In this work:
1. More simulations
2. Studying both uncertainties
3. Higher resolution data

28 models:

based on Neijssel+19

$$\mathrm{MSSFR}(z_{\mathrm{form}}) = \underbrace{\frac{d^2 M_{\mathrm{SFR}}}{dt_s dV_c}(z_{\mathrm{form}})} \times \underbrace{\frac{dM_{\mathrm{SFR}}}{dZ}(z_{\mathrm{form}})}.$$



Star formation rate
SFR(z)

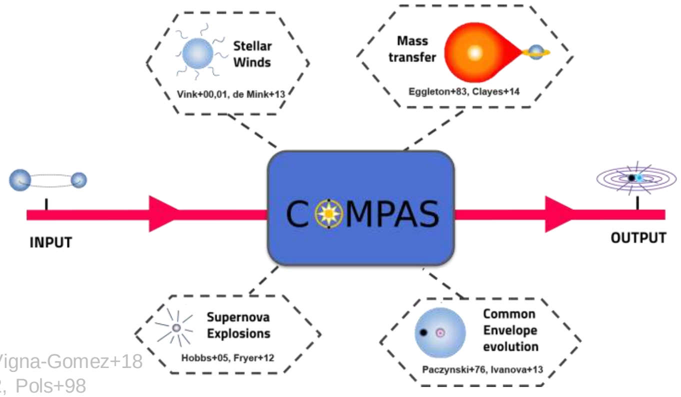
Madau & Dickinson 2014
Strolger +2004
Madau & Fragos 2017

Chemical evolution **Z(z)**

Mass-Z relation:
Langer & Norman 2006 (+offset)
Ma+2016

Galaxy-Stellar-Mass-function:
Panter+2004,
Furlong (single/double Schechter)+2015

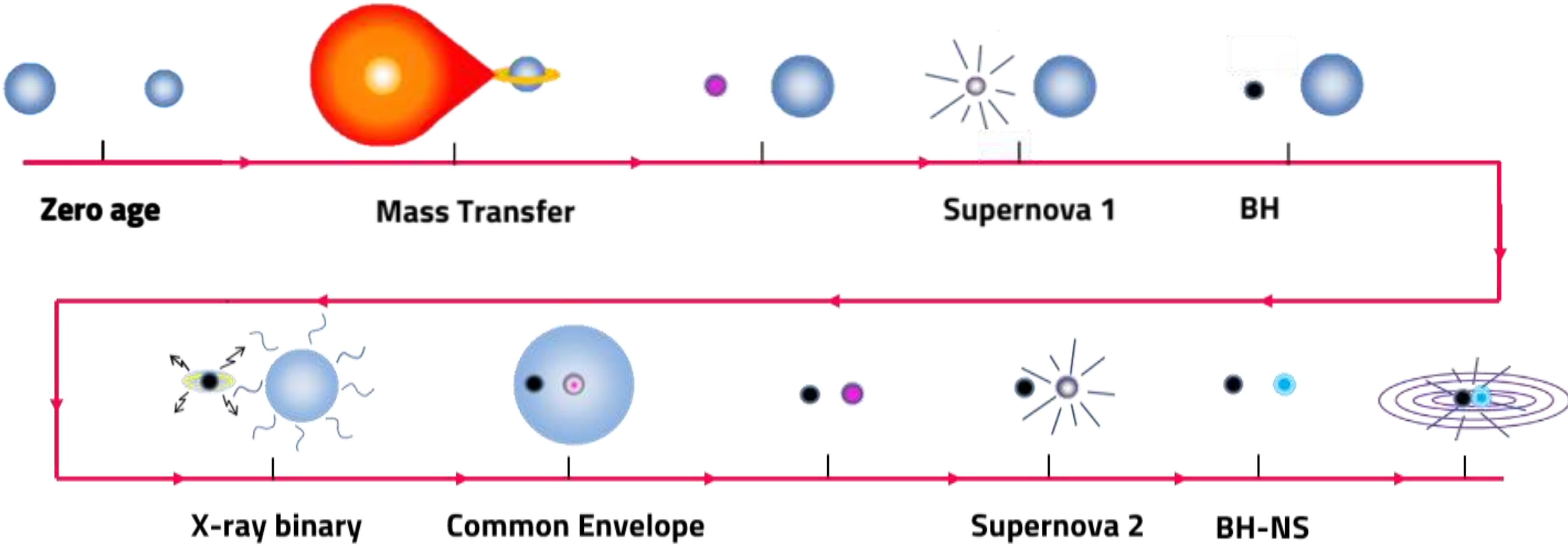
12 models:



Stevenson+15,17,19, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

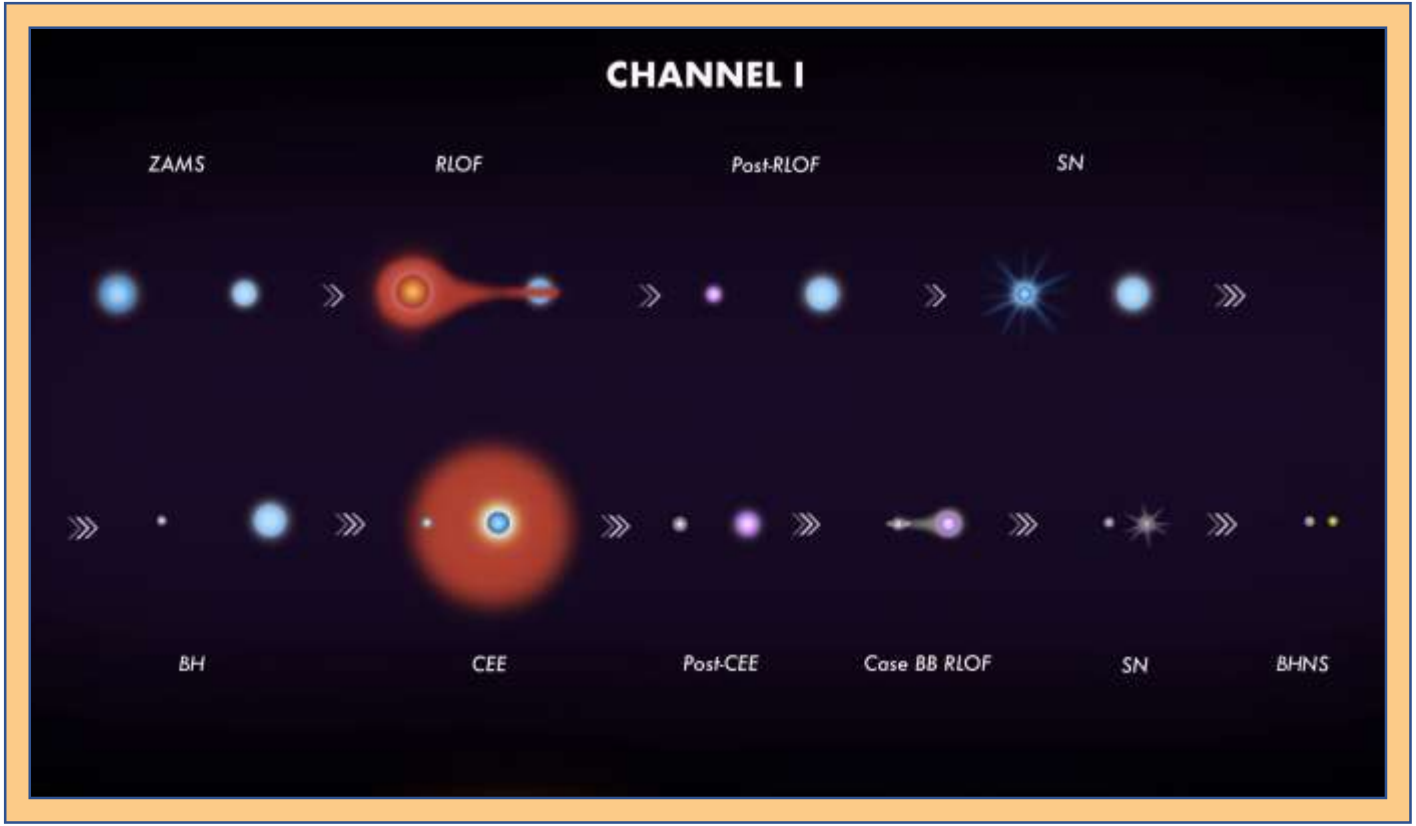
Classic channel BH-NS merger:

e.g. Paczynski+76, Smarr & Blandford+76 | Figure based on Tauris+17



Floor Broekgaarden

Modelling gravitational-wave sources



Floor Broekgaarden

How do they form?

$z = 0$

Image not to scale!

e.g. Paczynski+76, Smarr & Blandford+76

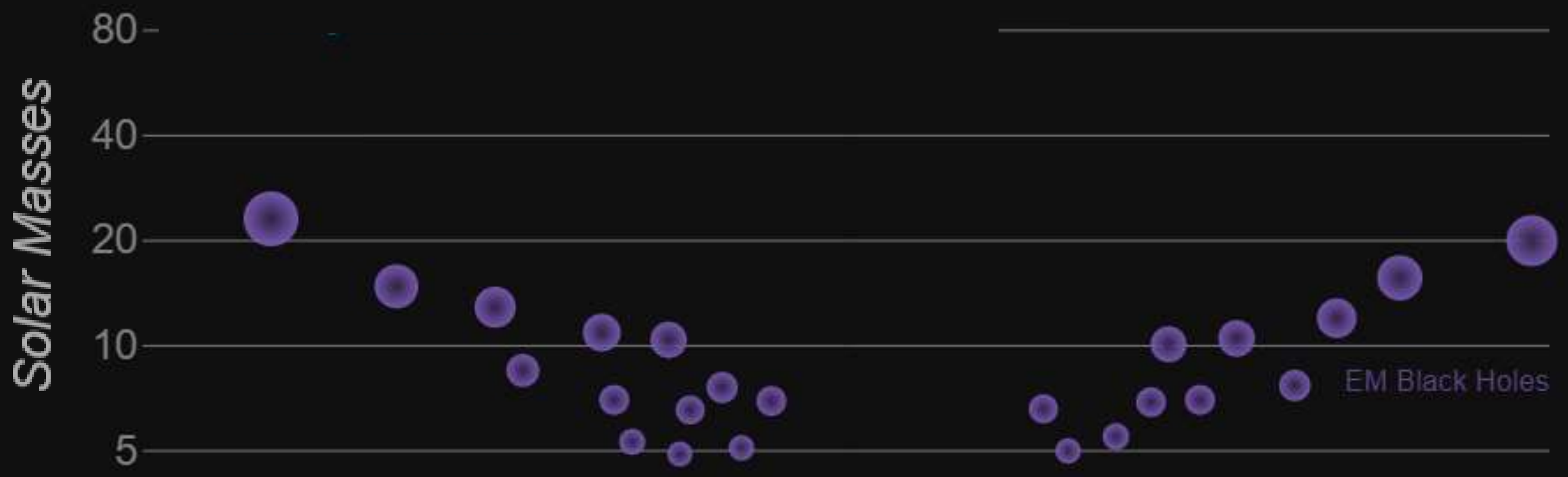
Oher puzzling methods

Q: How many stellar black holes are thought to exist in our Milky Way at this moment?

~100 000 000

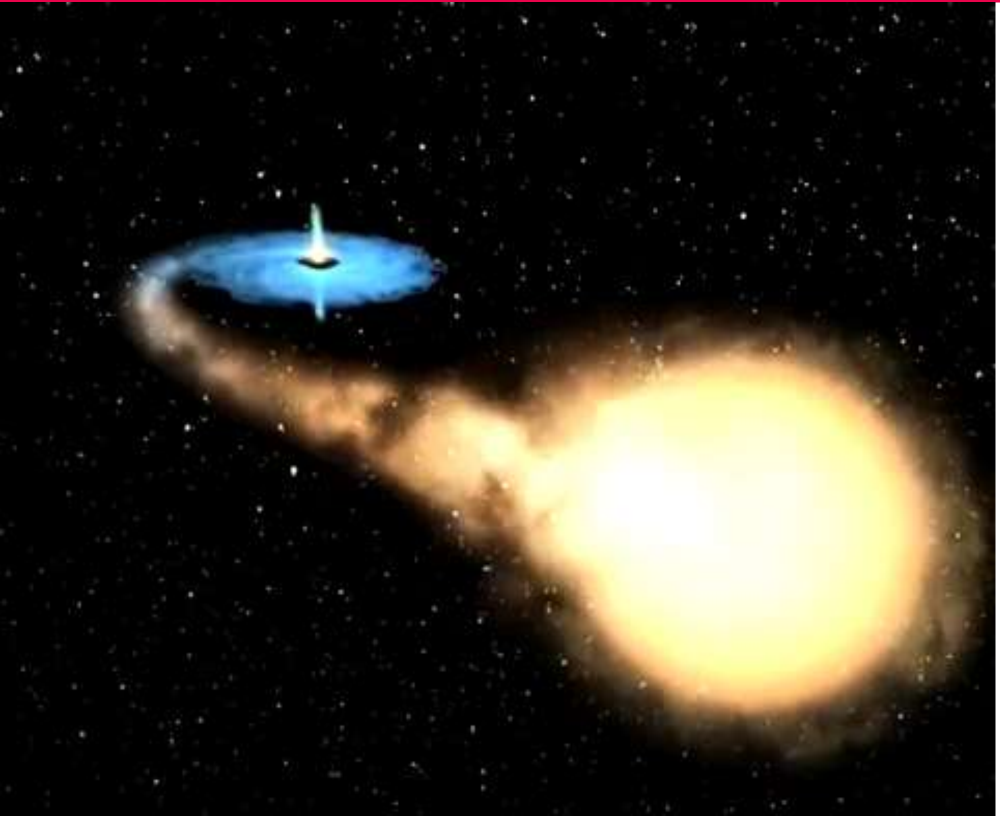
~50 000 000

Up to 2015:
~20 stellar BHs detected with known masses

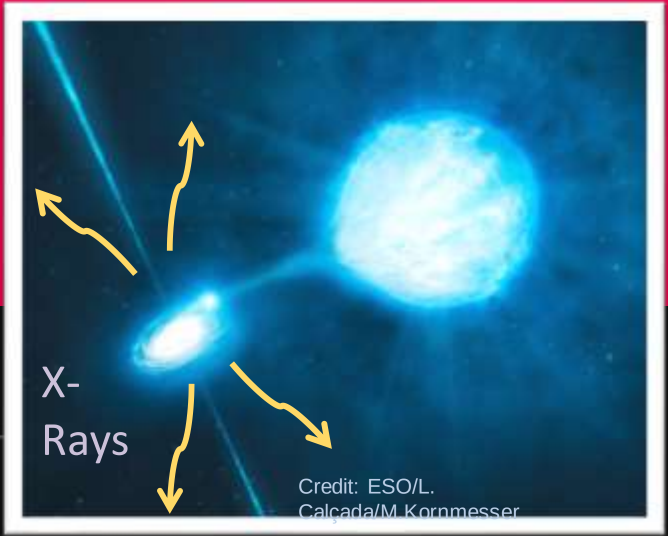


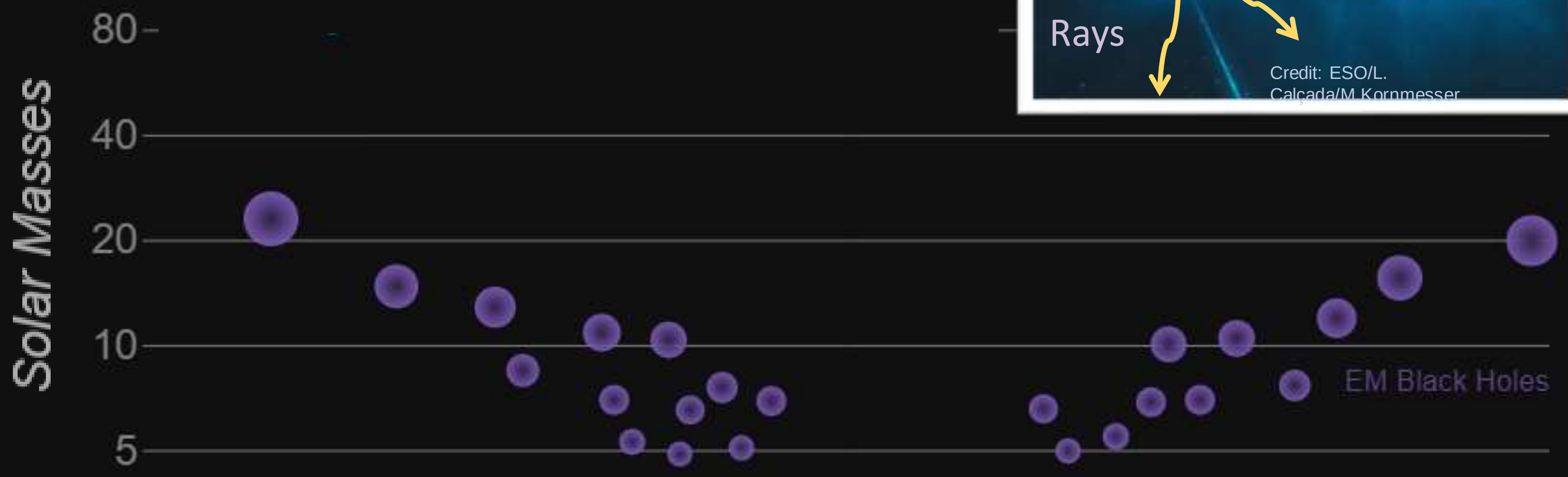
Q: The first detection of a stellar mass black hole in 1971 (through X-rays)

Bolton +72, Webster and Murdin+72

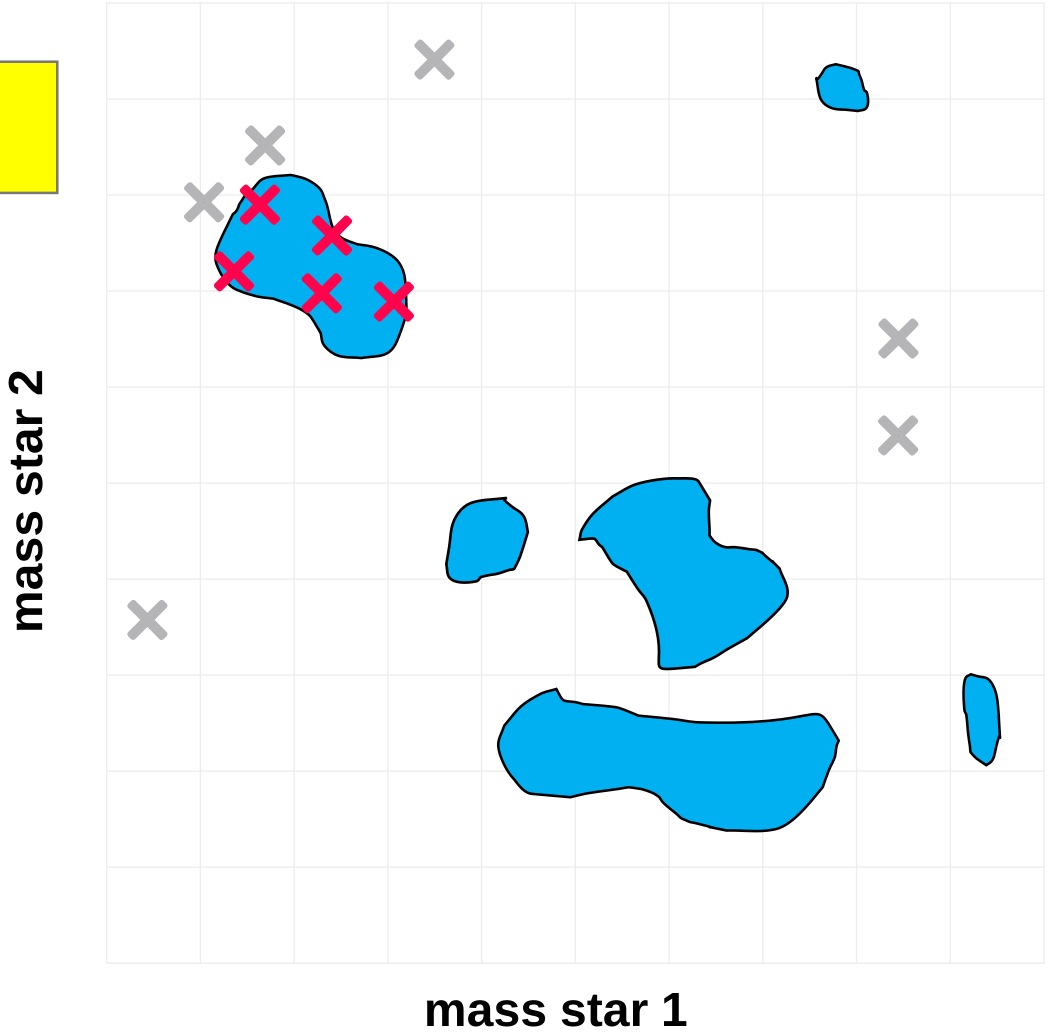


Up to 2015:
~20 stellar BHs detected with known masses

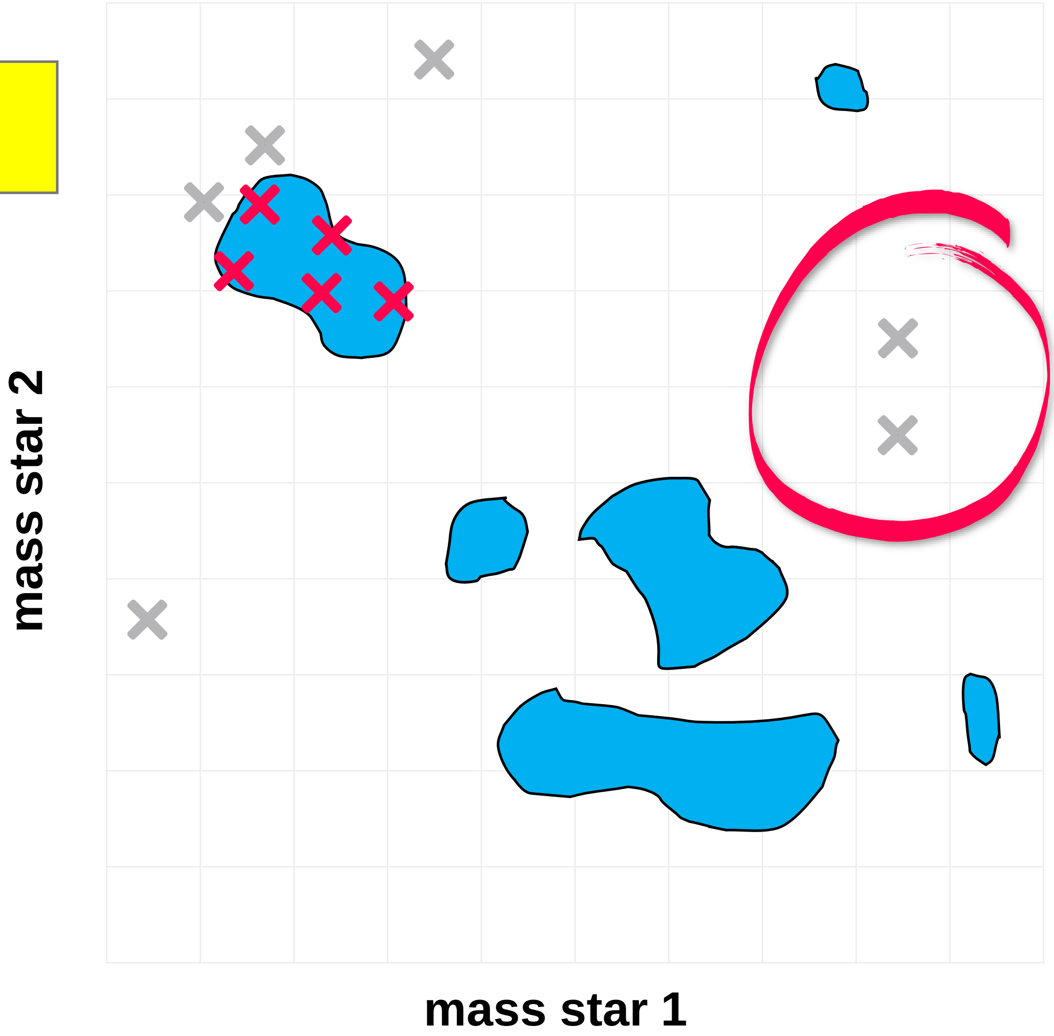




1. “Adaptive learning”

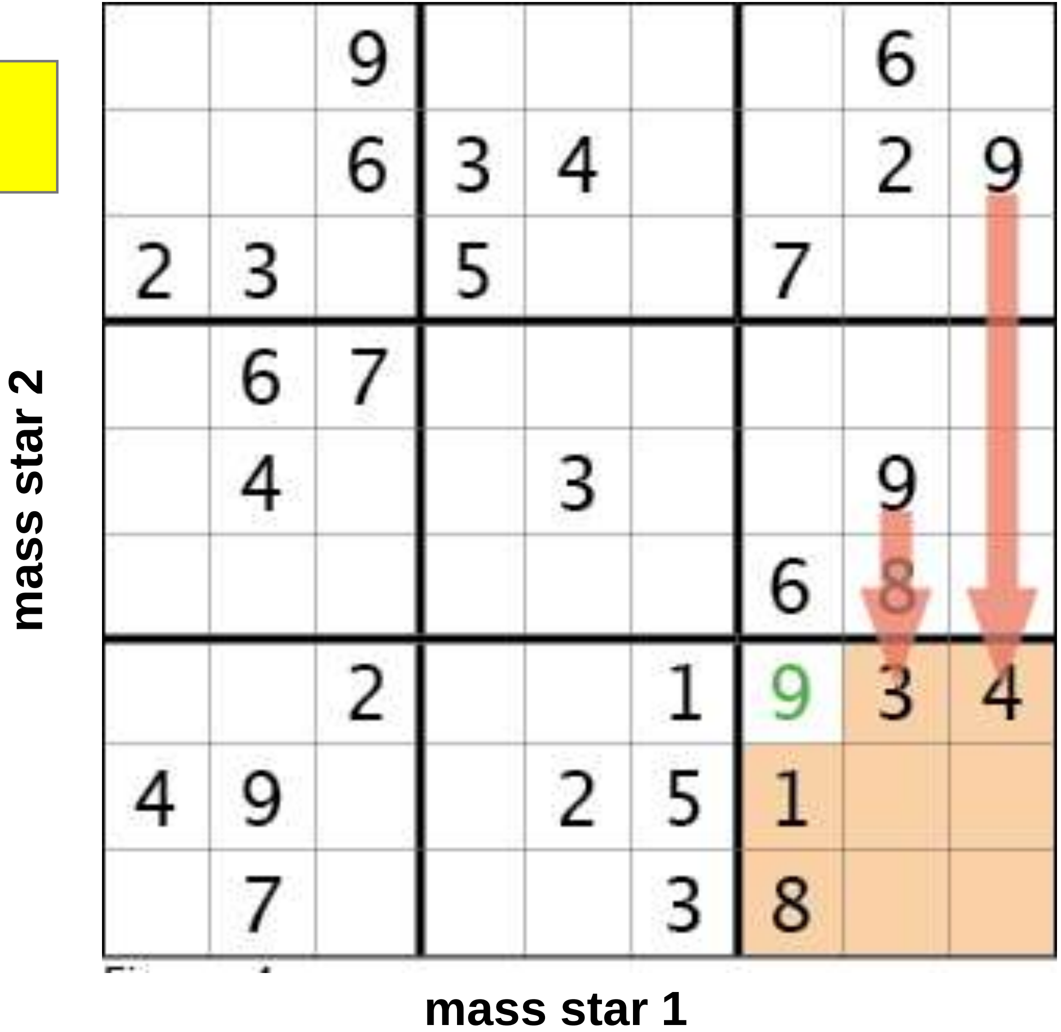


1. "*Adaptive learning*"



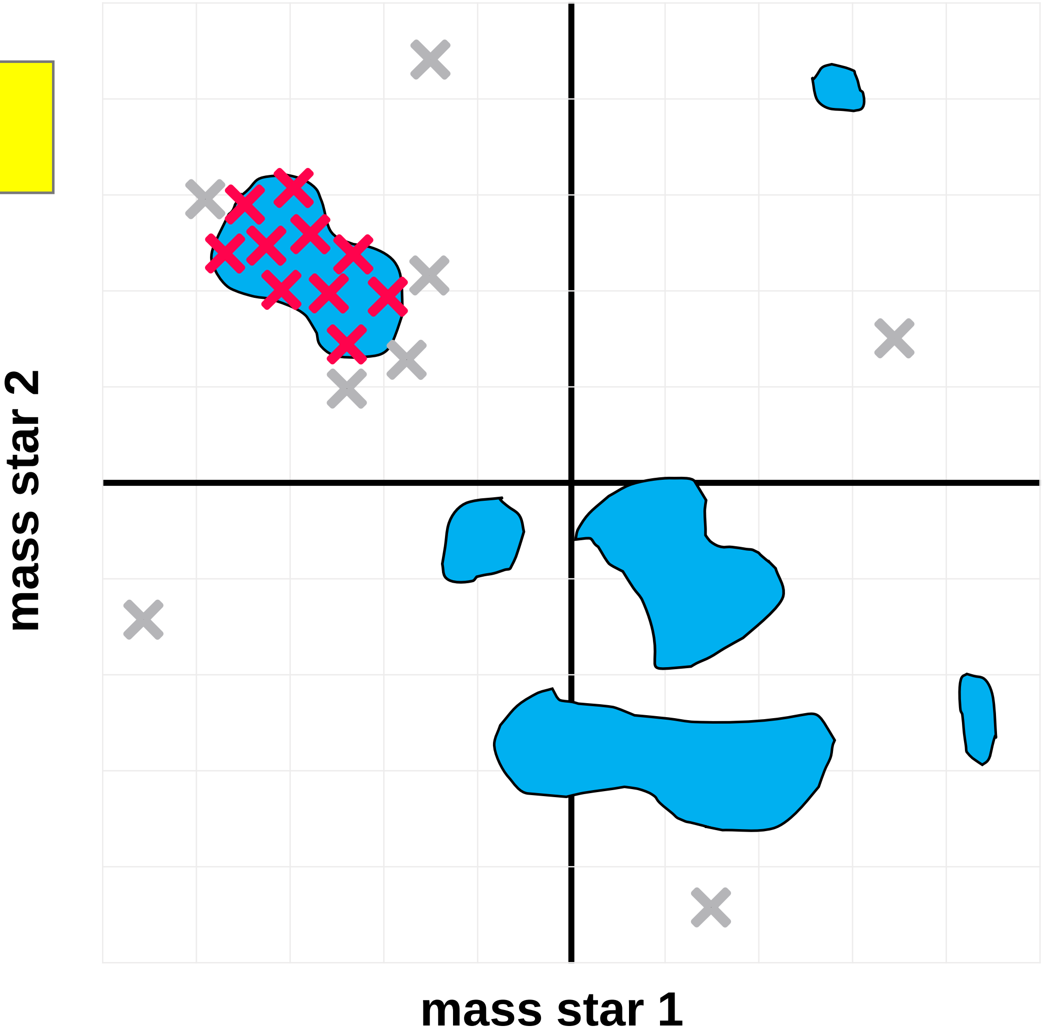
1. “Adaptive learning”

2. Latin hypercube (“Sudoku”) sampling



1. "Adaptive learning"

2. Latin hypercube ("Sudoku") sampling



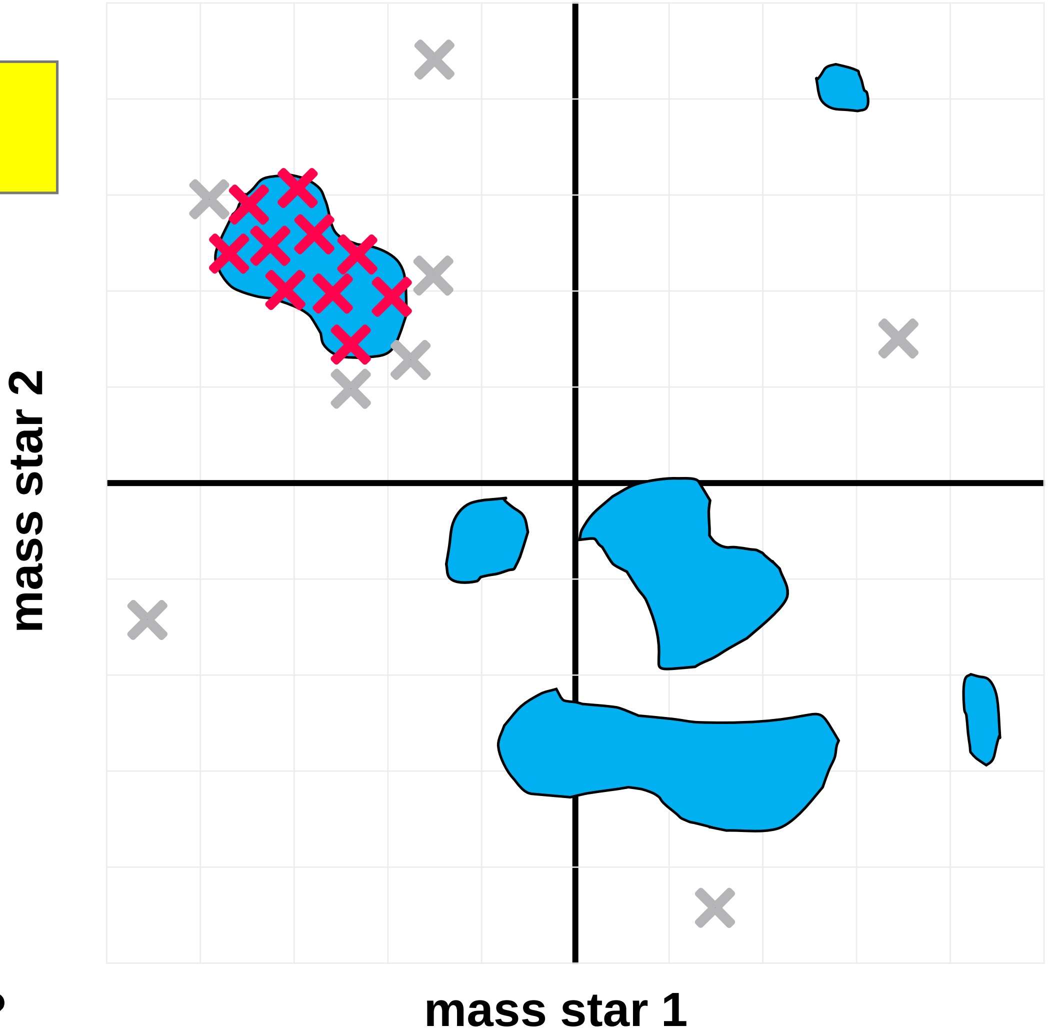
1. “Adaptive learning”

2. Latin hypercube (“Sudoku”) sampling

requires

<? >

to complete the game?



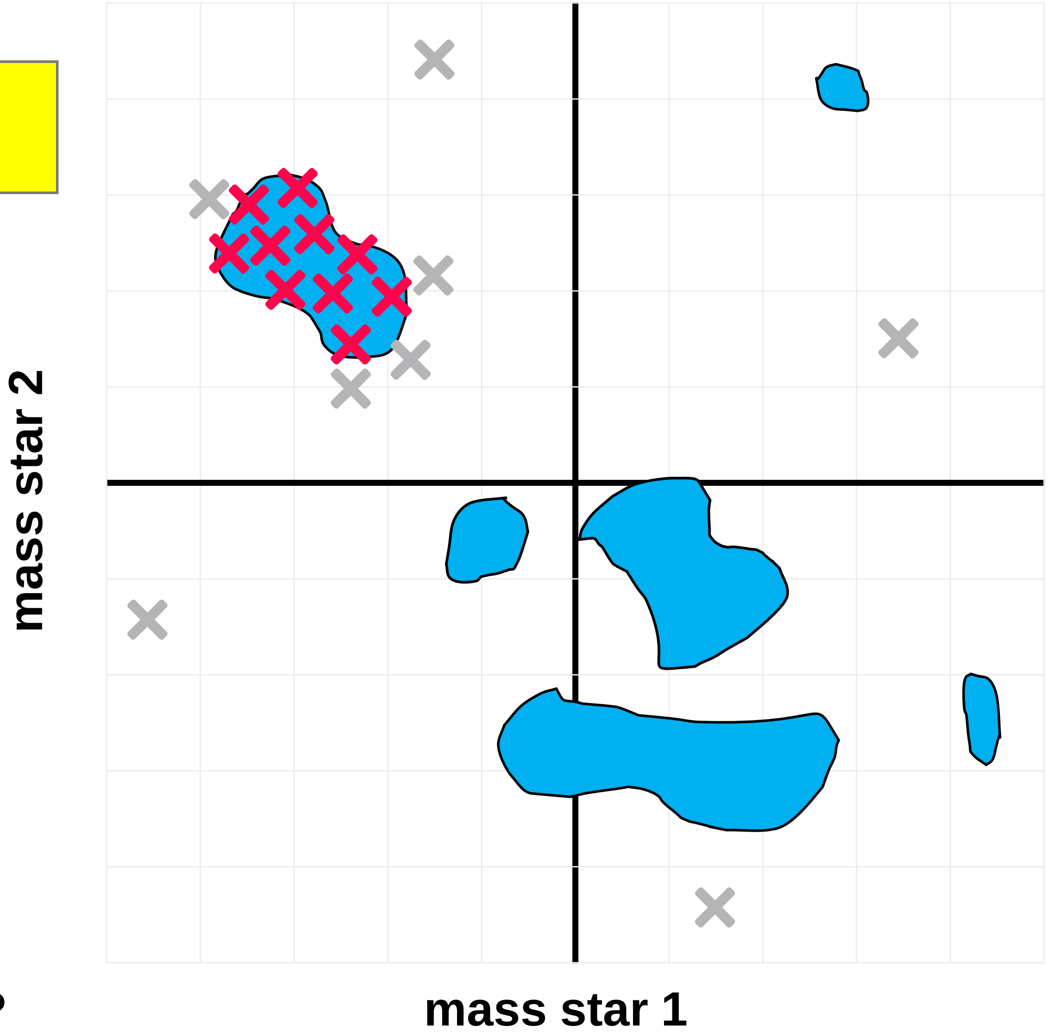
1. *“Adaptive learning”*

2. *Latin hypercube (“Sudoku”) sampling*

requires

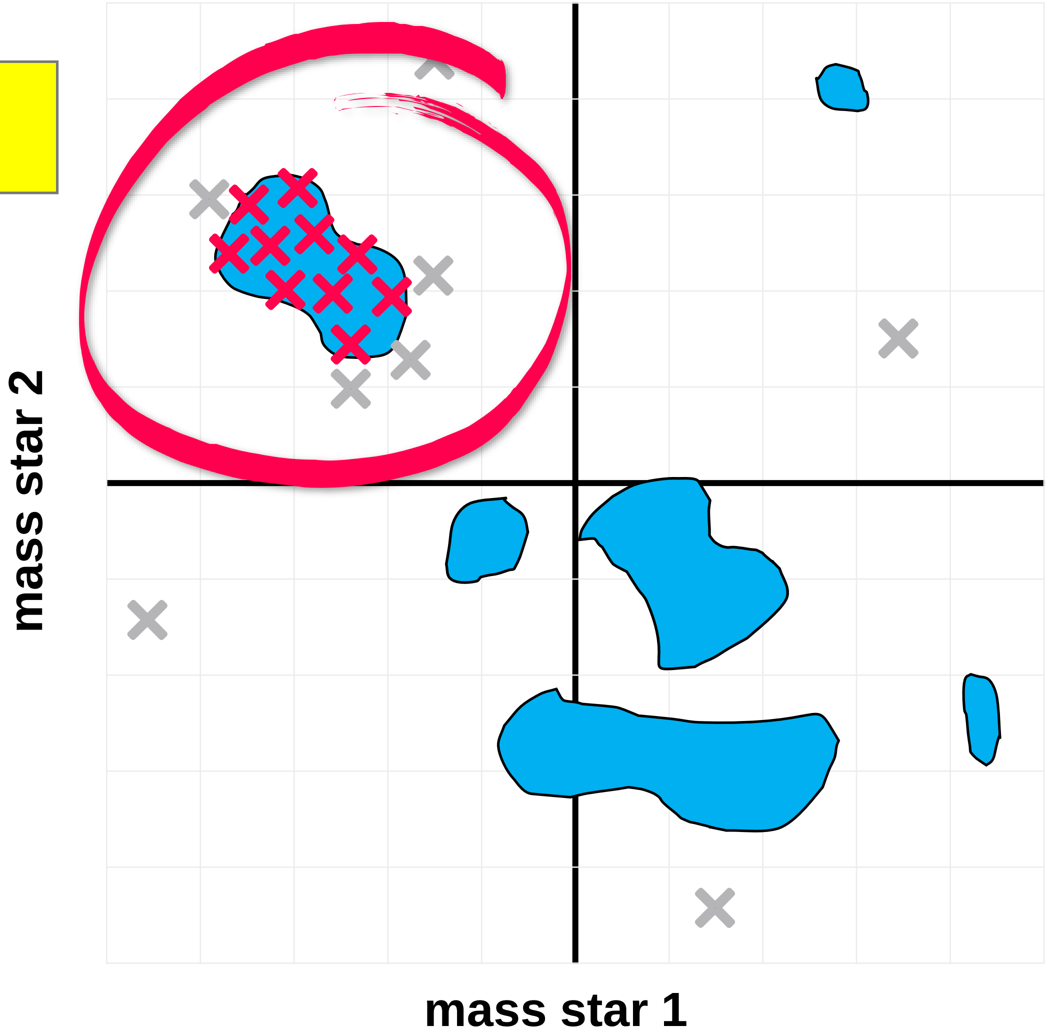
< 61 >

to complete the game?



1. “Adaptive learning”

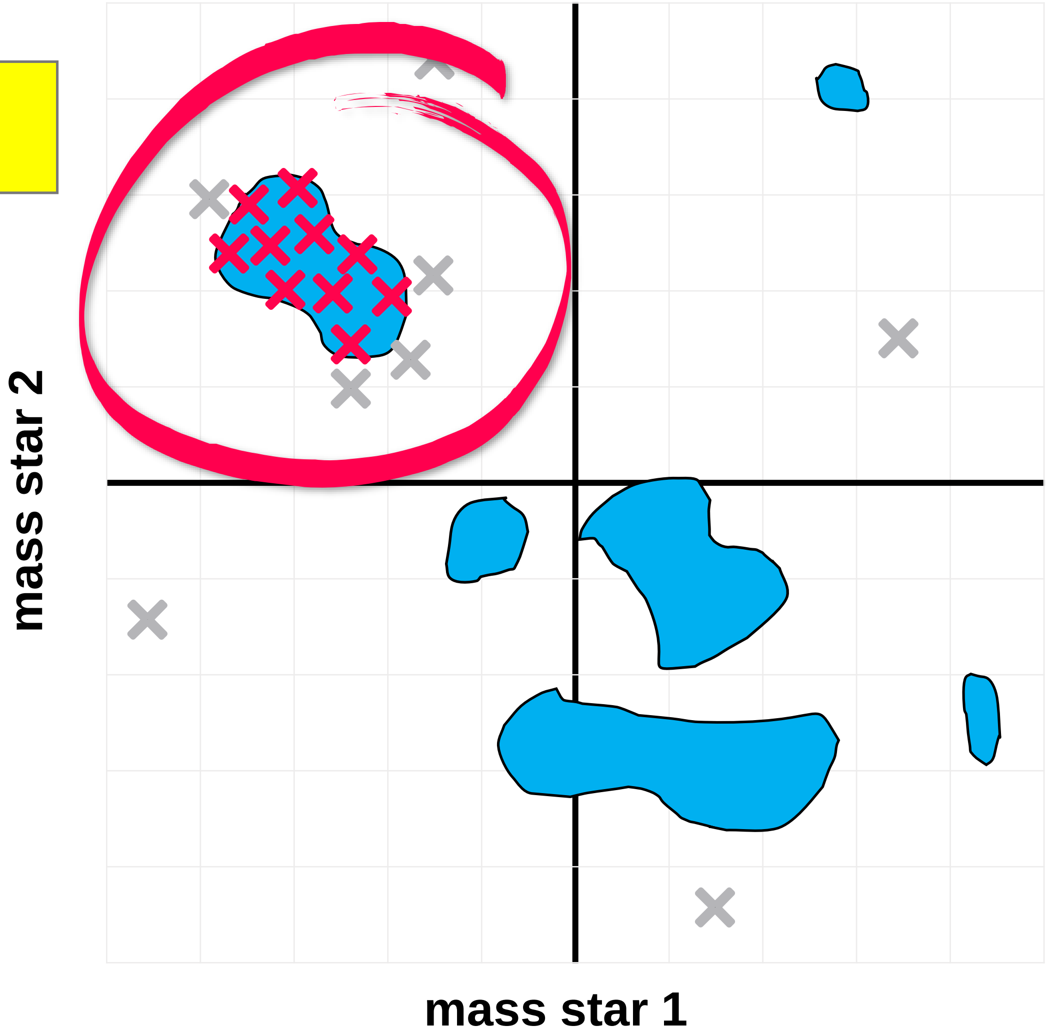
2. Latin hypercube (“Sudoku”) sampling



1. *"Adaptive learning"*

2. *Latin hypercube ("Sudoku") sampling*

3. *Predict outcome in between ("Emulators")*

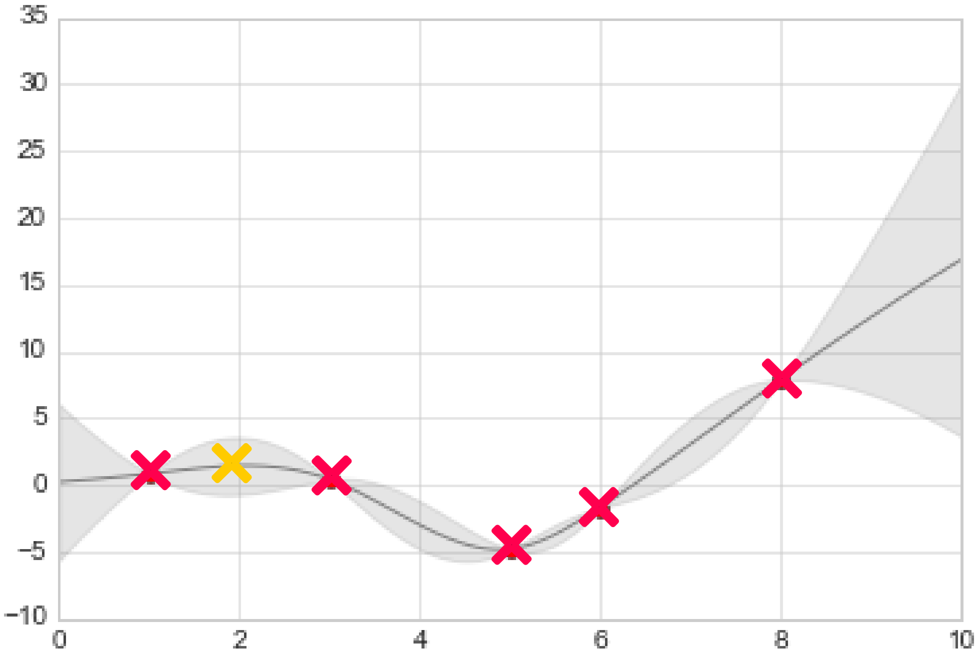


1. *“Adaptive learning”*

2. *Latin hypercube (“Sudoku”) sampling*

3. *Predict outcome in between (“Emulators”)*

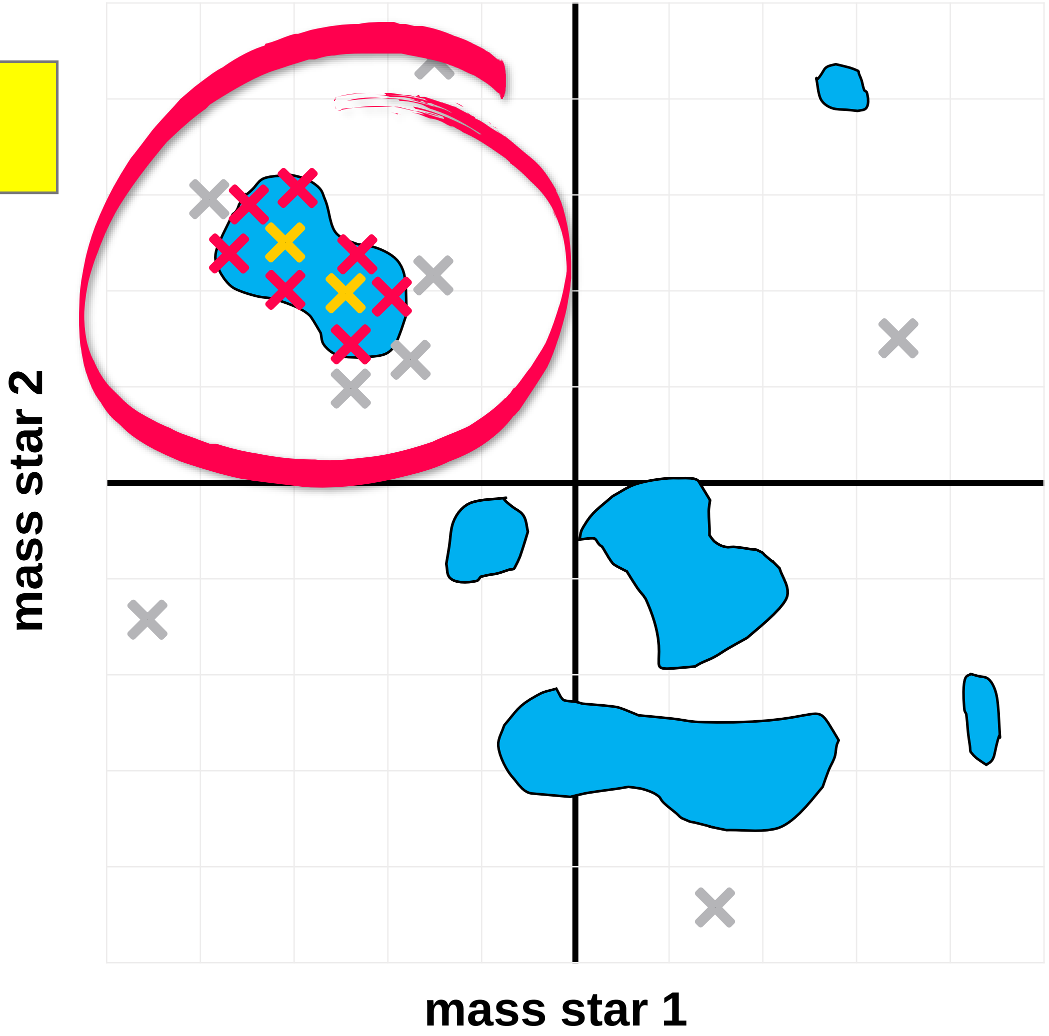
Emulators



1. *“Adaptive learning”*

2. *Latin hypercube (“Sudoku”) sampling*

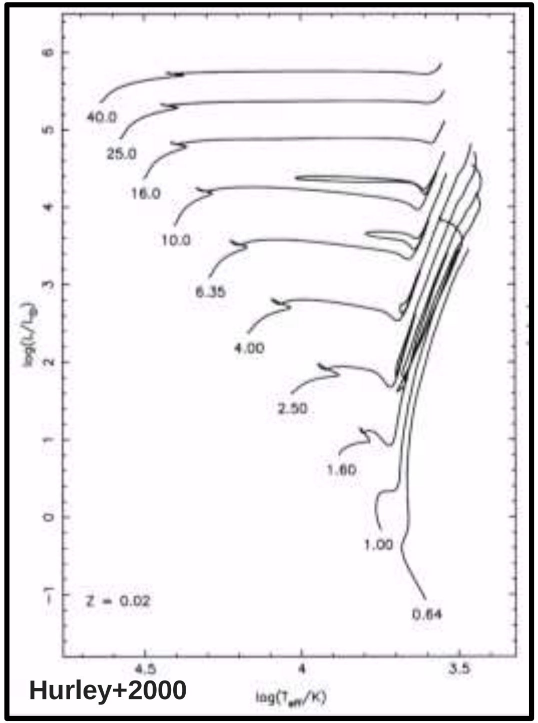
3. *Predict outcome in between (“Emulators”)*



2 ways to implement (single) stellar evolution:

“Full” stellar calculations

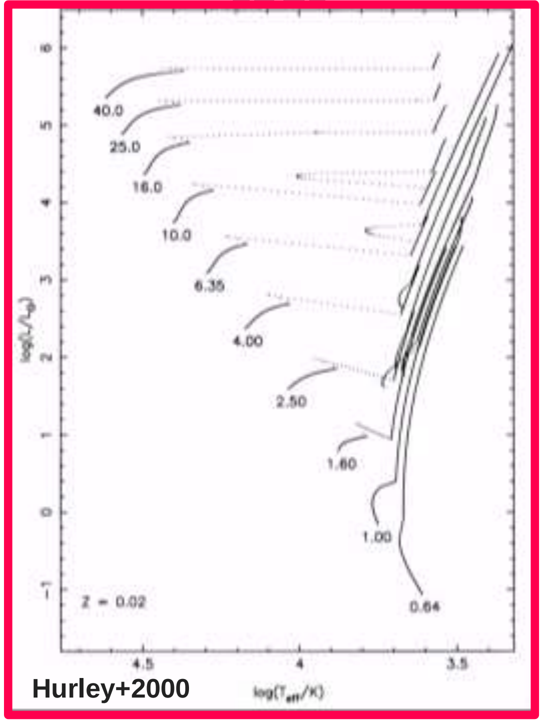
relatively slow



e.g. MESA

Analytical fits or interpolations

fast



e.g. StarTrack, binary_c, COMPAS