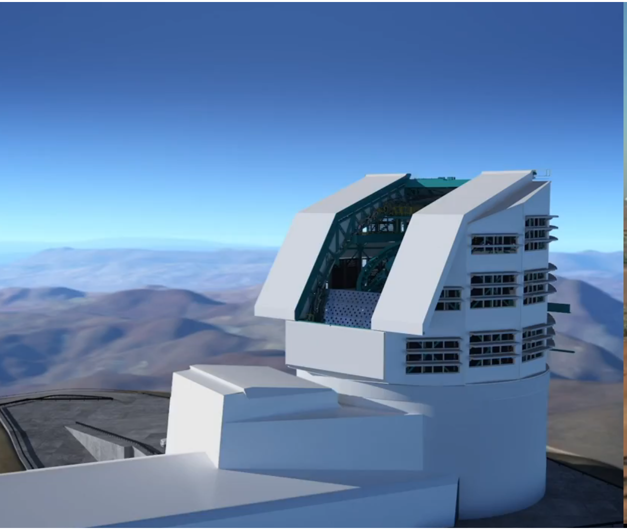
STROOPWAFEL: a Ducth cookie and an Adaptive Importance Sampling algorithm

Floor Broekgaarden
Center for Astrophysics | Harvard & Smithsonian

In collaboration with:
Floris Kummer, Lokesh Khandelwal , Stephen Justham, Luyau Lin, Edo Berger, Coen Neijssel, Alejandro Vigna-Gomez, Simon Stevenson, Tom Wagg, Lieke van Son, Ilya Mandel, Selma de Mink, Michelle Wassink

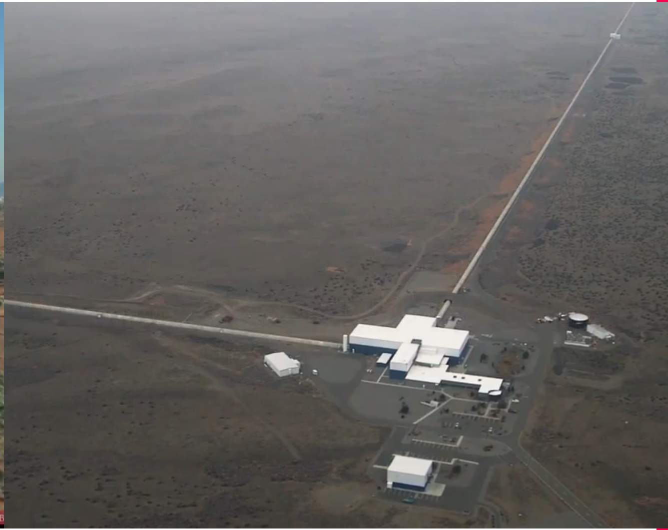
Topics in Astrostatistics meeting
28 January 2020

We are in the era of Big Data

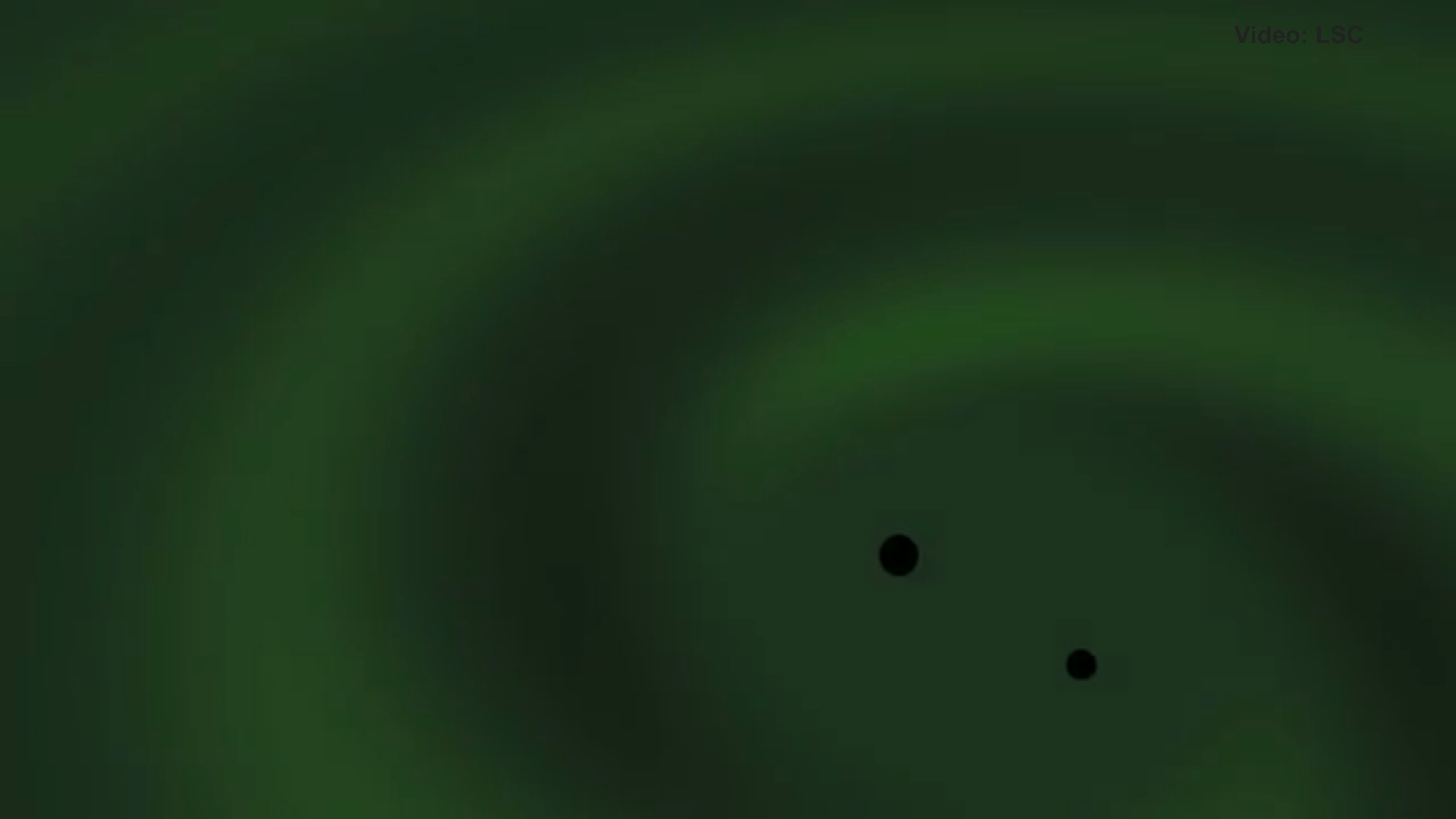


LSST/Vera Rubin Survey Telescope
(Optical)

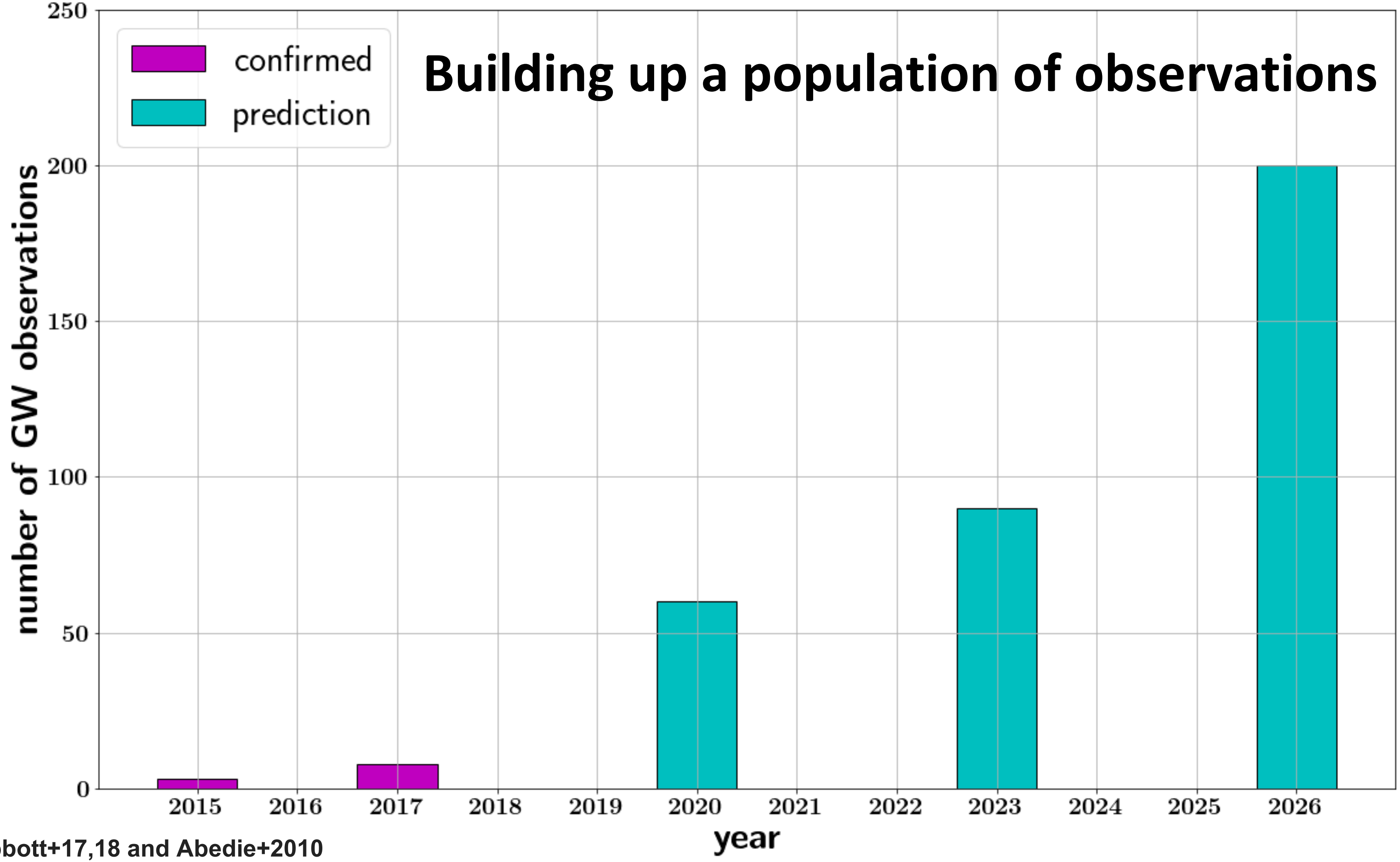
SKA
(Radio)



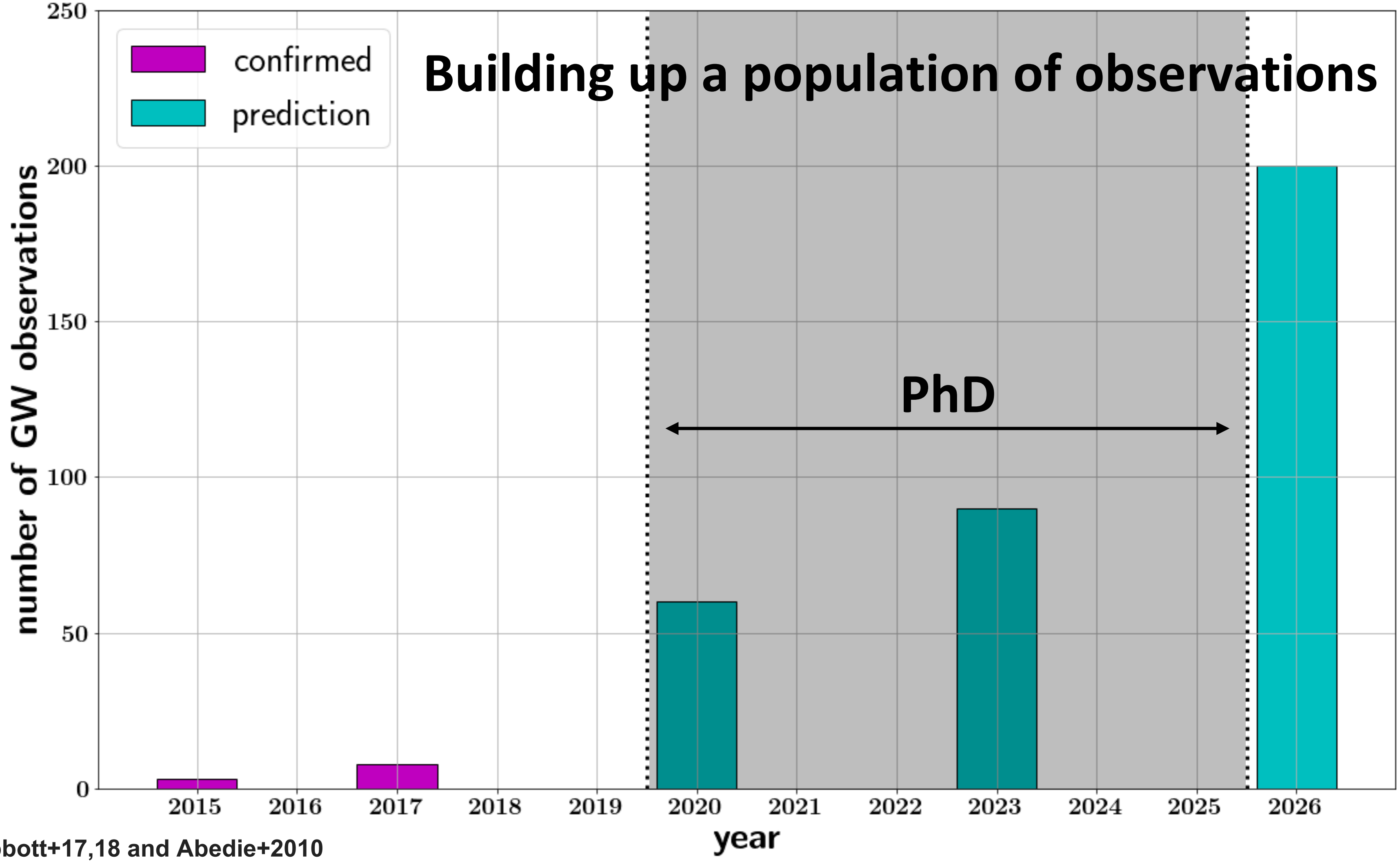
LIGO
(Gravitational Waves)


Video: LSC

Building up a population of observations



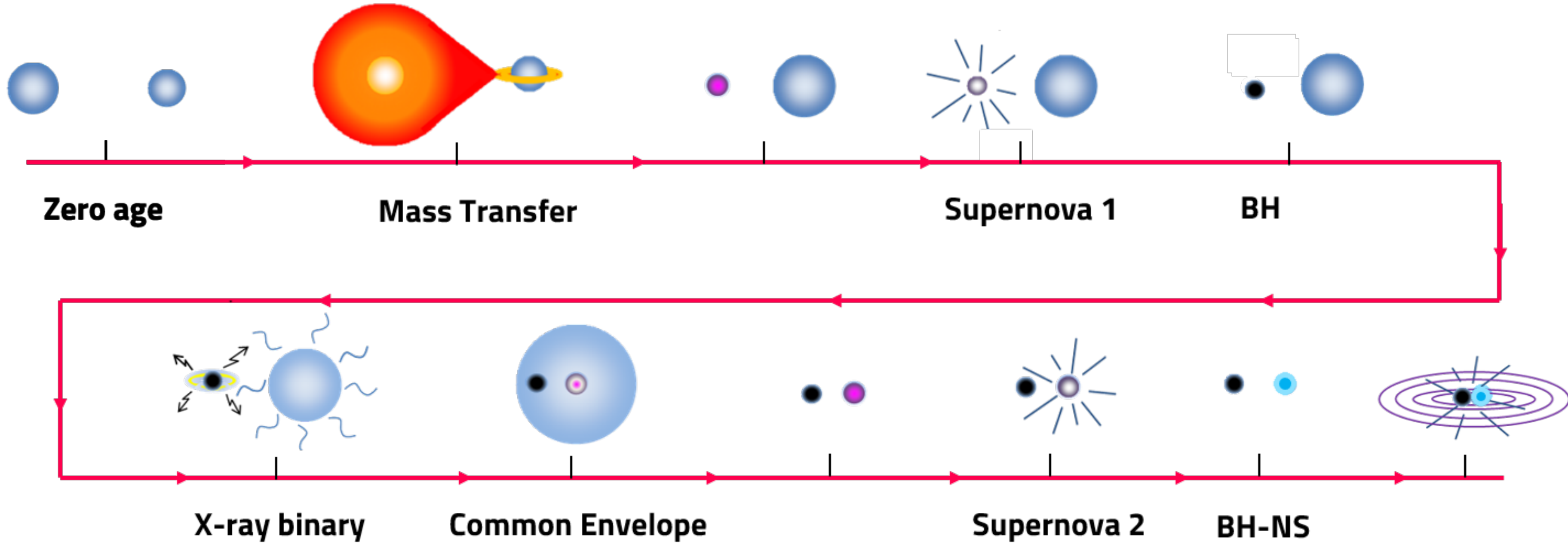
Based on Abbott+17,18 and Abedie+2010



Based on Abbott+17,18 and Abedie+2010

Classic channel BH-NS merger:

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



Searching for the progenitors of GWs

Floor Broekgaarden

Rapid binary population synthesis

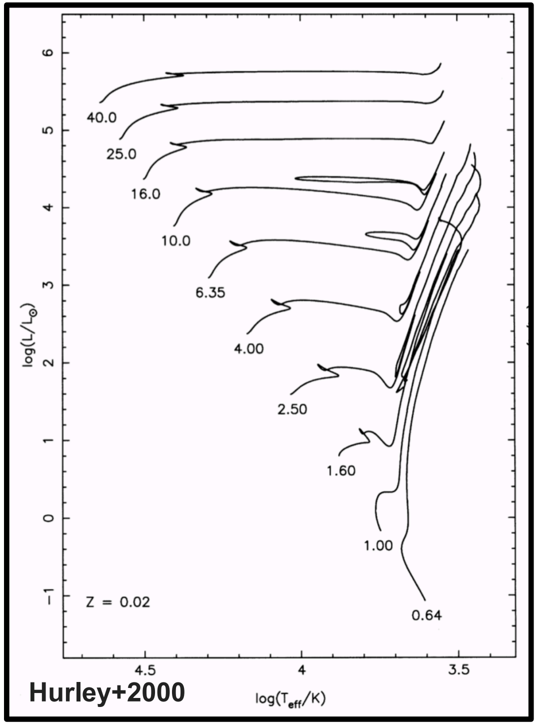
Stevenson+17, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

2 ways to implement (single) stellar evolution:

Slide from Stephen Justham

"Full" stellar calculations

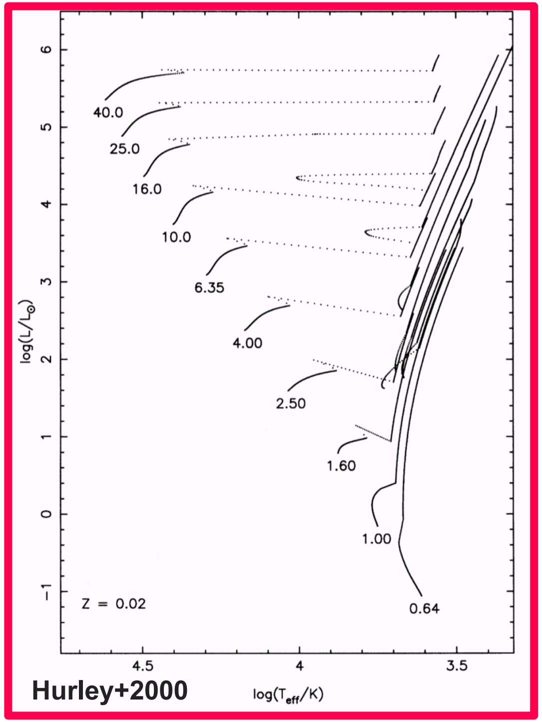
relatively slow



e.g. MESA

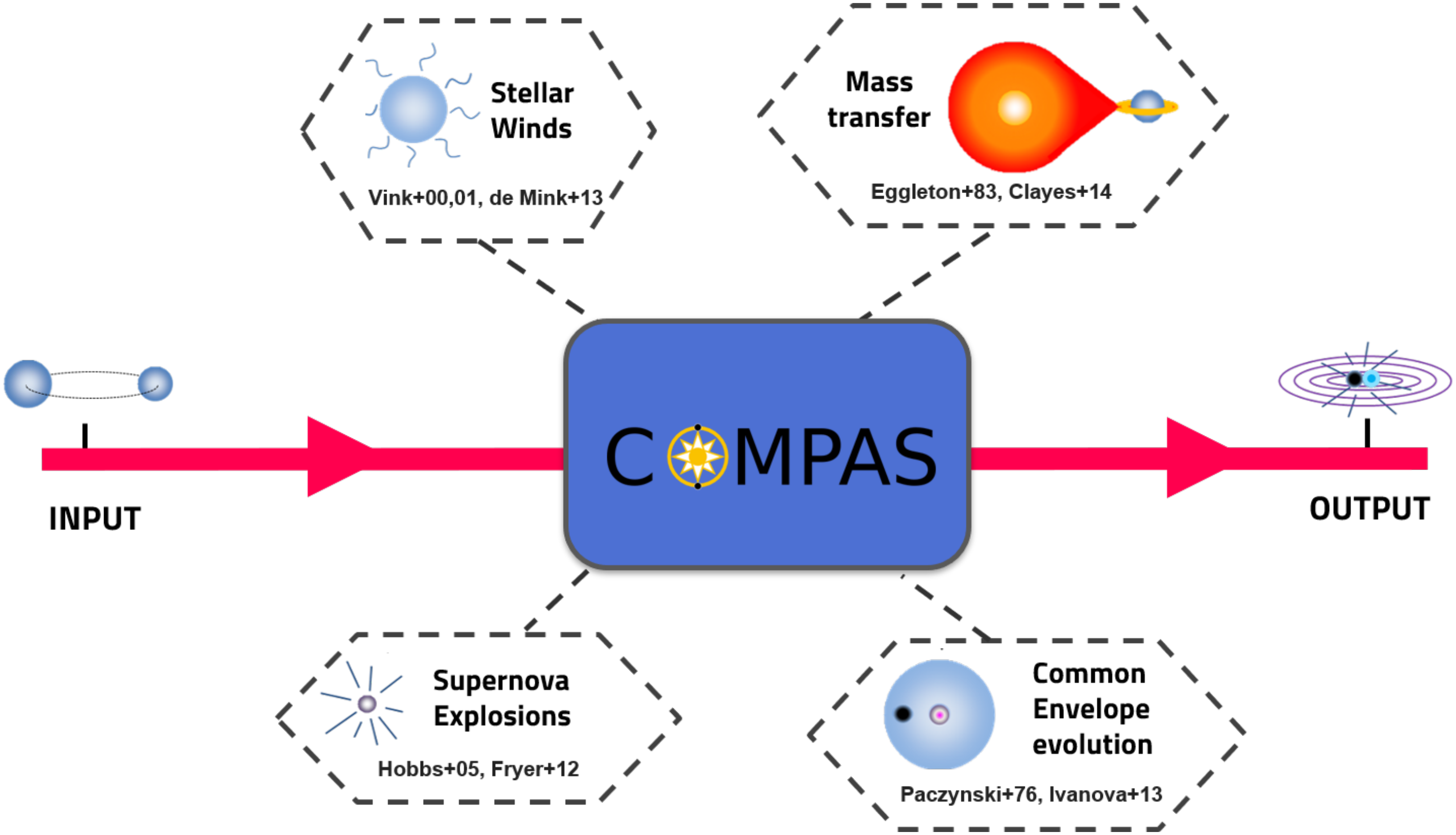
Analytical fits or interpolations

fast



e.g. StarTrack, binary_c, COMPAS

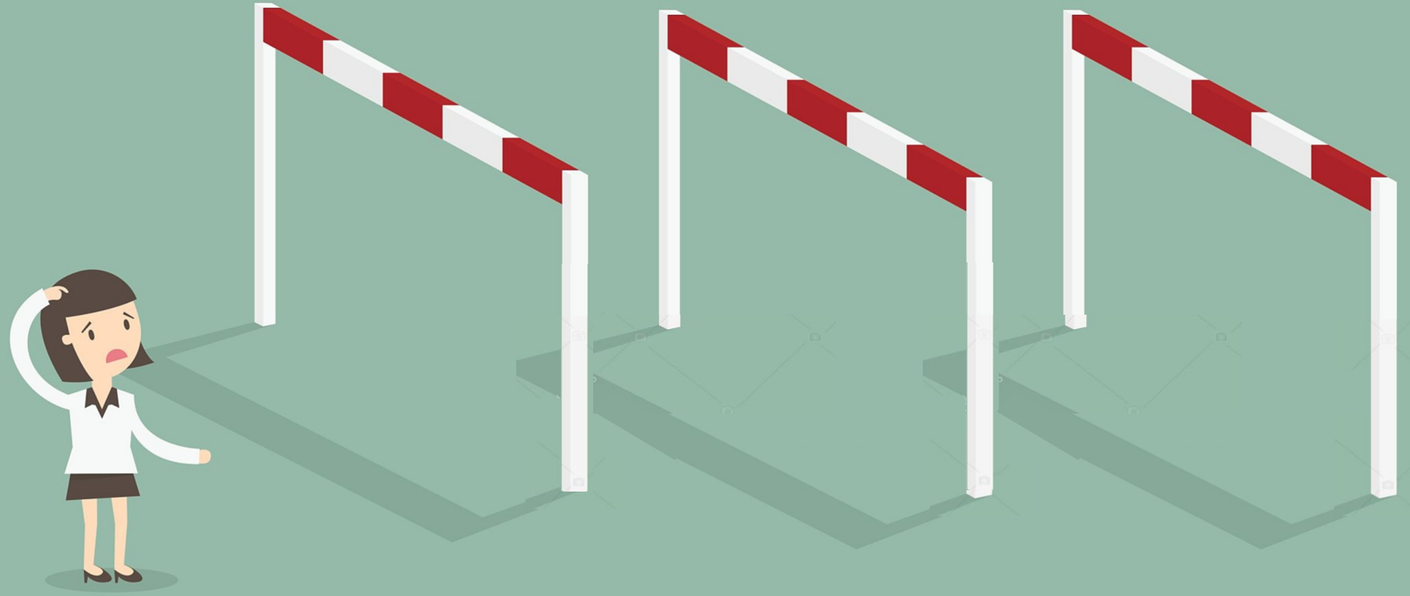
Rapid binary population synthesis

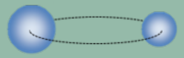


Stevenson+17, Barrett+18, Vigna-Gomez+18
Based on tracks from Hurley+00,02, Pols+98

3 computational challenges

1 Simulating populations





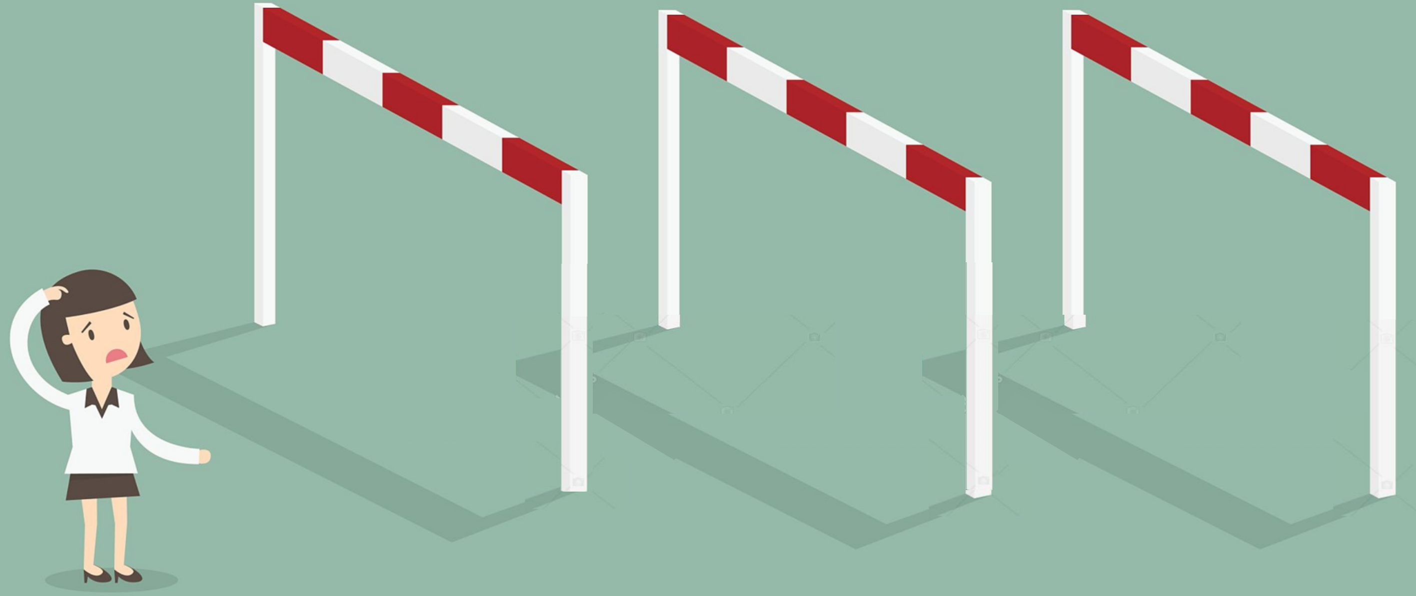
~1 sec

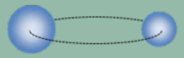
Cost:

3 computational challenges

1 Simulating populations

2 Testing physics





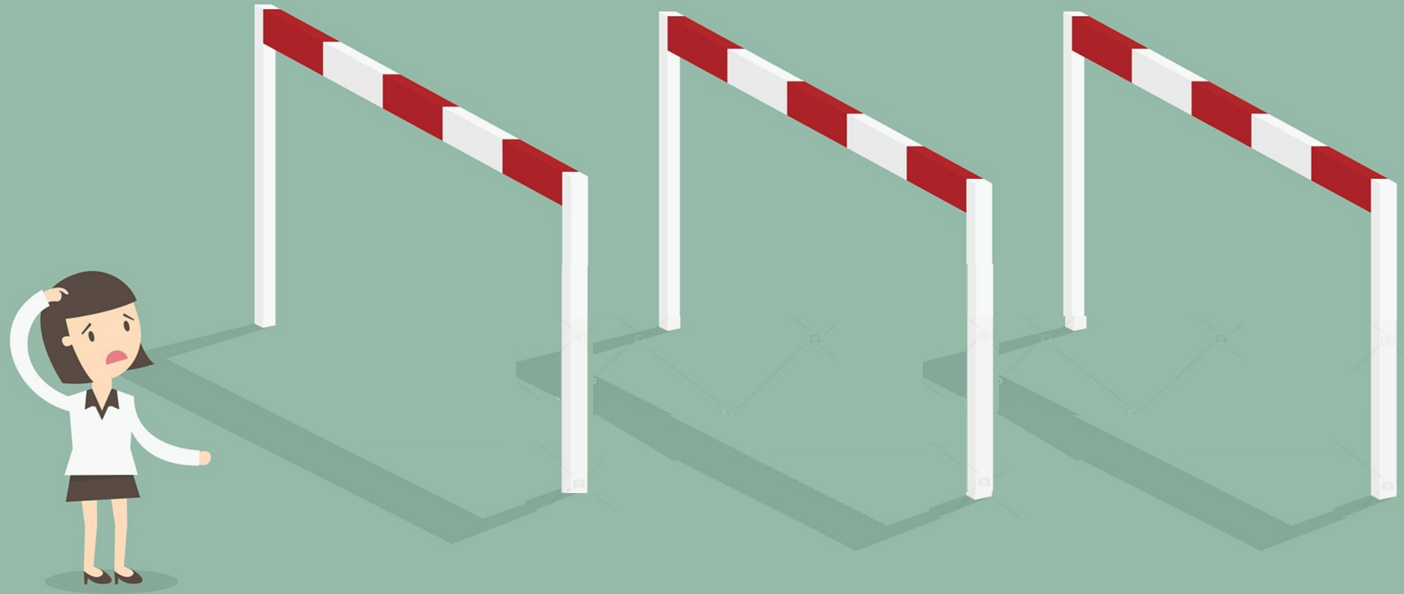
Cost: ~ 1 **sec** $\times 10^6$

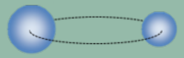
3 computational challenges

1 Simulating populations

2 Testing physics

3 Rare events

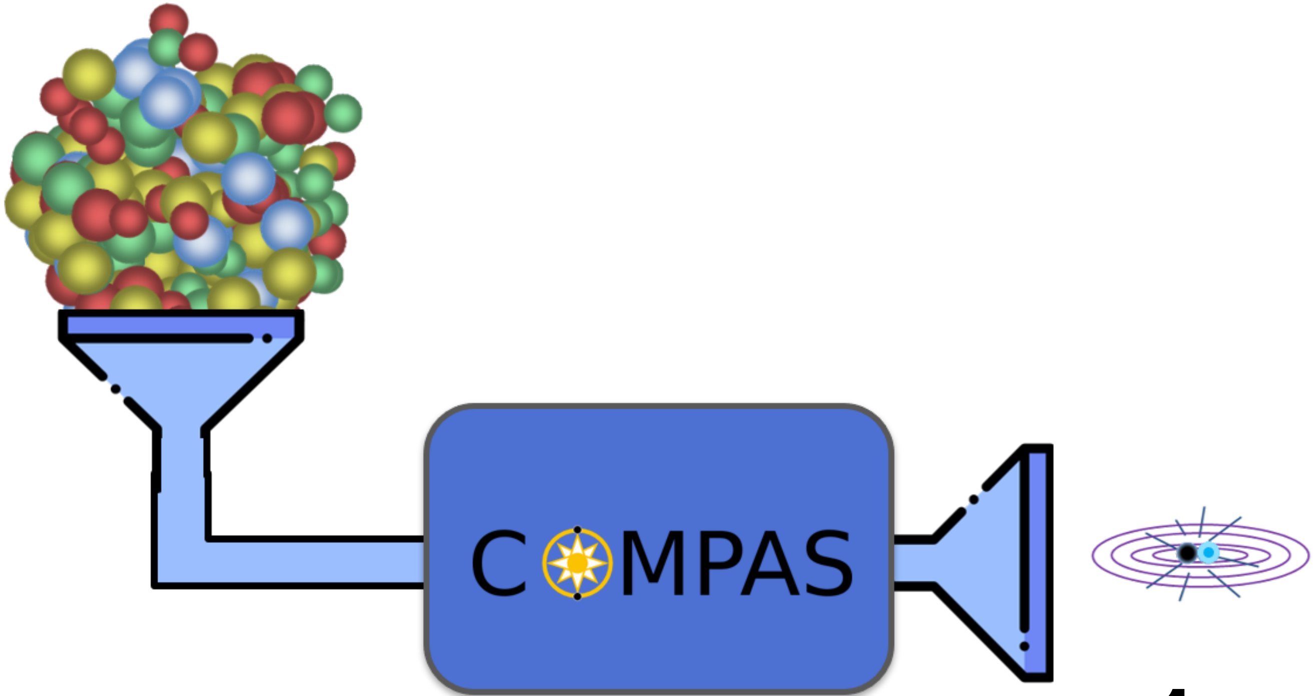




Cost: ~ 1 sec $\times 10^6$ $\times 10 - 10^6$

3 Rare events:

1000 binaries



only ~1 BH-NS merger!

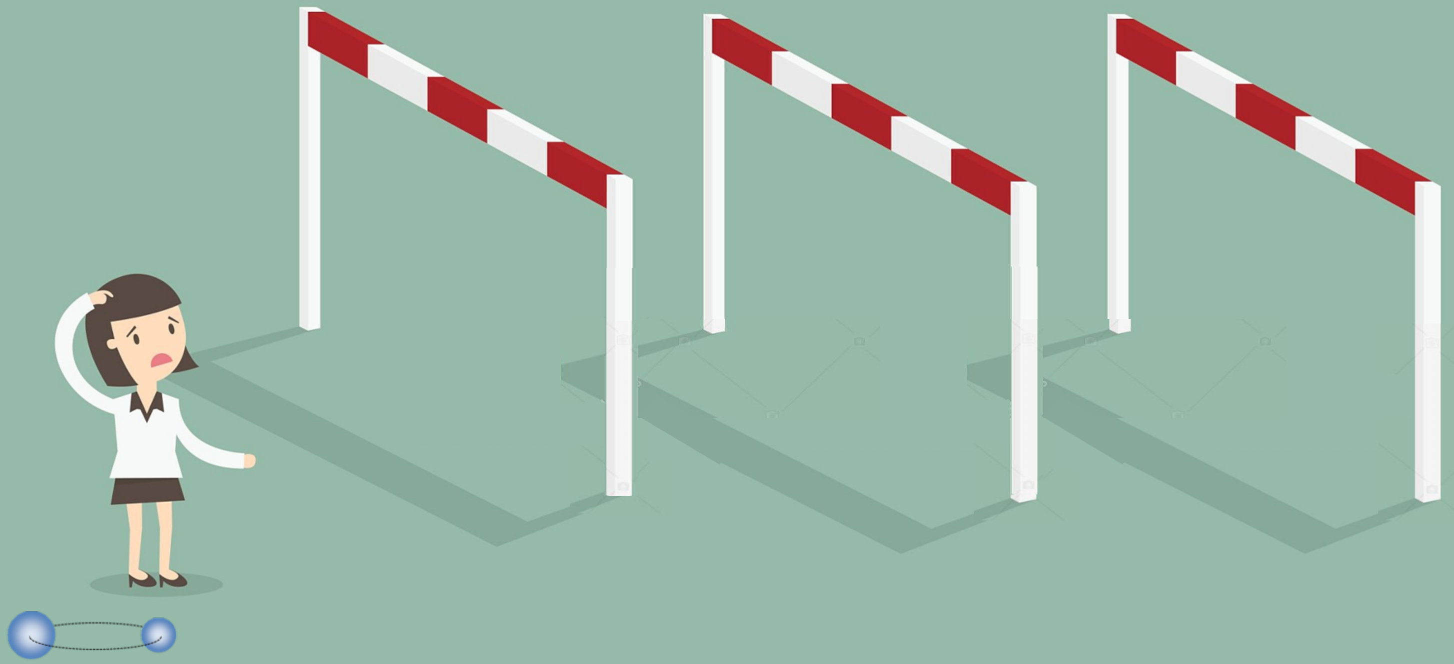
E.g. Belczynski+02

3 computational challenges

1 Simulating populations

2 Testing physics

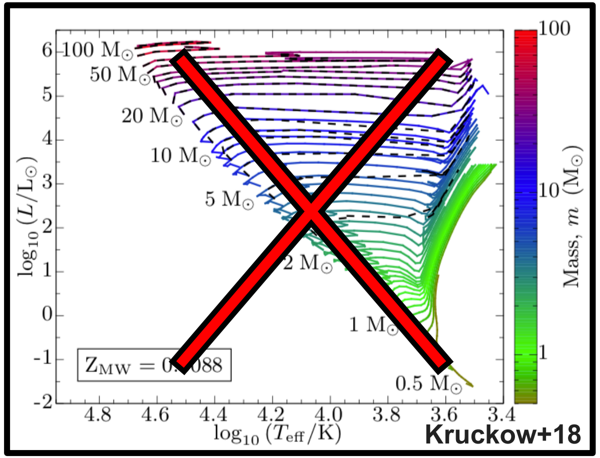
3 Rare events



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

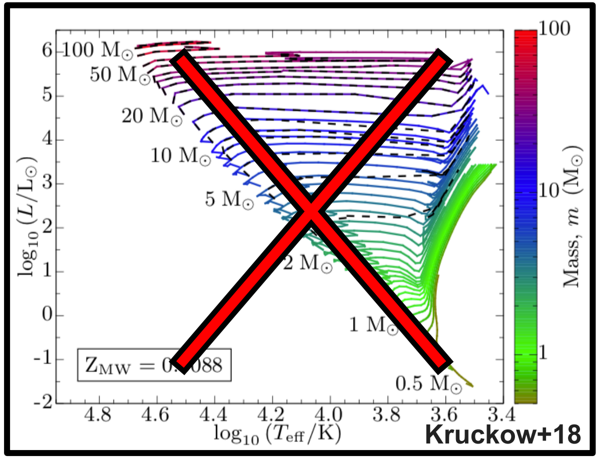
$\sim 3 - 3\times 10^4$ years computing time

Current binary population synthesis models pay a high price...

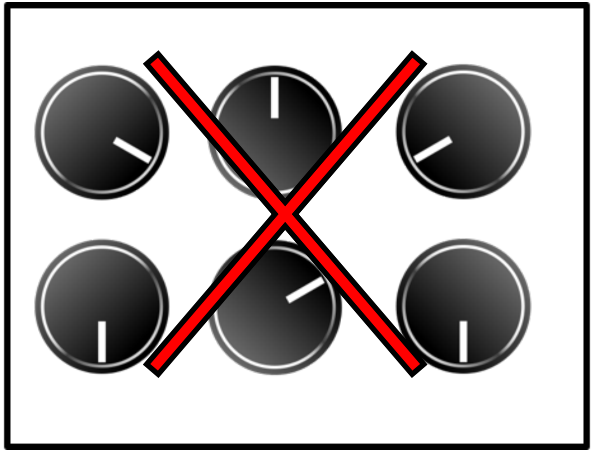


Do not include detailed prescriptions

Current binary population synthesis models pay a high price…

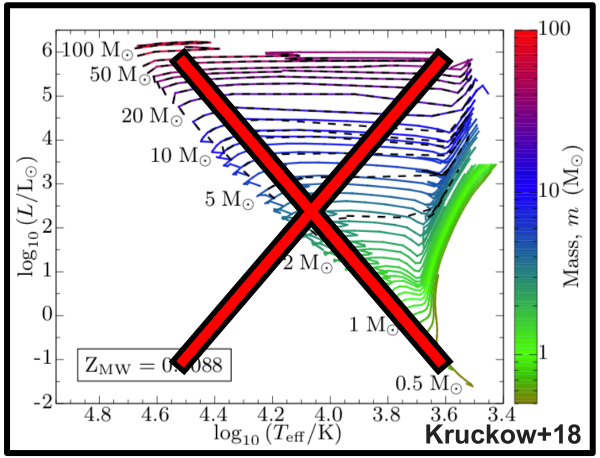


Do not include detailed prescriptions

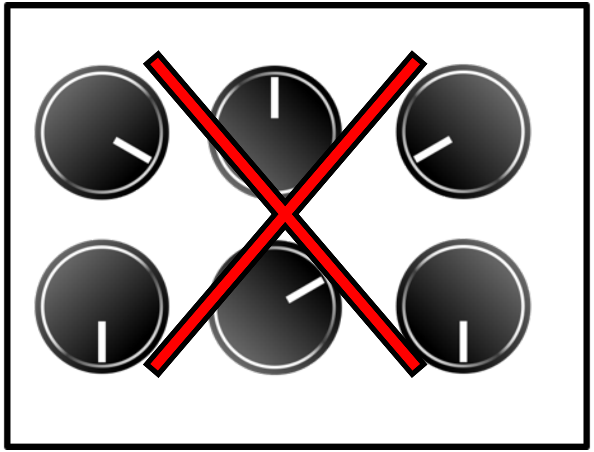


Perform only a small parameter study

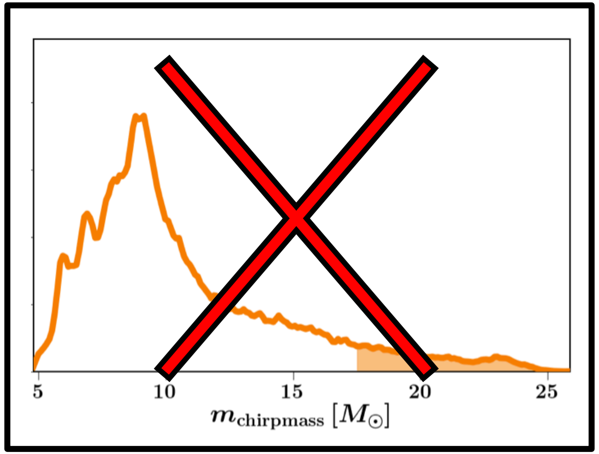
Current binary population synthesis models pay a high price...



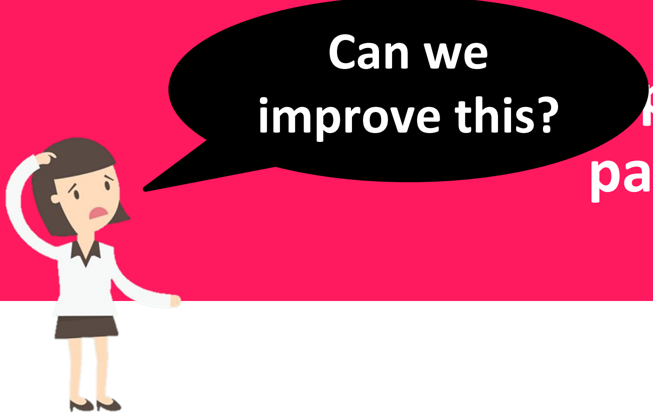
Do not include detailed prescriptions



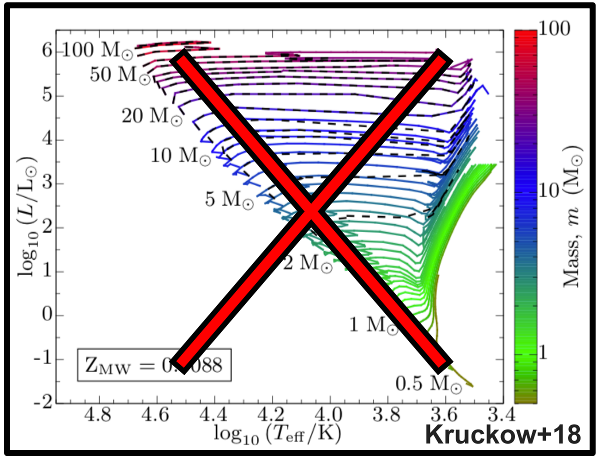
Perform only a small parameter study



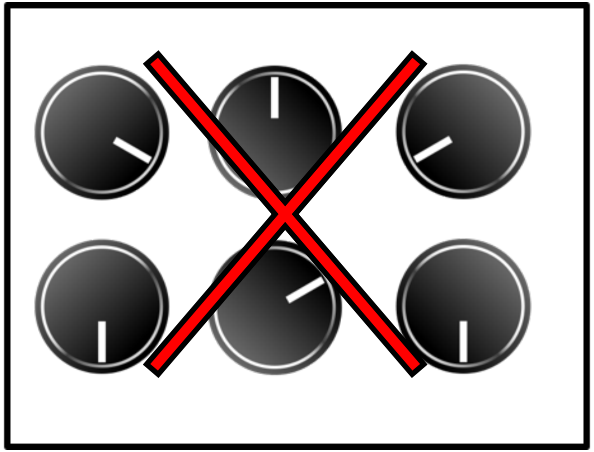
Do not explore tails of distributions



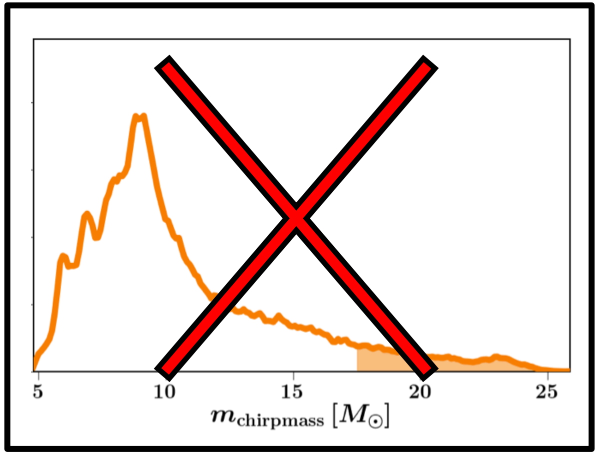
population synthesis models pay a high price…



Do not include detailed prescriptions



Perform only a small parameter study

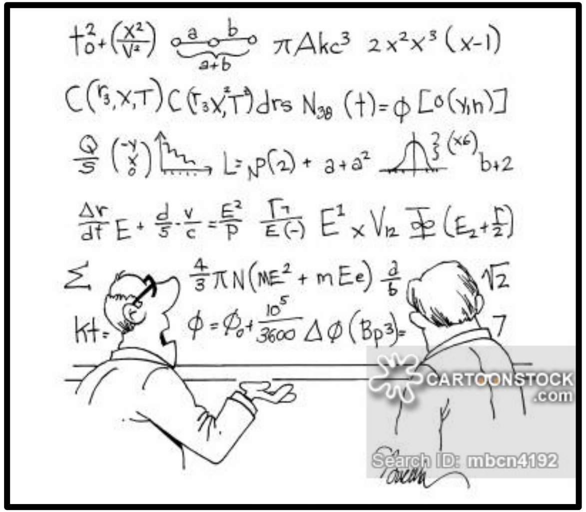


Do not explore tails of distributions

Previous work

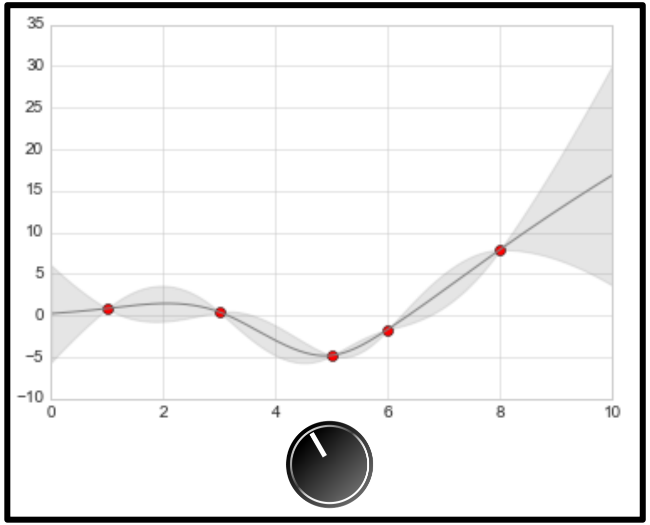
Analytical formalisms

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



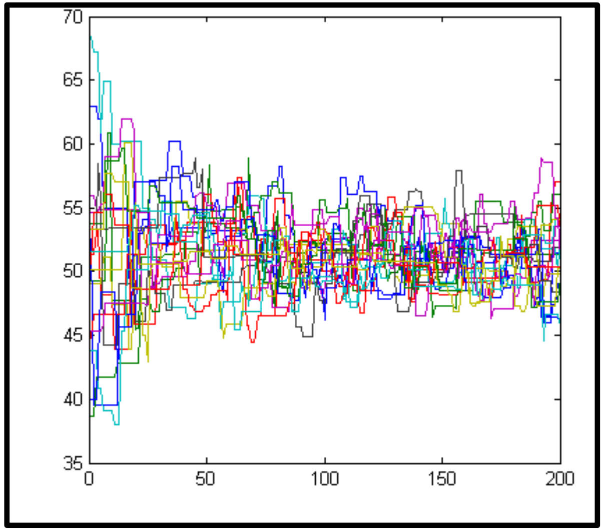
Emulators (GPR)

Barret+2017, Taylor & Gerosa 2018



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

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)

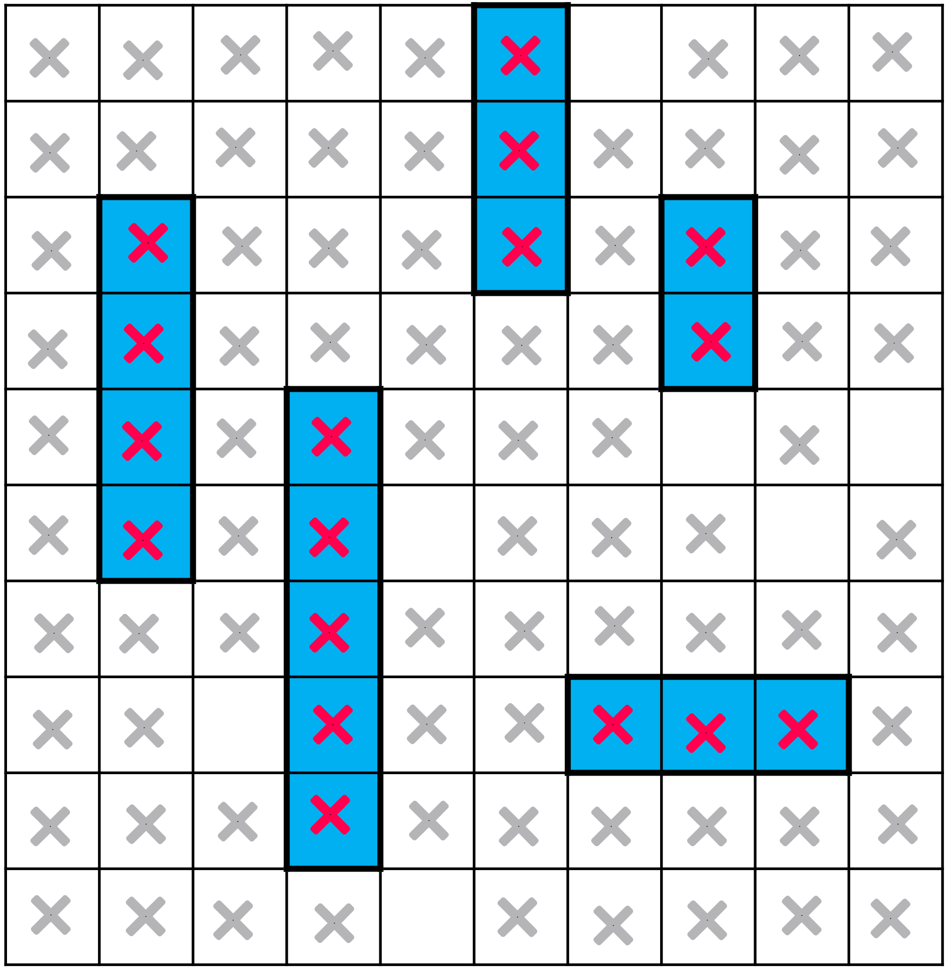
Traditional models use *"random shooting"*:

requires

$$< 96 >$$

shots

to complete the game



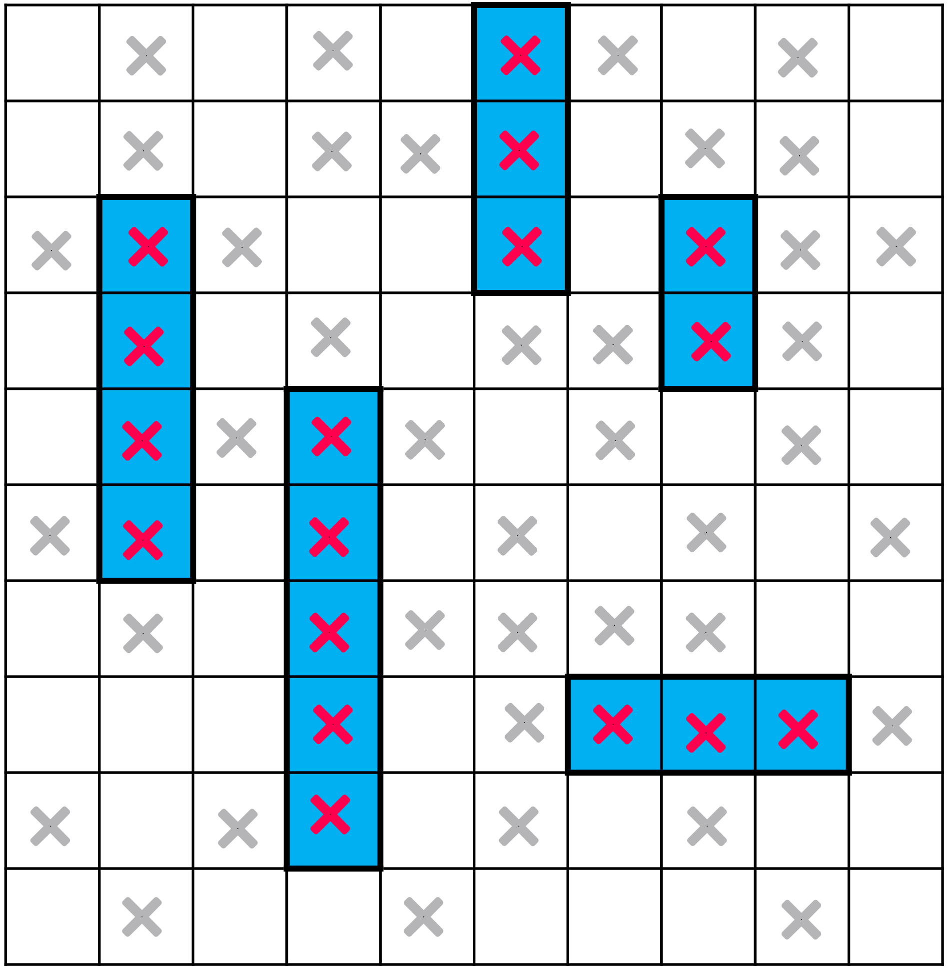
STROOPWAFEL
uses *"explore/refine"*:

requires

< 65 >

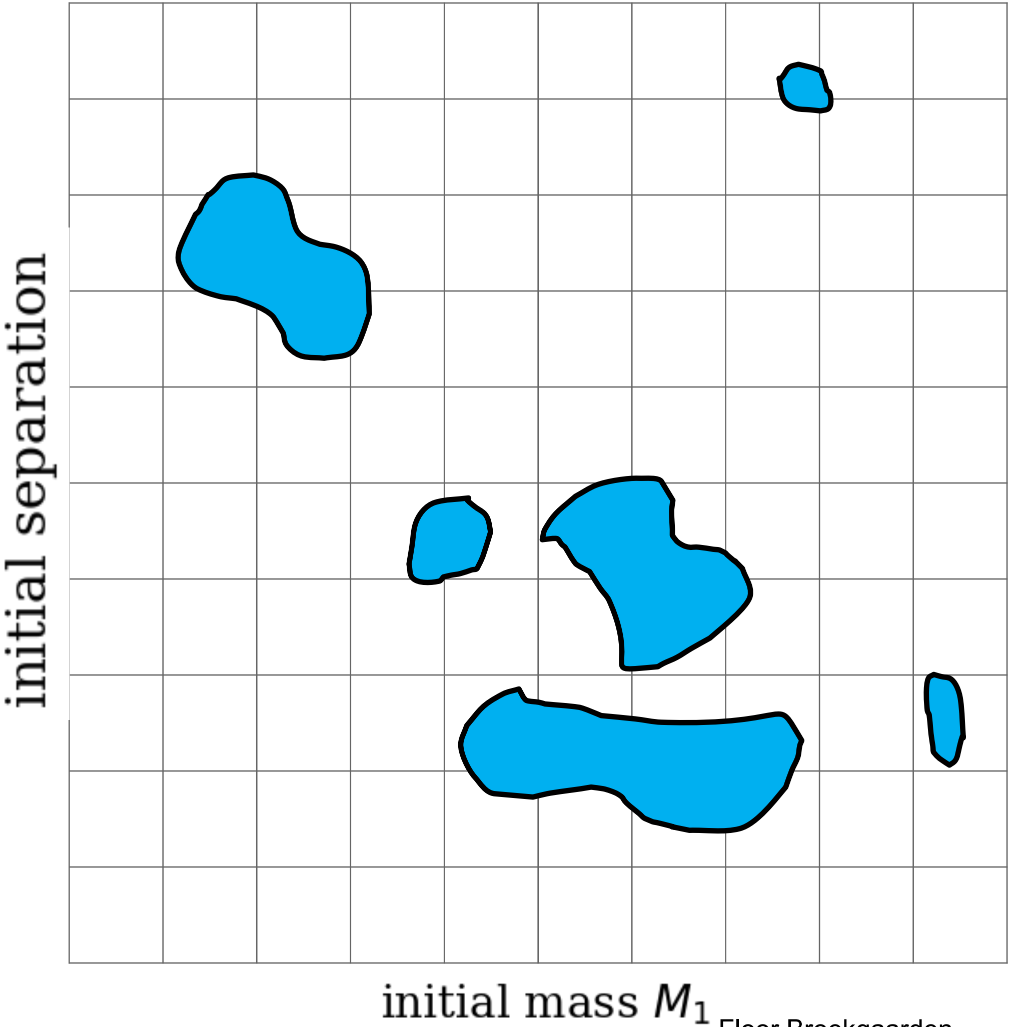
shots

to complete the game



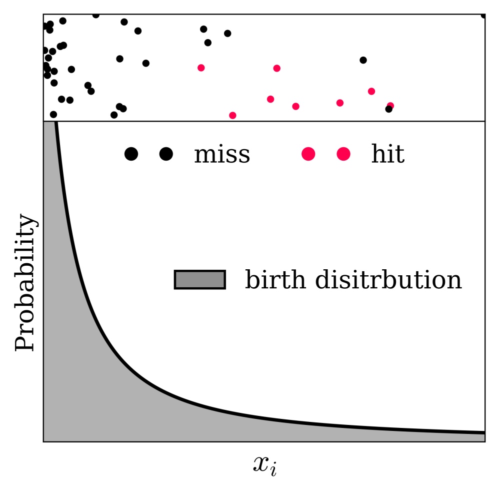
Binary population synthesis:

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



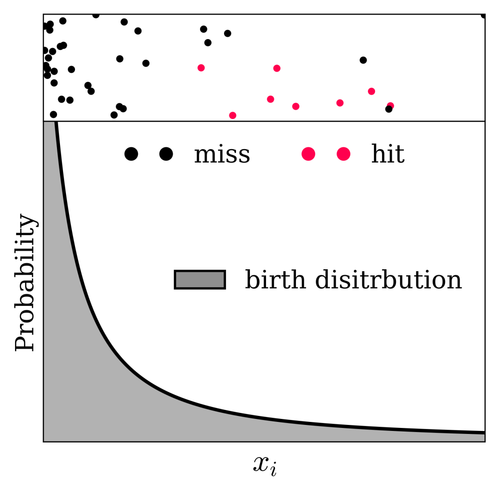
STROOPWAFEL:

1) Exploring phase

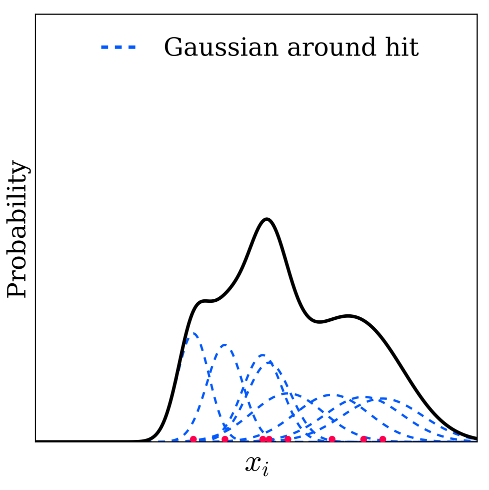


STROOPWAFEL:

1) Exploring phase

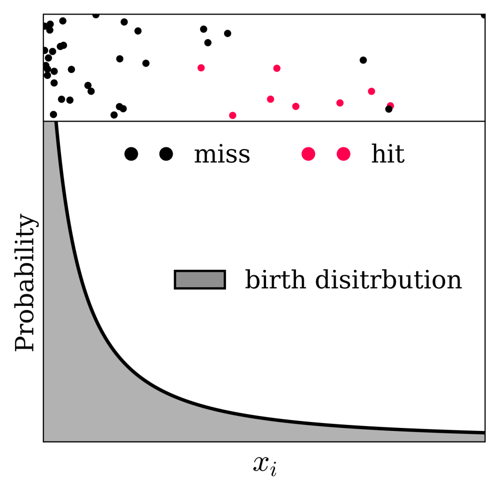


2) Create adapted distribution

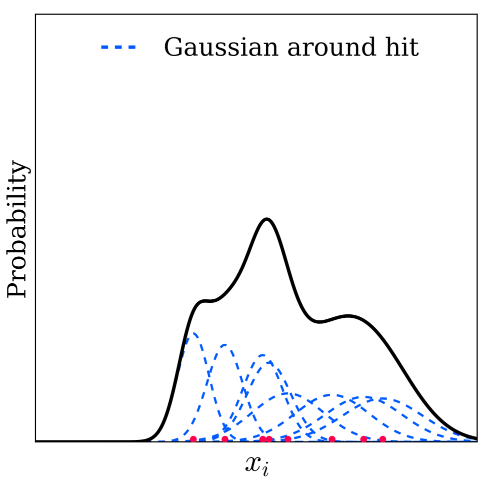


STROOPWAFEL:

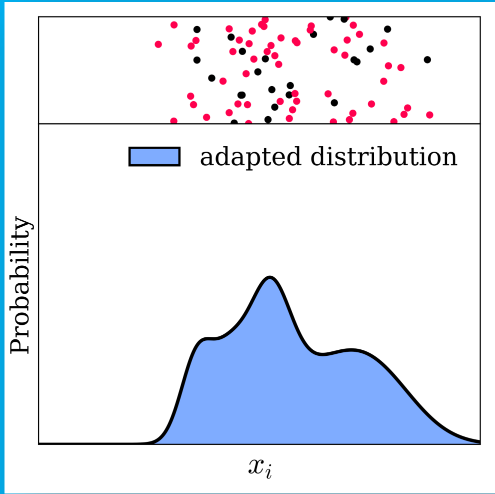
1) Exploring phase



2) Create adapted distribution



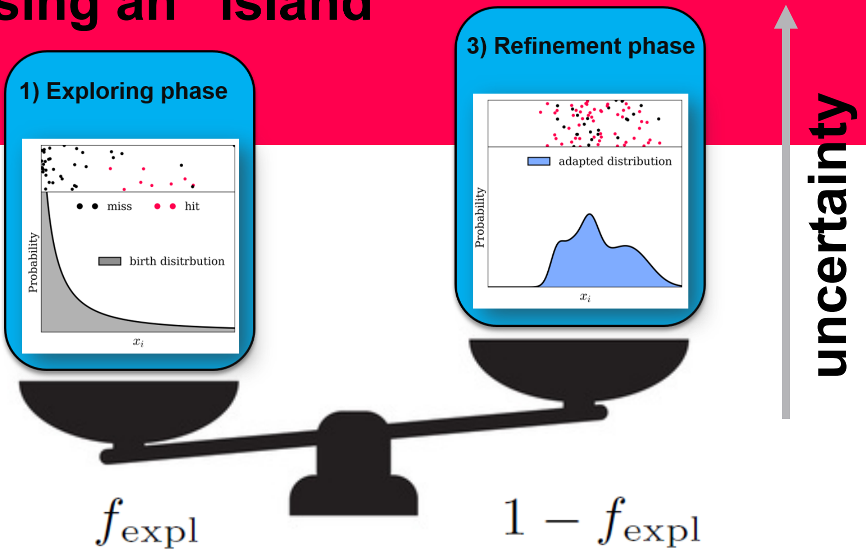
3) Refinement phase

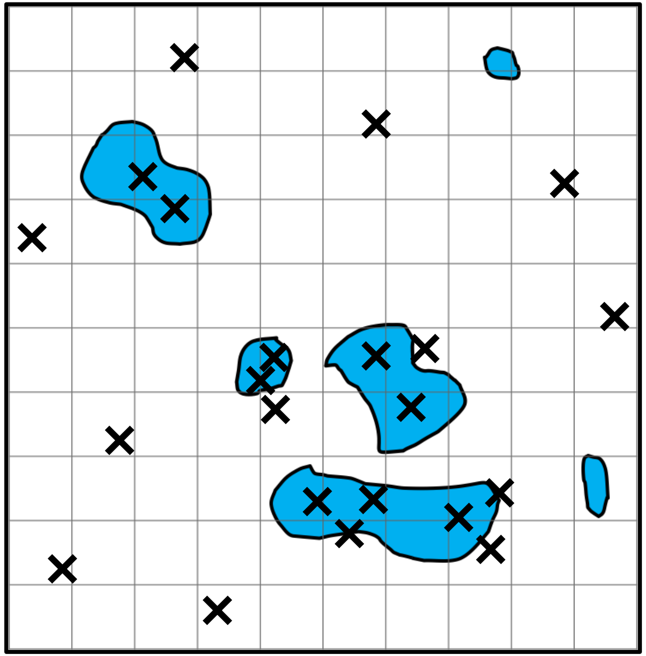


When to switch from exploring to refinement?

Getting more "hits"

Missing an "island"



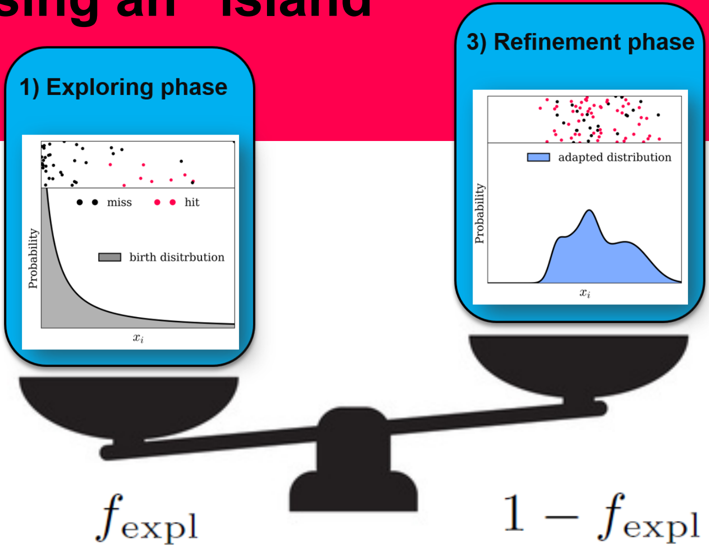


When to switch from exploring to refinement?

Getting more “hits”

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

Missing an “island”



Uncertainty from refining

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

uncertainty from missing an island

Results

Ingredients

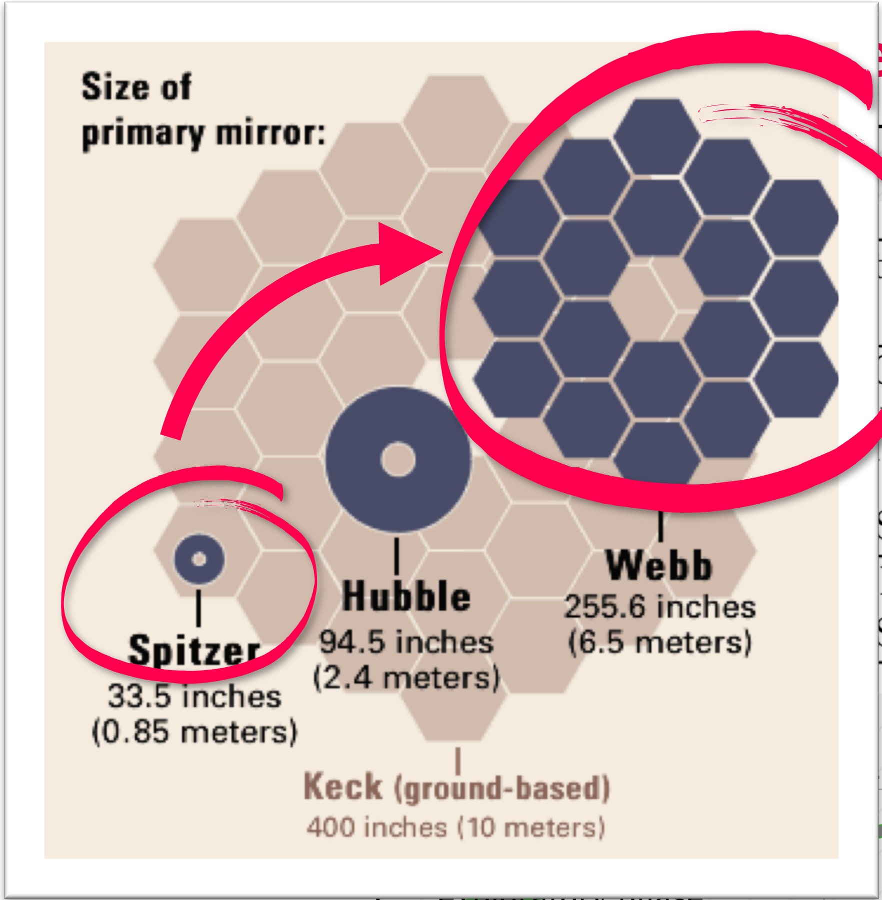
- 6 different target populations

- 3 parameters:
  m1, m2 & separation

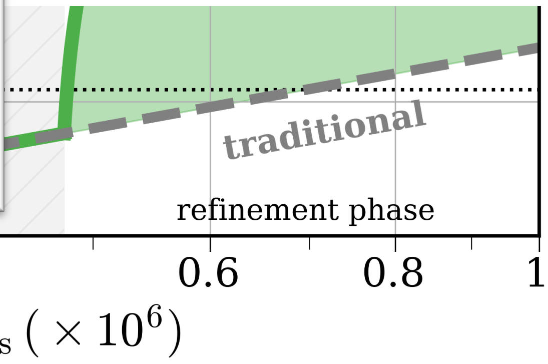
- traditional vs STROOPWAFEL sampling

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 $m_{tot} > 50\,M_{\odot}$
6	NS–NS mergers with $t_{coal} < 50$ Myrs

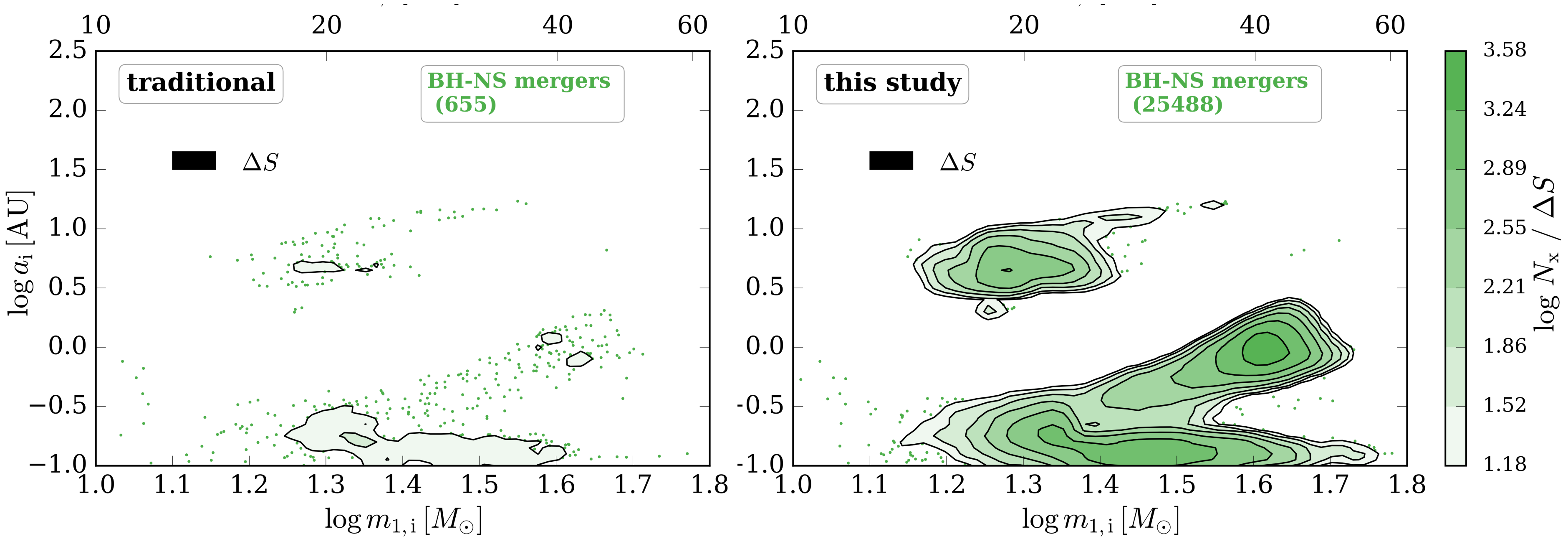
efficiency



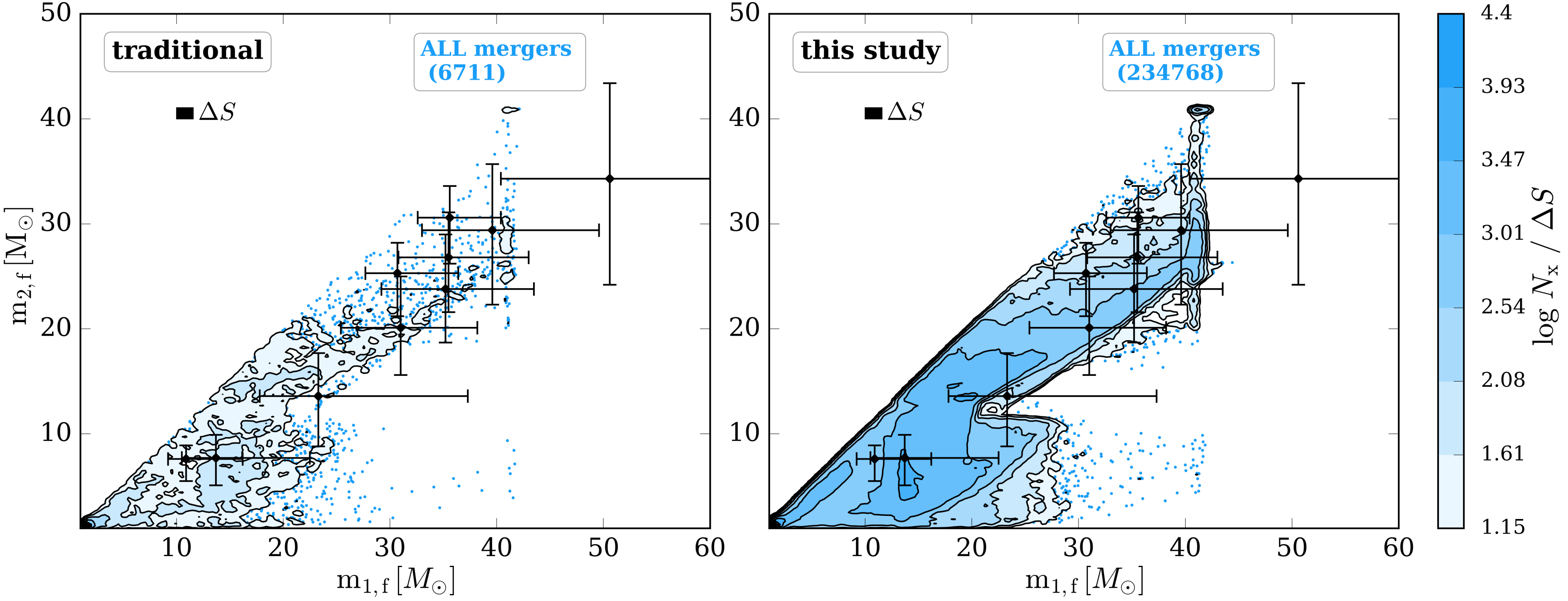
subpopulation	gain
CO mergers in a Hubble time	35×
H mergers in a Hubble time	53×
S mergers in a Hubble time	39×
S mergers in a Hubble time	45×
H mergers $m_{tot} > 50\,M_{\odot}$	202×
S mergers with $t_{coal} < 50$ Myrs	24×



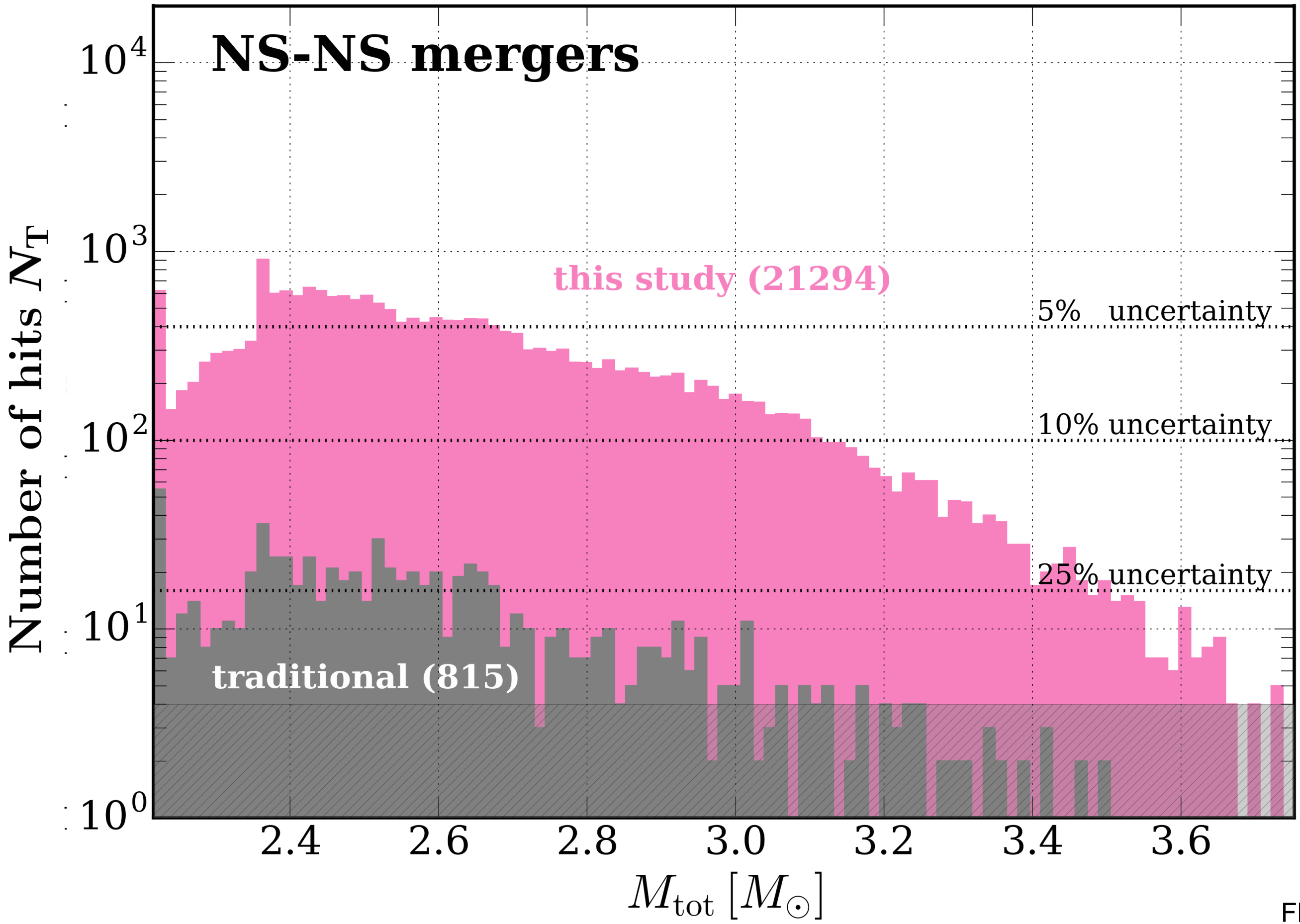
Higher resolution on input parameters



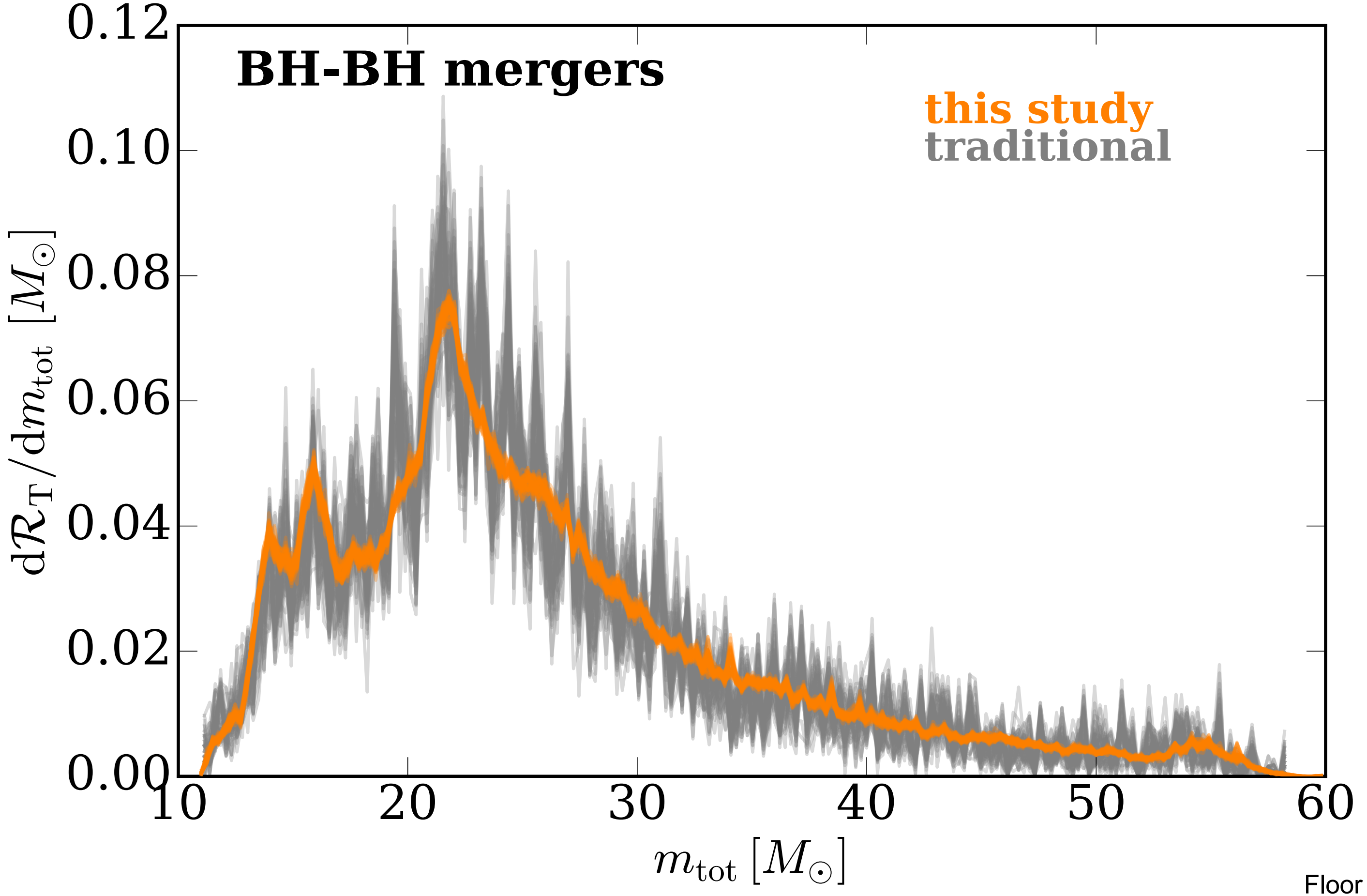
Higher resolution output



Resolve tails of distributions



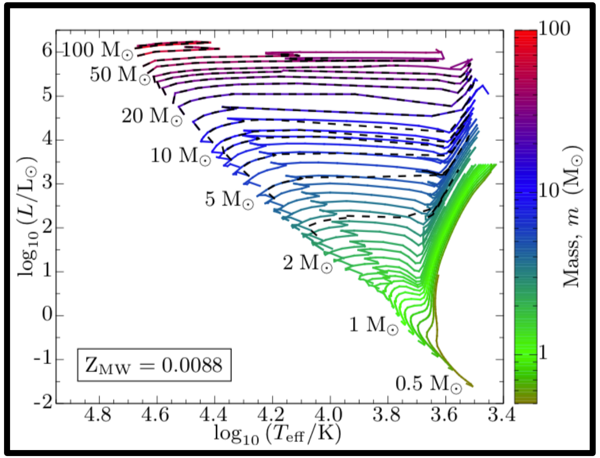
Better distribution functions



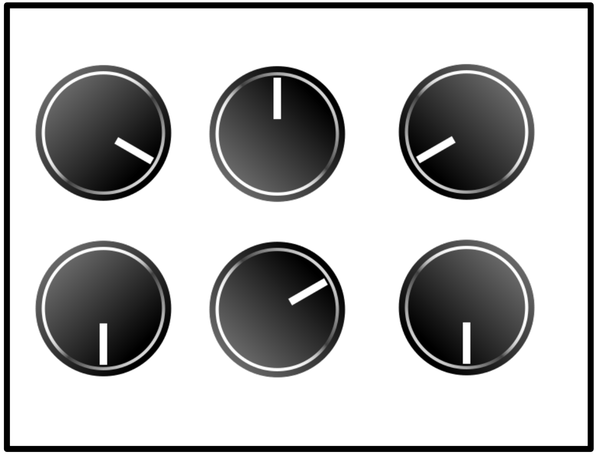
Conclusions

- **STROOPWAFEL obtains $\times 30 - 200$ more hits or: provides $\times 30 - 200$ speed up**

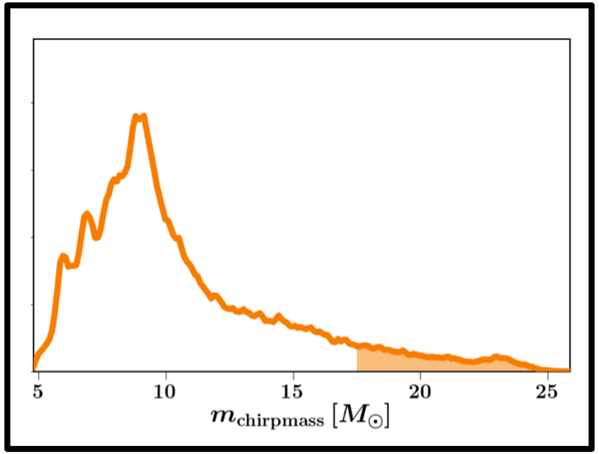
STROOPWAFEL helps next generation simulations:



Include detailed prescriptions



Perform larger parameter study



Explore tails of distributions

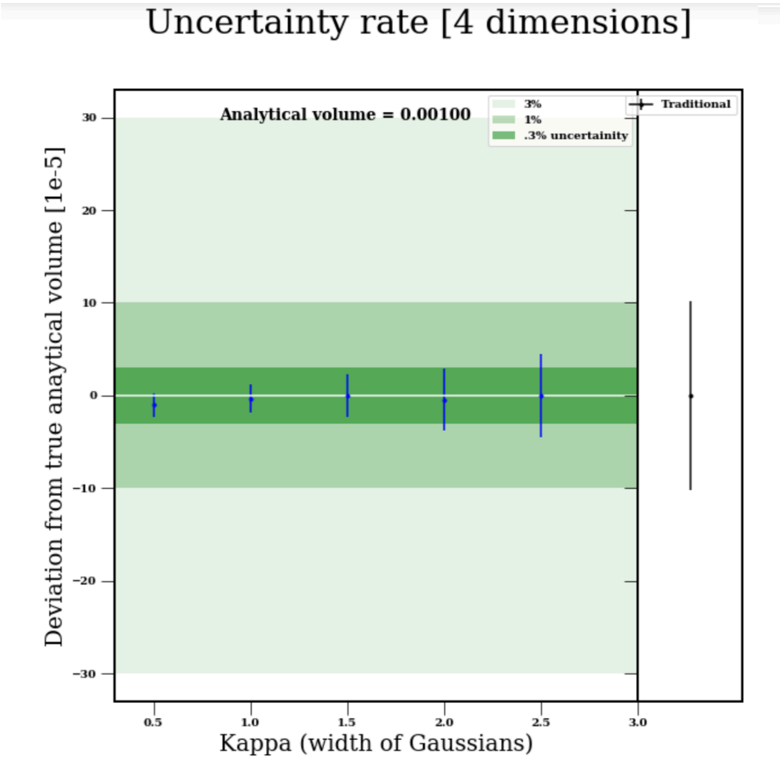
Floor Broekgaarden

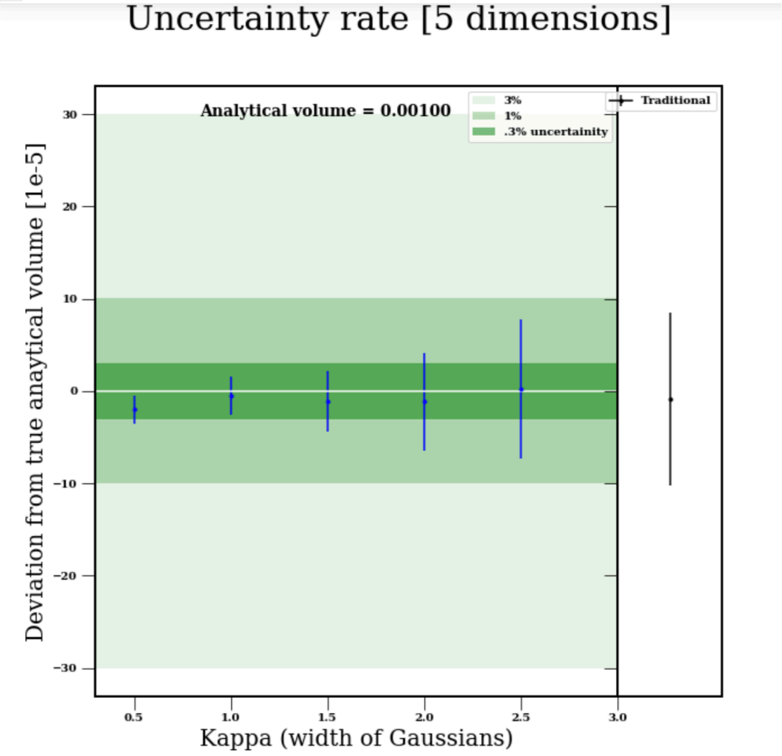
Conclusions

- **STROOPWAFEL obtains $\times 30 - 200$ more hits or: provides $\times 30 - 200$ speed up**
- **Improves population studies of rare events**

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

Future Prospects: more dimensions





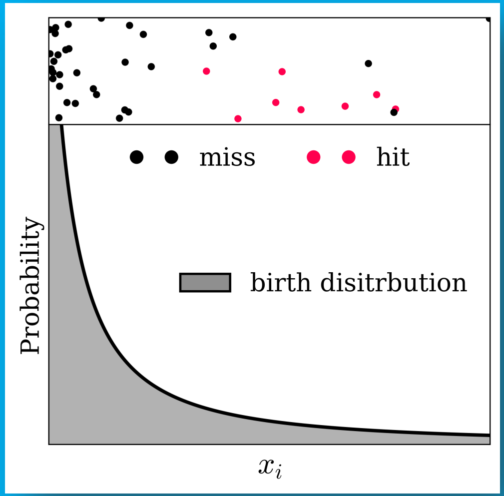
Floris Kummer

Lokesh Khandelwal

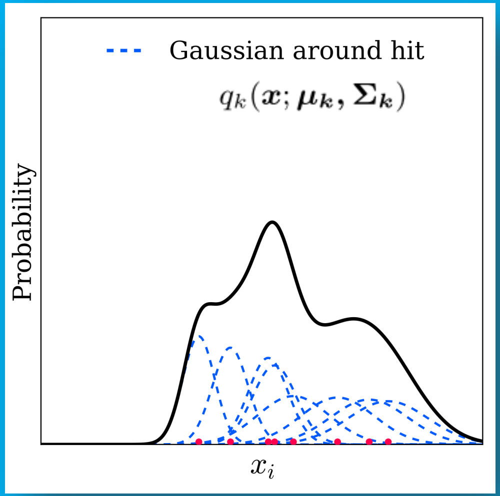
Future prospects: non-diagonal Gaussians?

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

1) Exploring phase



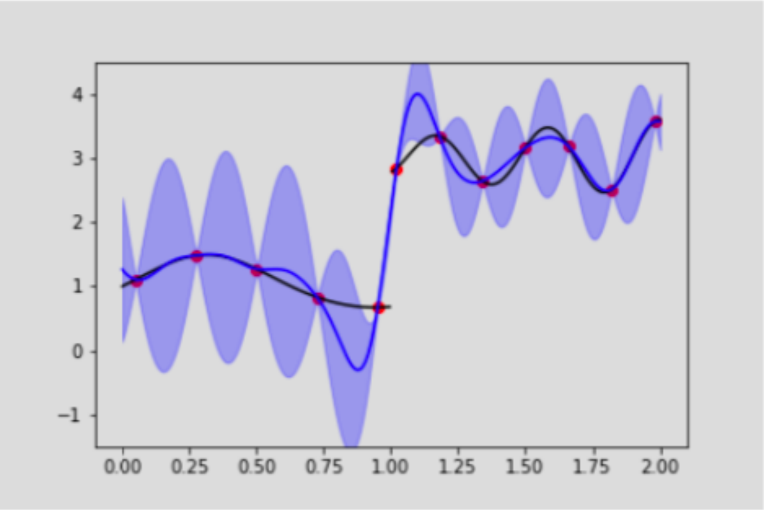
2) Create adapted distribution



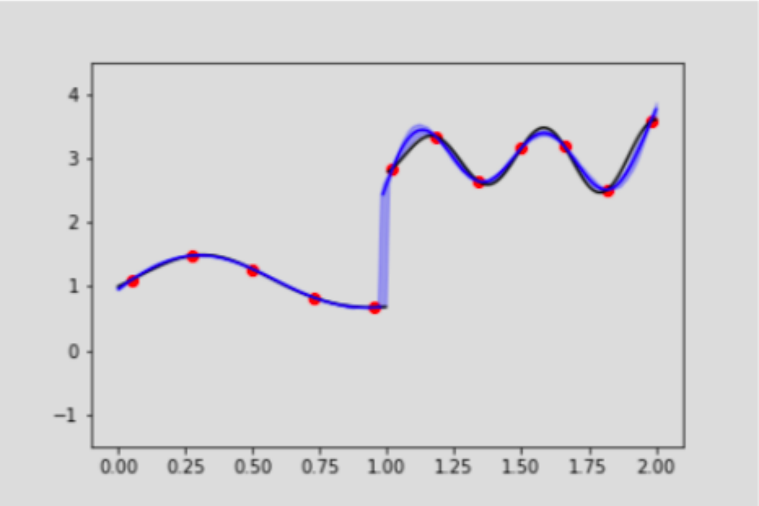
$$\sigma_{j,k} = \kappa \frac{1}{\pi_j(x_k) N_{\text{expl}}^{1/d}},$$

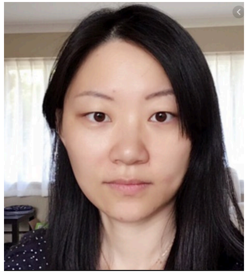
Future Prospects: Gaussian process classifiers & emulation

Traditional Gaussian Process Regression



Luyao Lin (in prep.)



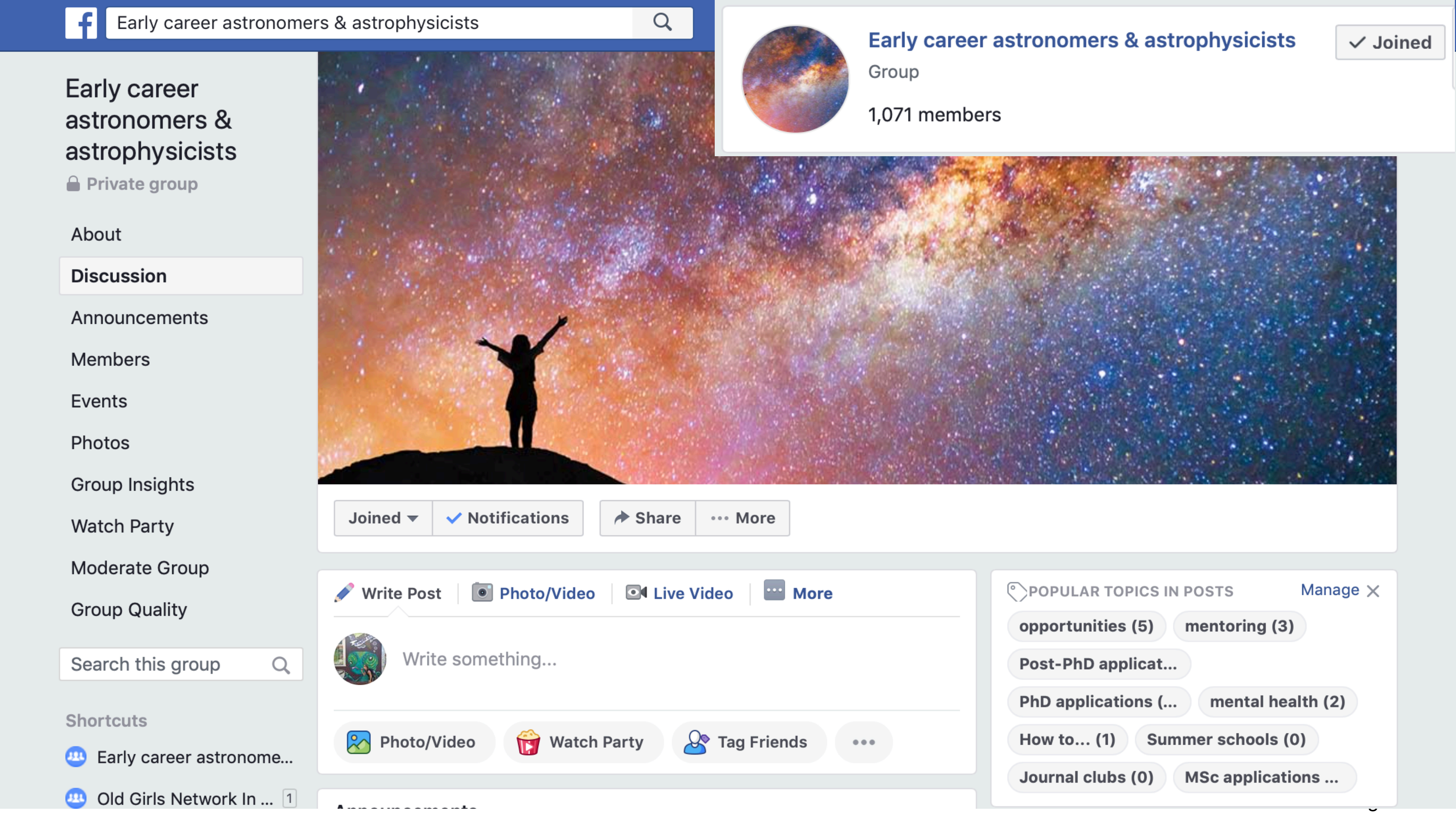


Luyao Lin

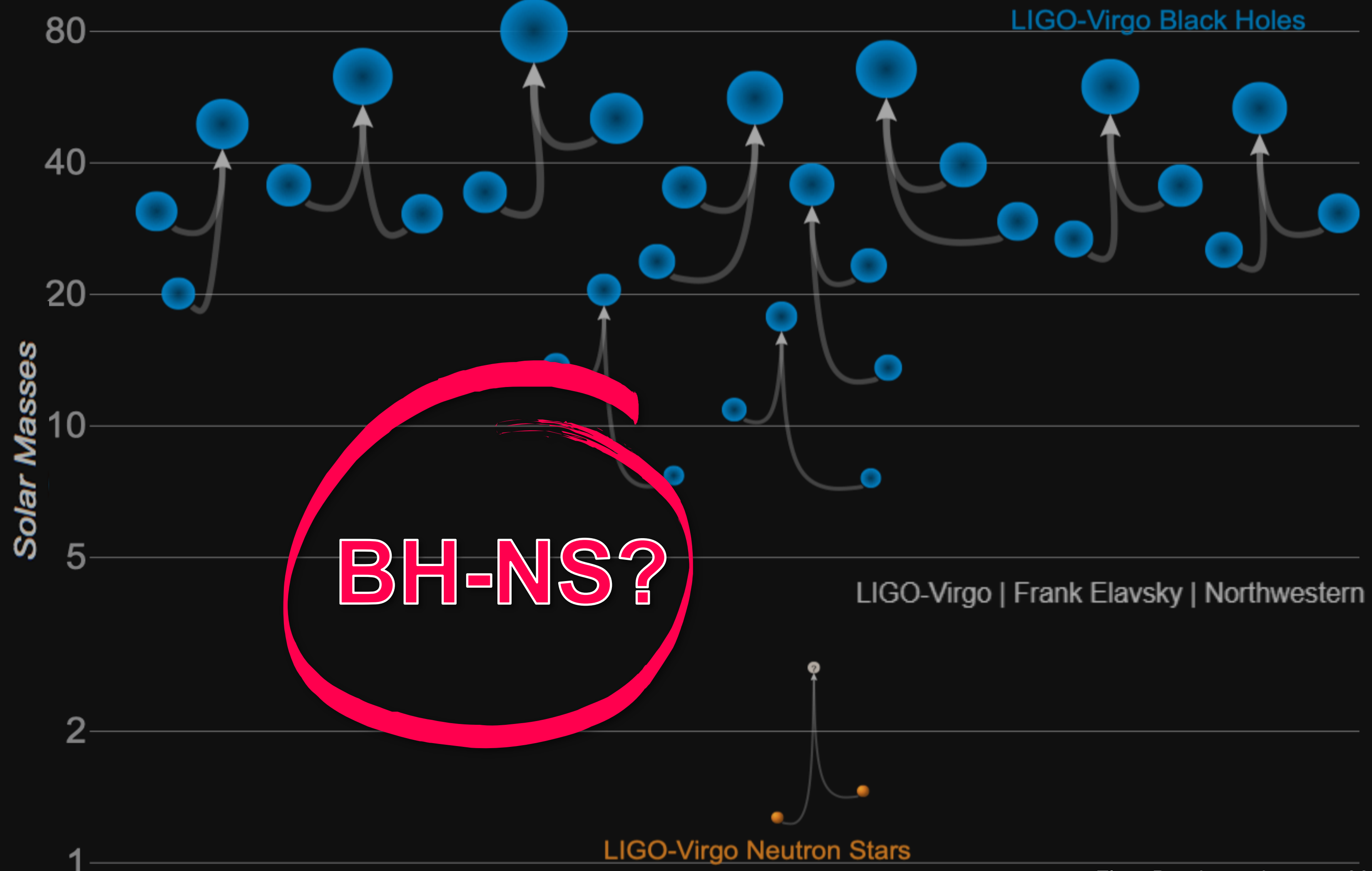
Derek Bingham

Thank you!


TIME COST
STRATEGY A
STRATEGY B
ANALYZING WHETHER STRATEGY A OR B IS MORE EFFICIENT
THE REASON I AM SO INEFFICIENT

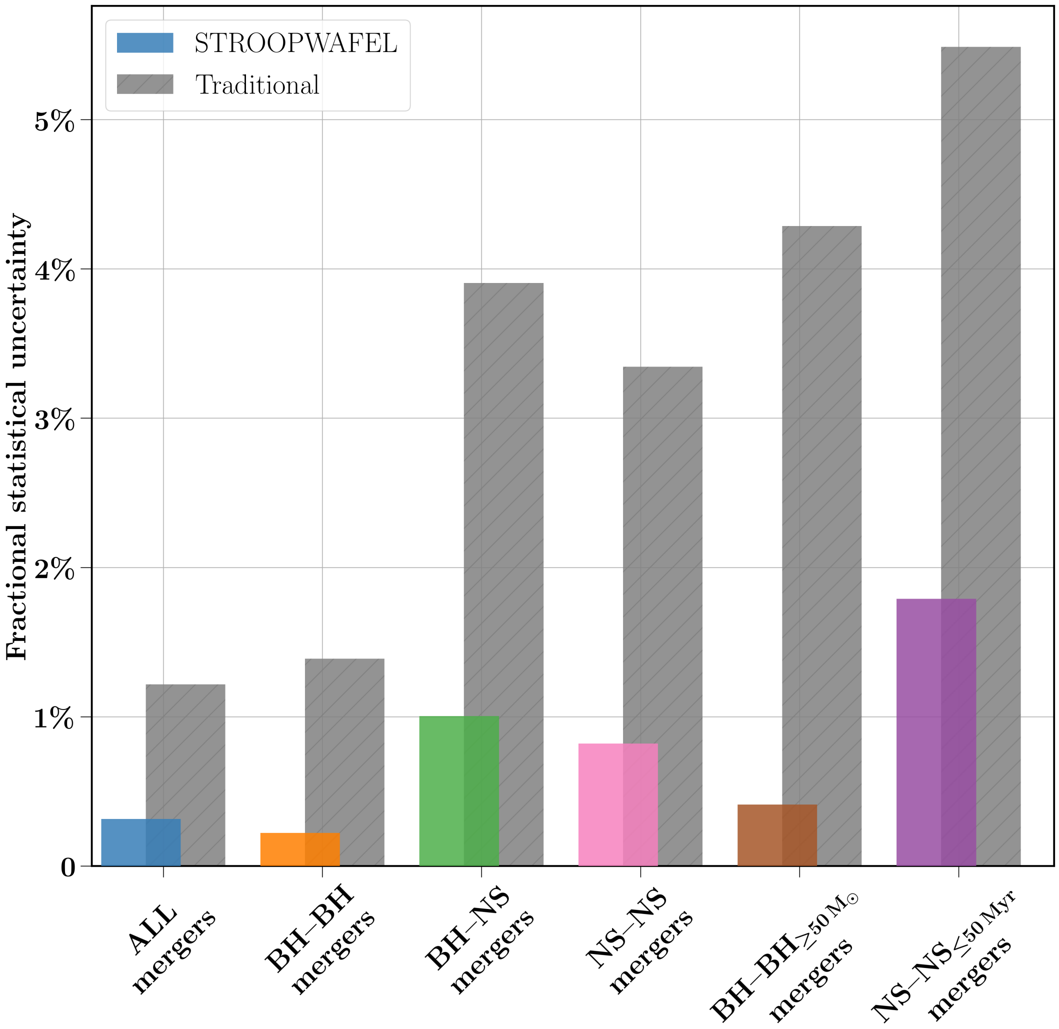

Early career astronomers & astrophysicists
Early career astronomers & astrophysicists
Group
1,071 members
Joined
Early career astronomers & astrophysicists
Private group
About
Discussion
Announcements
Members
Events
Photos
Group Insights
Watch Party
Moderate Group
Group Quality
Search this group
Shortcuts
Early career astronome...
Old Girls Network In ...
1
Joined
Notifications
Share
More
Write Post
Photo/Video
Live Video
More
Write something...
Photo/Video
Watch Party
Tag Friends
POPULAR TOPICS IN POSTS
Manage
opportunities (5)
mentoring (3)
Post-PhD applicat...
PhD applications (...
mental health (2)
How to... (1)
Summer schools (0)
Journal clubs (0)
MSc applications ...

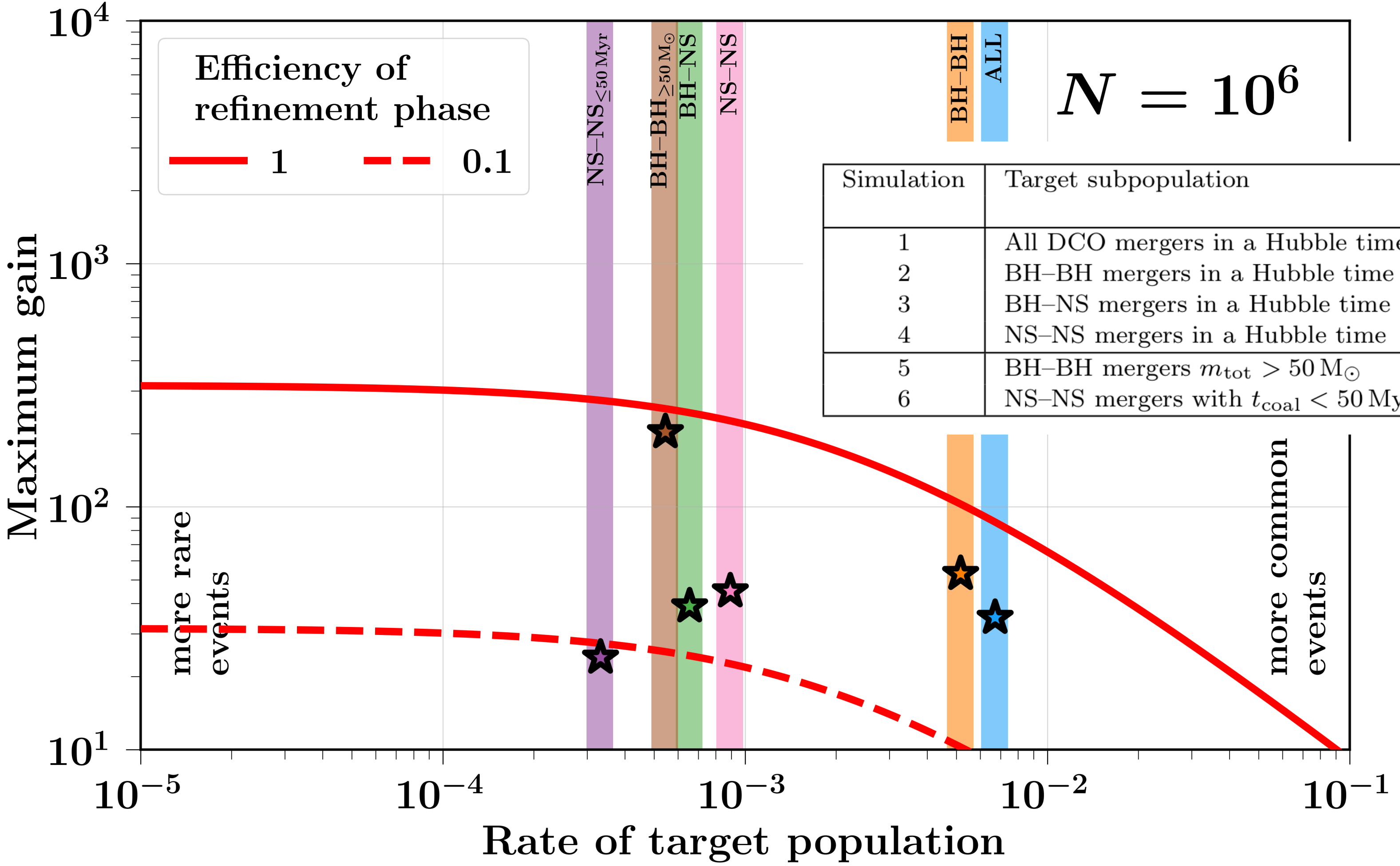
Era Of GWs


80
40
20
10
5
2
1
Solar Masses
LIGO-Virgo Black Holes
BH-NS?
LIGO-Virgo | Frank Elavsky | Northwestern
LIGO-Virgo Neutron Stars

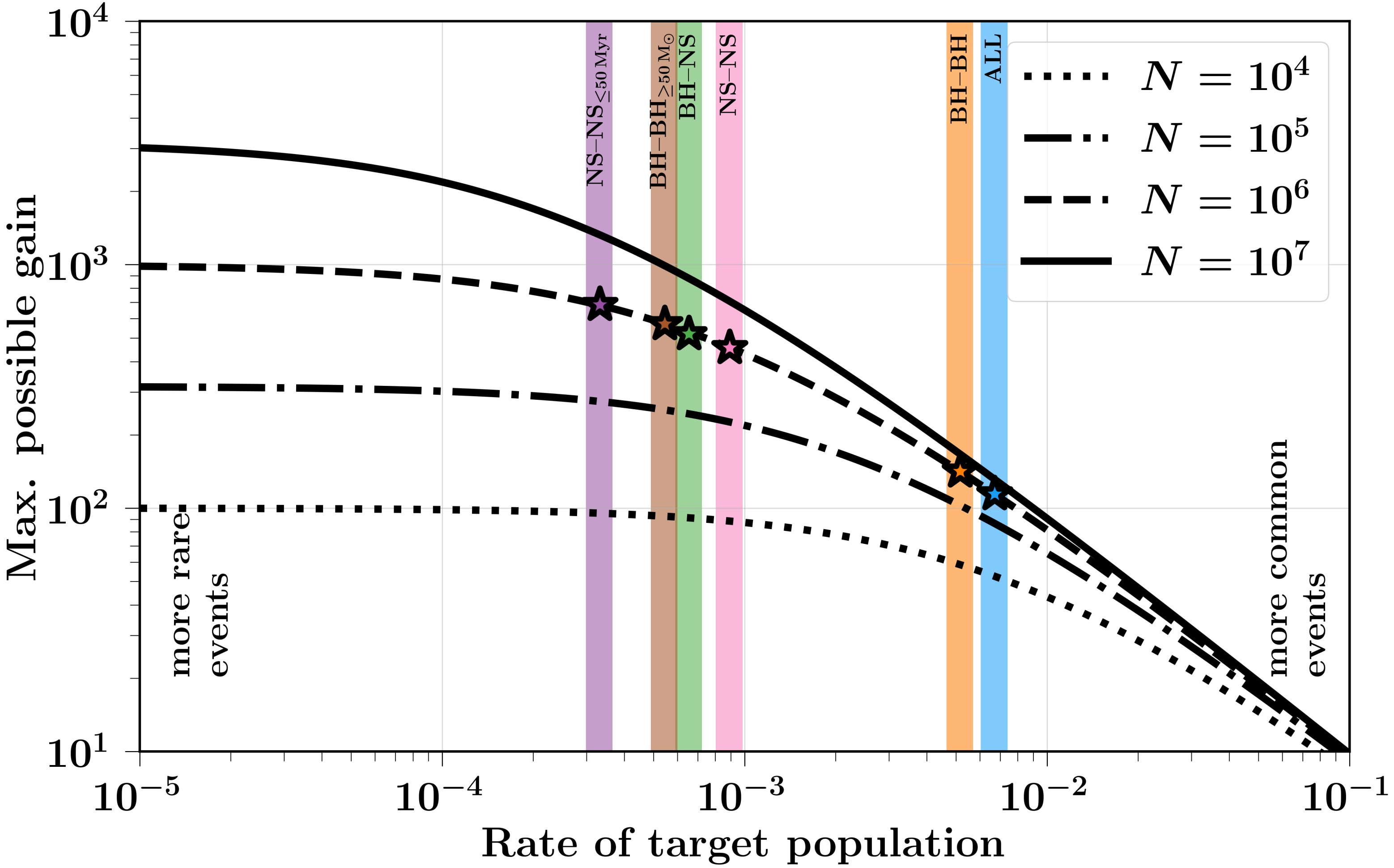
Building up a population of observations

Credit: ESO

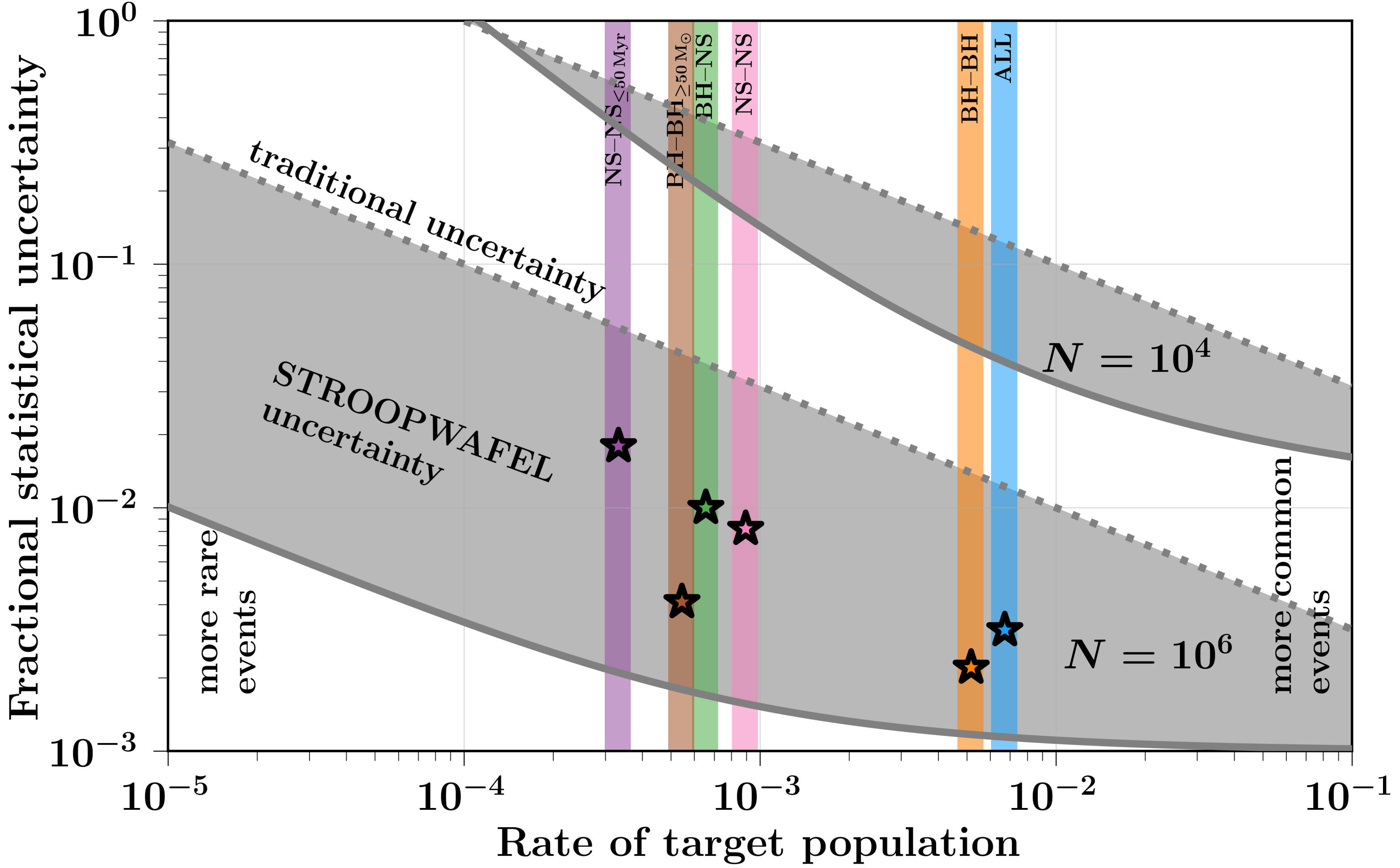

STROOPWAFEL
Traditional
Fractional statistical uncertainty
5%
4%
3%
2%
1%
0
ALL mergers
BH–BH mergers
BH–NS mergers
NS–NS mergers
BH–BH$_{\geq 50\,M_\odot}$ mergers
NS–NS$_{\leq 50\,Myr}$ mergers



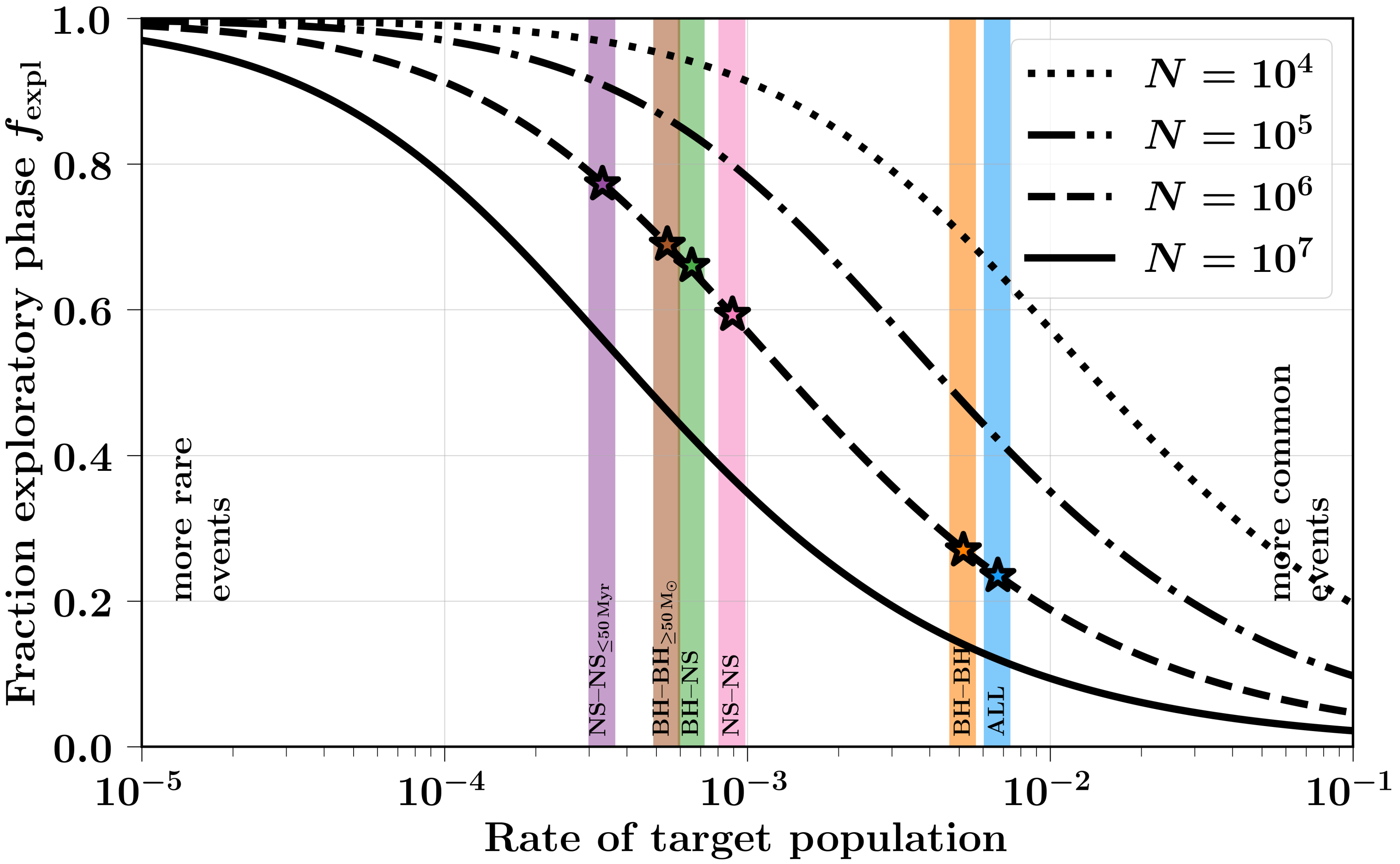
Simulation	Target subpopulation	gain
1	All DCO mergers in a Hubble time	35×
2	BH–BH mergers in a Hubble time	53×
3	BH–NS mergers in a Hubble time	39×
4	NS–NS mergers in a Hubble time	45×
5	BH–BH mergers $m_{\rm tot} > 50\,{\rm M}_\odot$	202×
6	NS–NS mergers with $t_{\rm coal} < 50\,{\rm Myrs}$	24×



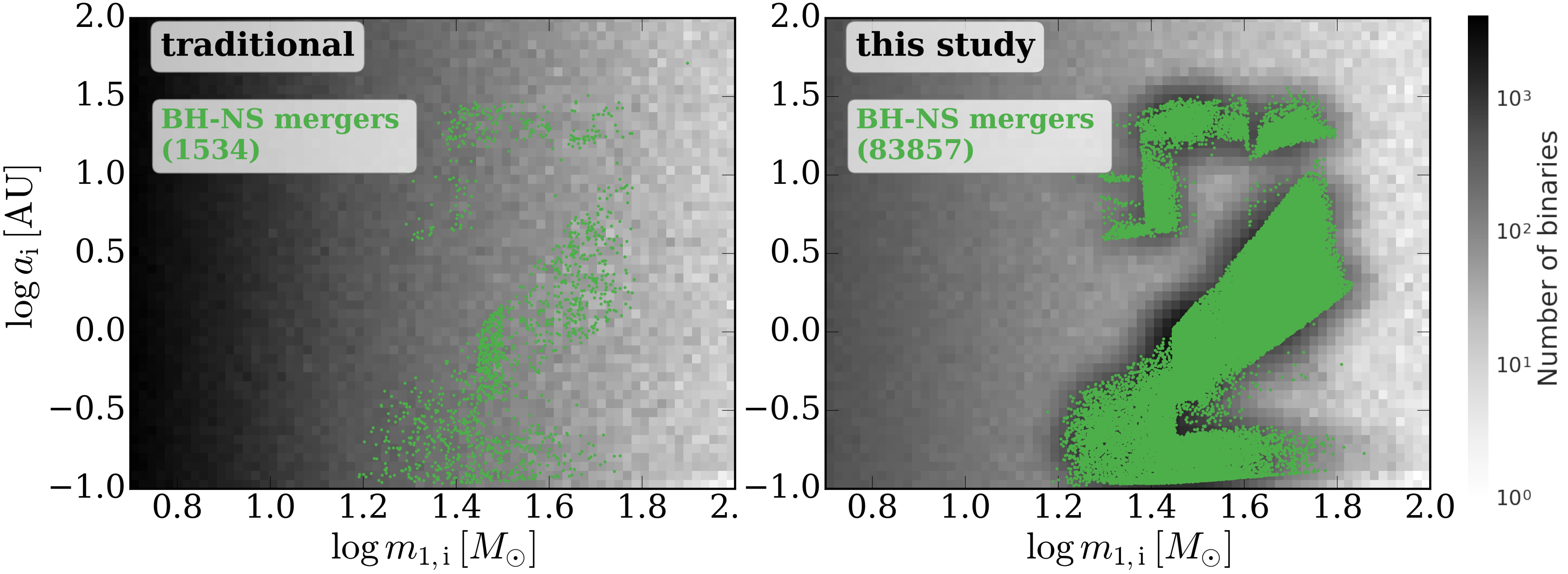
Max. possible gain
Rate of target population
NS–NS≤50 Myr
BH–BH≥50 M⊙
BH–NS
NS–NS
BH–BH
ALL
N = 10^4
N = 10^5
N = 10^6
N = 10^7
more rare events
more common events



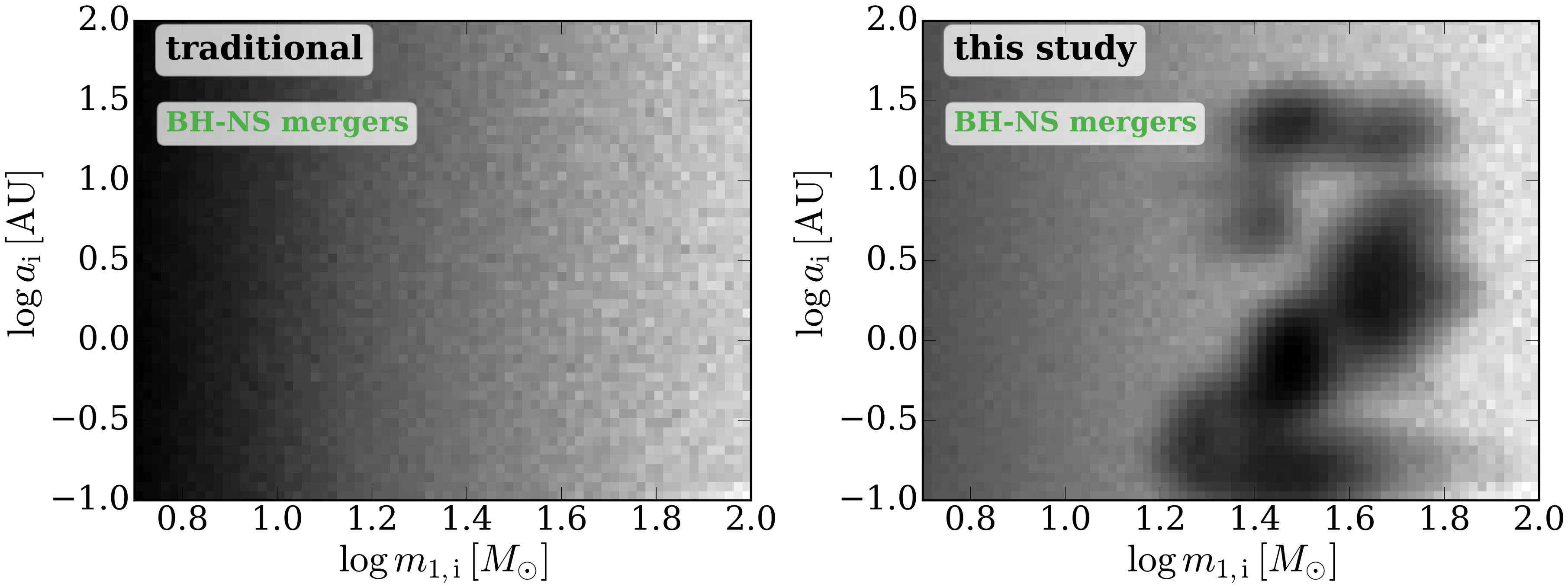
Fractional statistical uncertainty
Rate of target population
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
NS–NS$_{\leq 50\,\mathrm{Myr}}$
BH–BH$_{\geq 50\,\mathrm{M}_\odot}$
BH–NS
NS–NS
BH–BH
ALL
traditional uncertainty
STROOPWAFEL uncertainty
$N = 10^4$
$N = 10^6$
more rare events
more common events

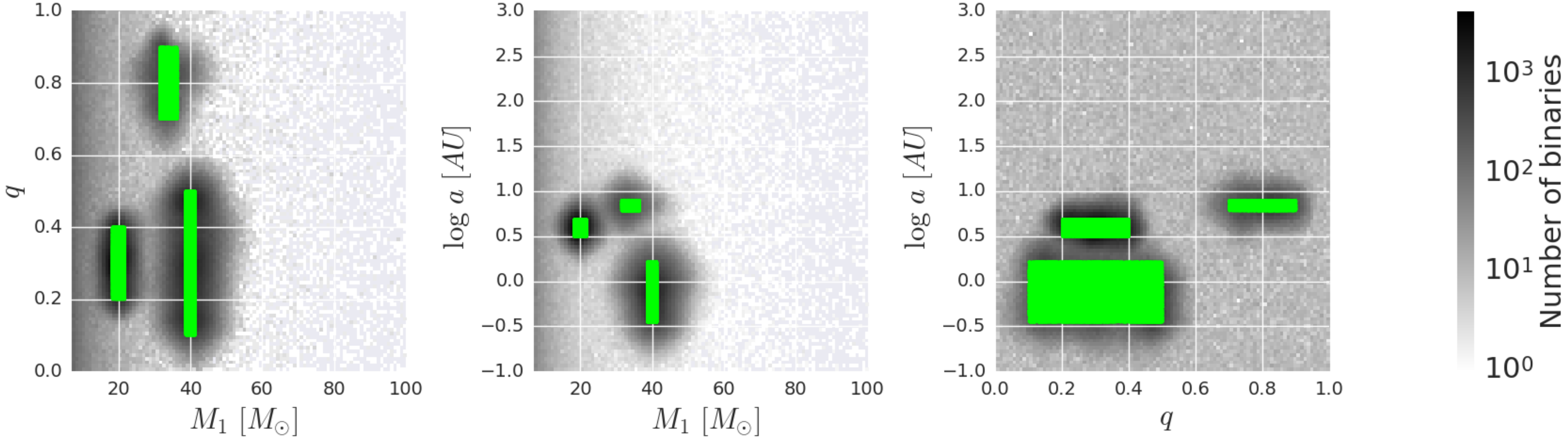

Fraction exploratory phase $f_{\rm expl}$
Rate of target population
$N = 10^4$
$N = 10^5$
$N = 10^6$
$N = 10^7$
more rare events
more common events
NS–NS$_{\leq 50\,\rm Myr}$
BH–BH$_{\geq 50\,\rm M_\odot}$
BH–NS
NS–NS
BH–BH
ALL
1.0
0.8
0.6
0.4
0.2
0.0
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}

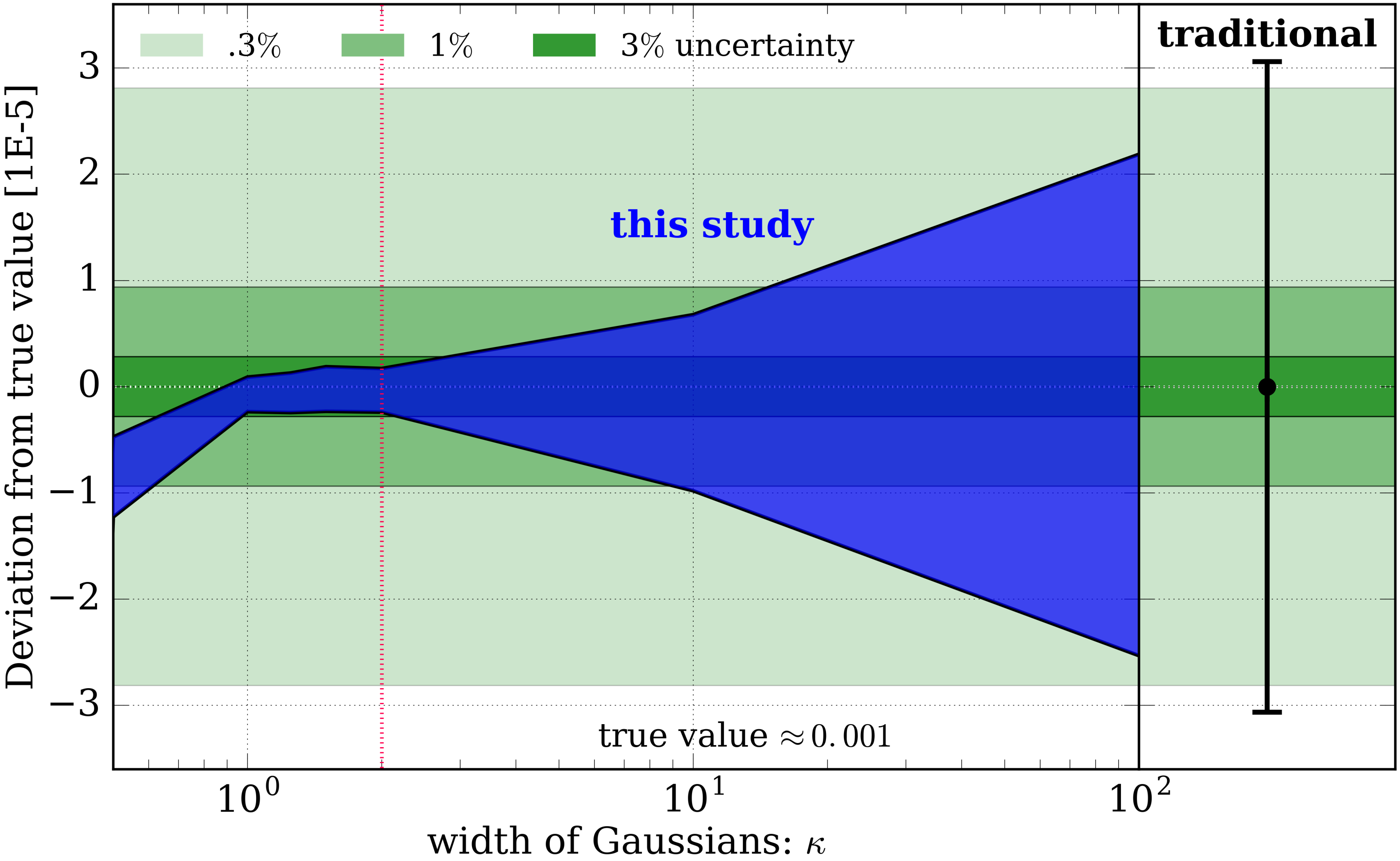
Better resolution progenitors



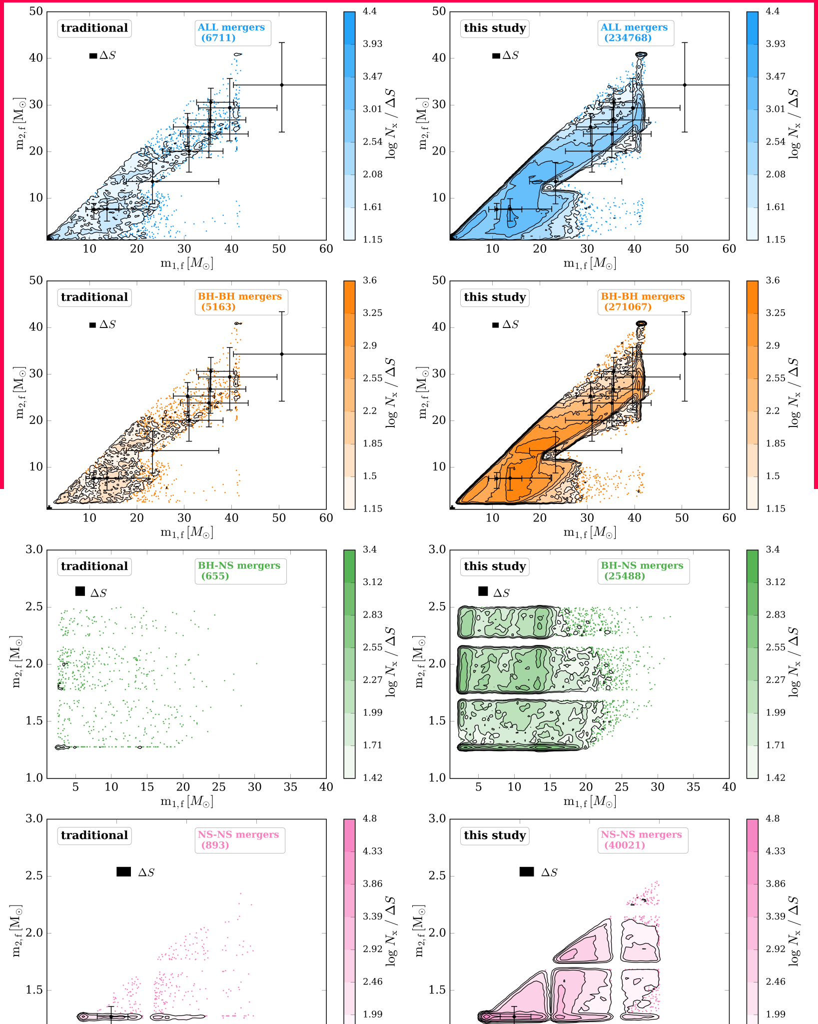
Refinement phase



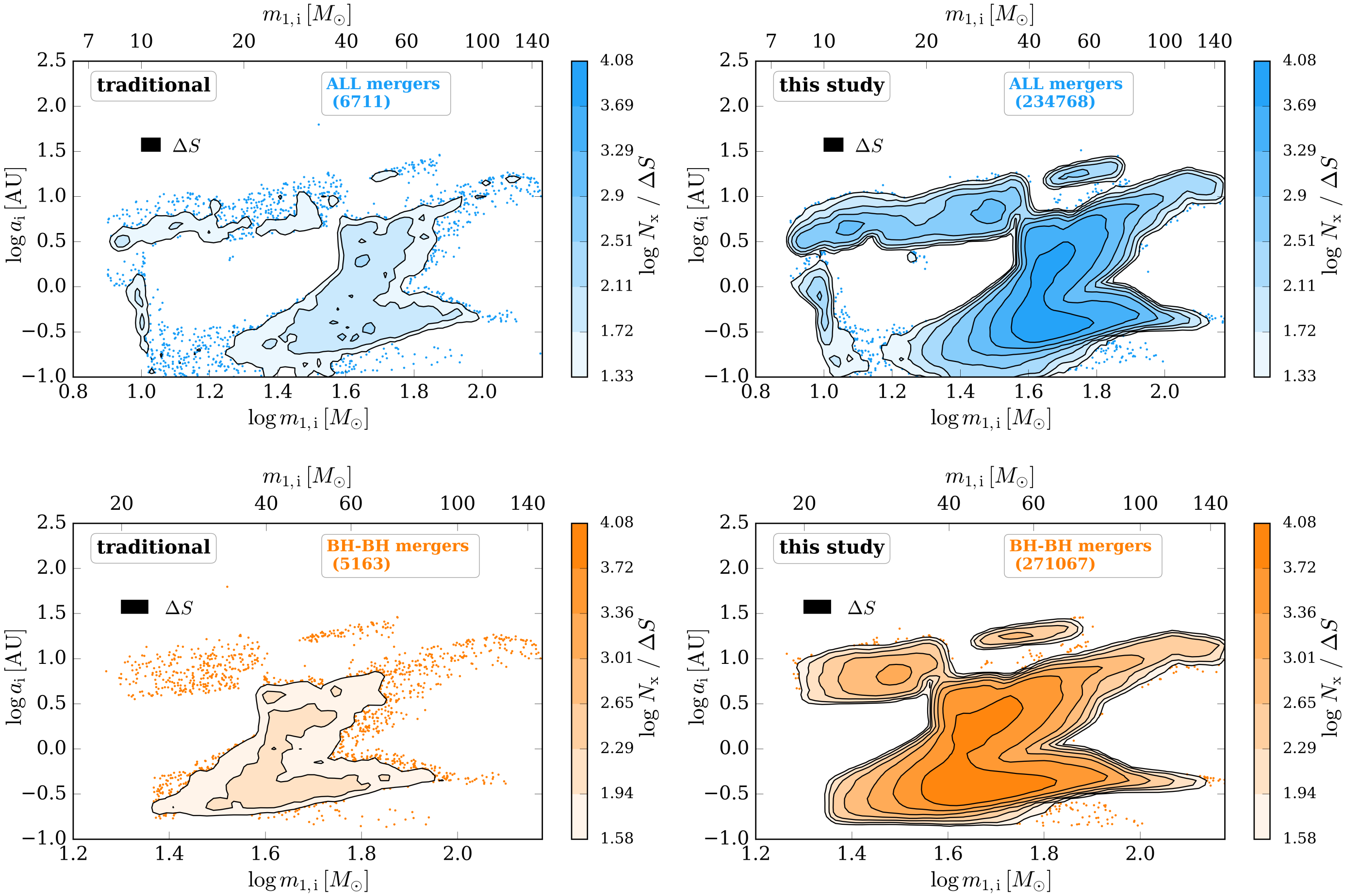

1.0
0.8
0.6
0.4
0.2
0.0
q
20
40
60
80
100
$M_1\ [M_\odot]$
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
$\log a\ [AU]$
20
40
60
80
100
$M_1\ [M_\odot]$
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
$\log a\ [AU]$
0.0
0.2
0.4
0.6
0.8
1.0
q
10^3
10^2
10^1
10^0
Number of binaries

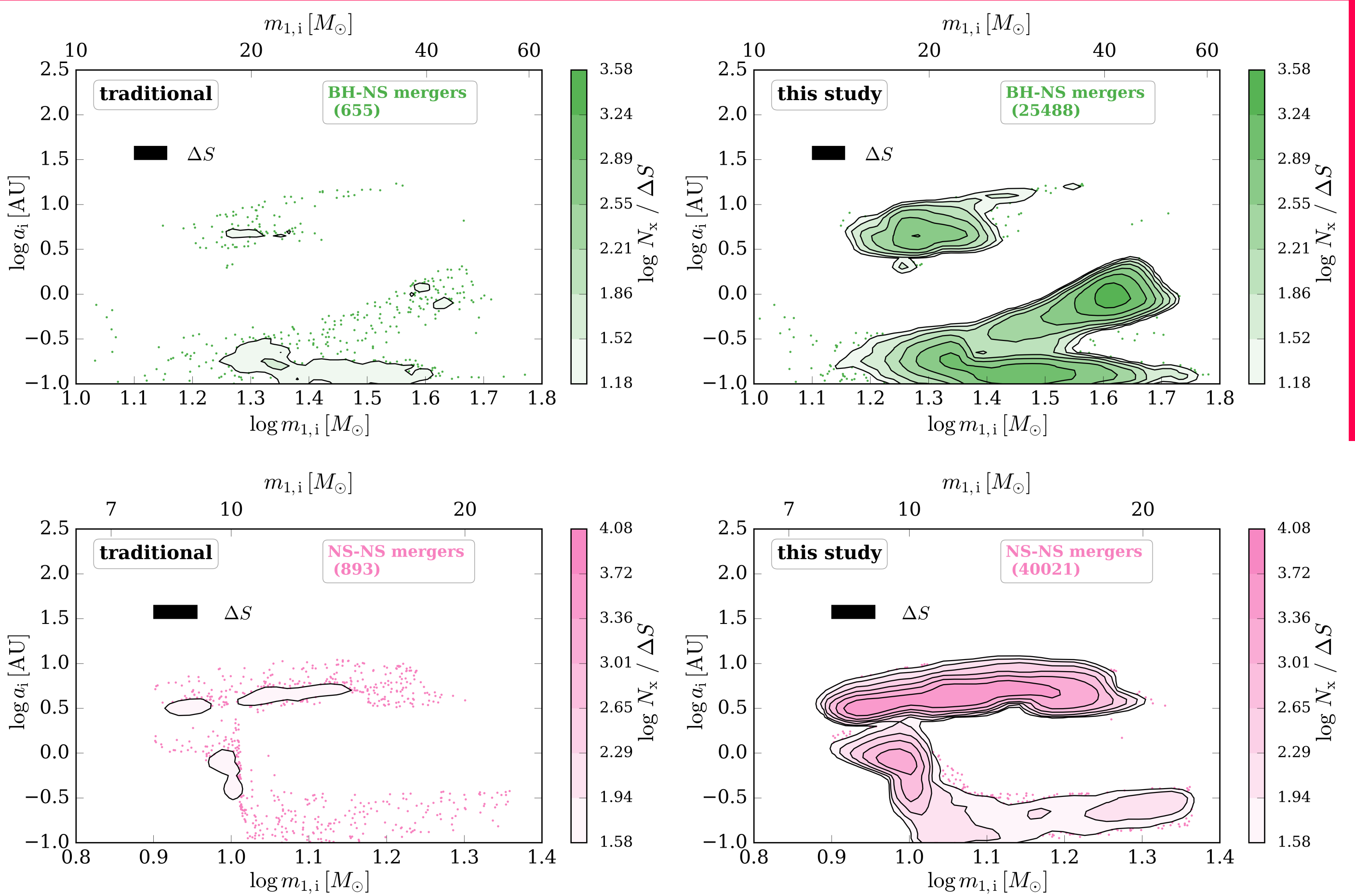


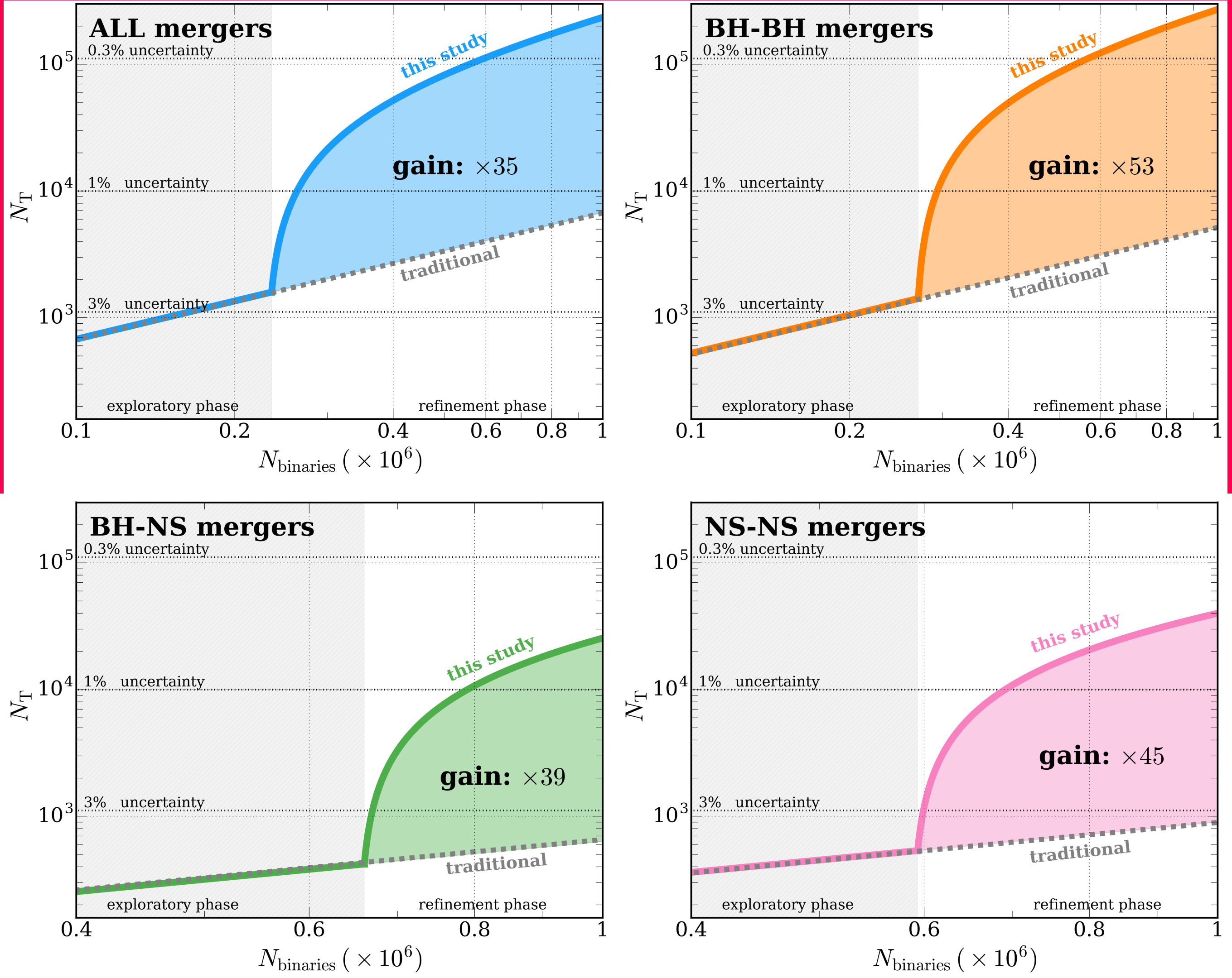
.3%
1%
3% uncertainty
traditional
this study
true value ≈0.001
Deviation from true value [1E-5]
width of Gaussians: κ
3
2
1
0
−1
−2
−3
10^0
10^1
10^2



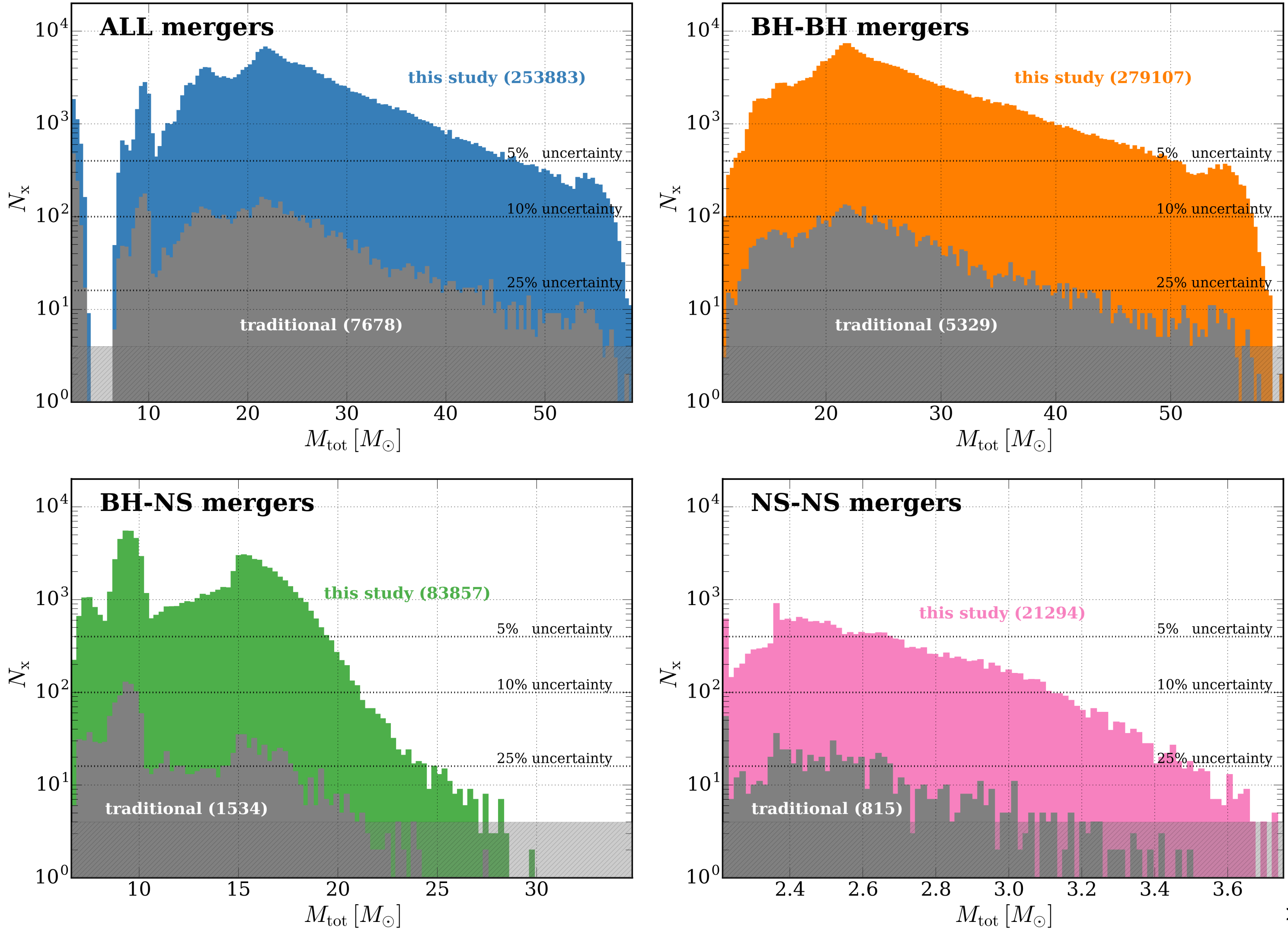
traditional
ALL mergers (6711)
this study
ALL mergers (234768)
traditional
BH-BH mergers (5163)
this study
BH-BH mergers (271067)
traditional
BH-NS mergers (655)
this study
BH-NS mergers (25488)
traditional
NS-NS mergers (893)
this study
NS-NS mergers (40021)
ΔS
$m_{1,f}\,[M_\odot]$
$m_{2,f}\,[M_\odot]$
$\log N_x / \Delta S$

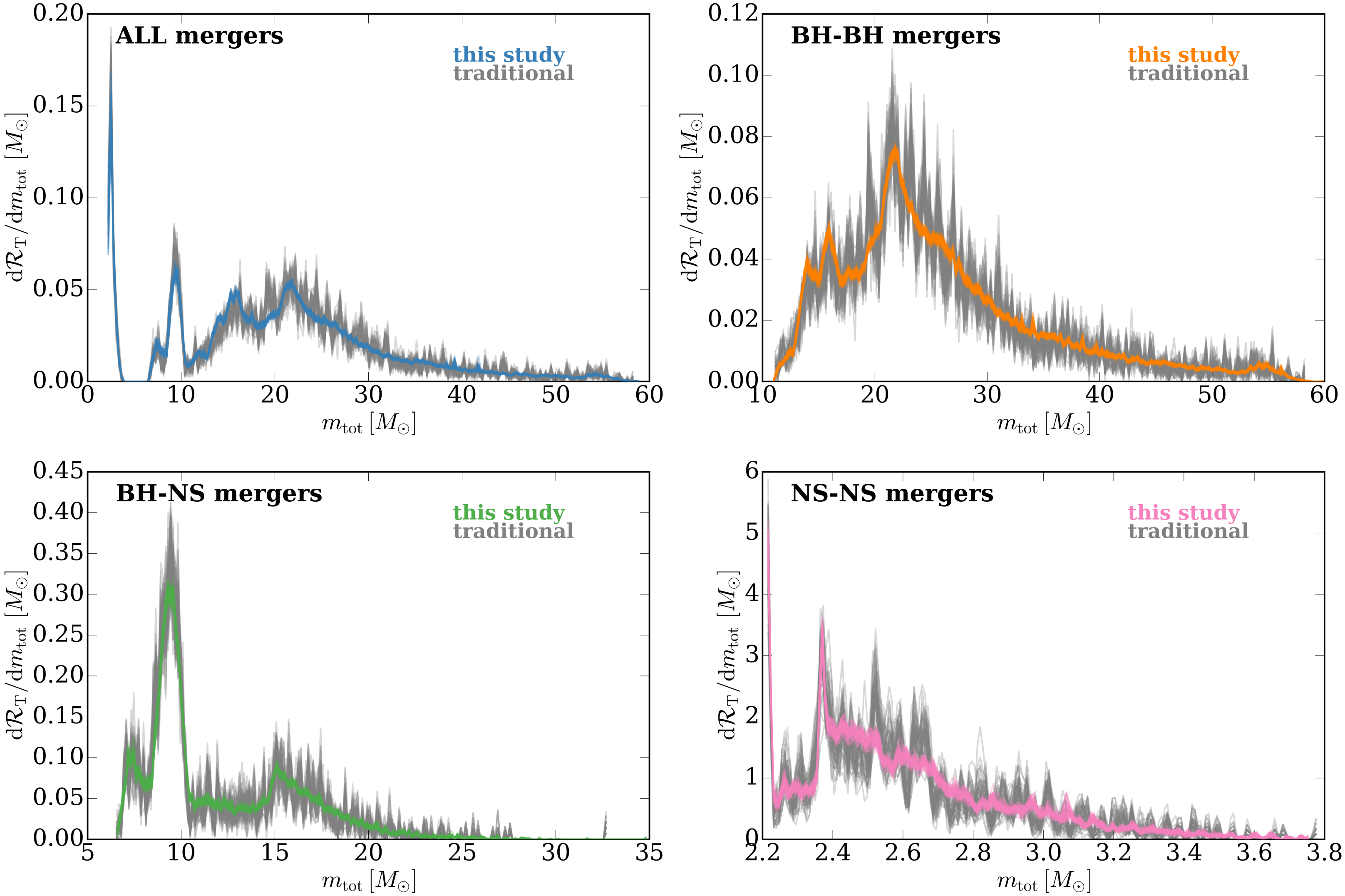

traditional
ALL mergers (6711)
this study
ALL mergers (234768)
traditional
BH-BH mergers (5163)
this study
BH-BH mergers (271067)
ΔS
$m_{1,\mathrm{i}}\,[M_\odot]$
$\log m_{1,\mathrm{i}}\,[M_\odot]$
$\log a_\mathrm{i}$ [AU]
$\log N_\mathrm{x}\,/\,\Delta S$


traditional
BH-NS mergers (655)
this study
BH-NS mergers (25488)
traditional
NS-NS mergers (893)
this study
NS-NS mergers (40021)
ΔS
$\log a_i$ [AU]
$\log m_{1,i}$ [$M_\odot$]
$m_{1,i}$ [$M_\odot$]
$\log N_x / \Delta S$

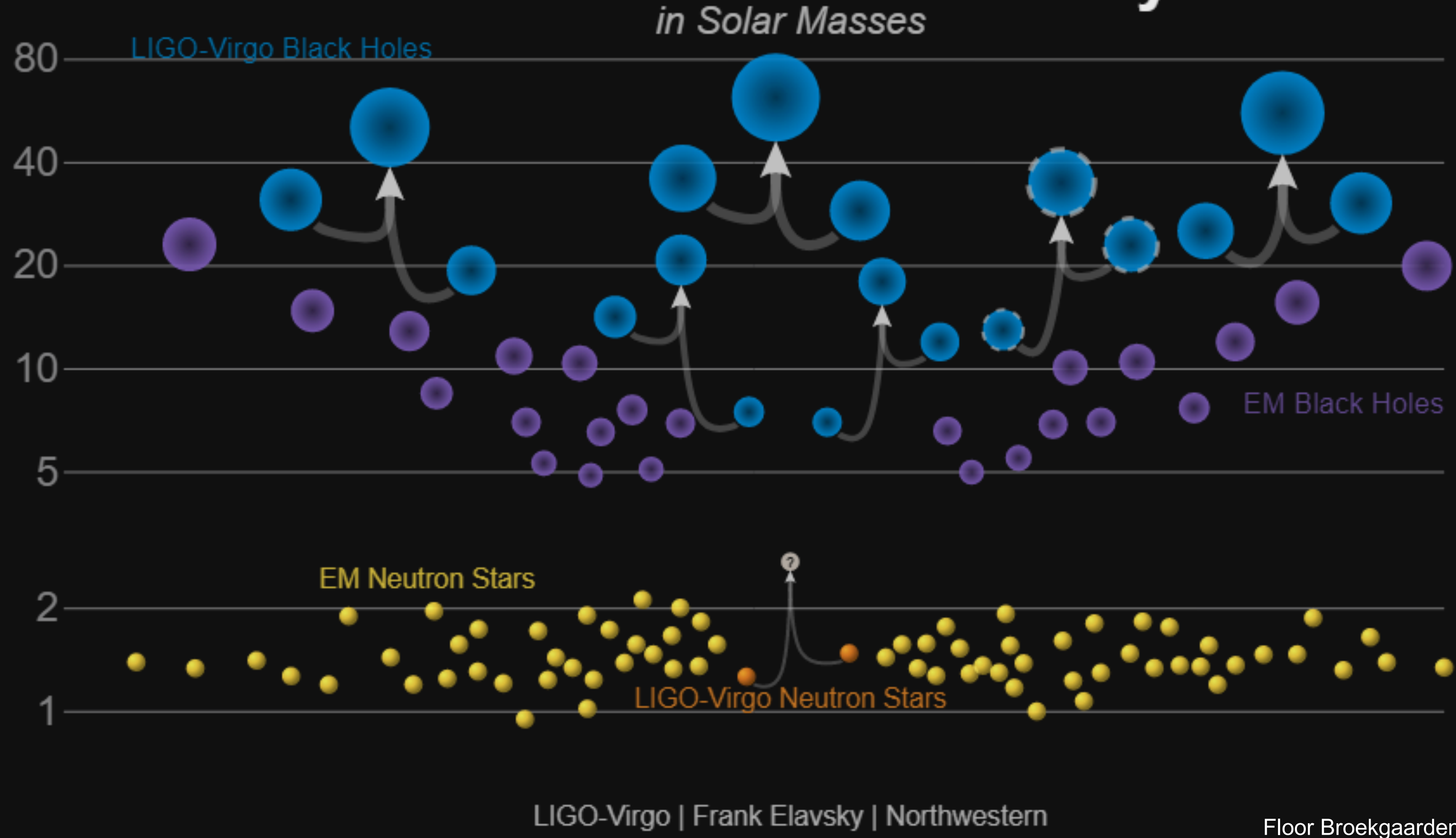


ALL mergers
BH-BH mergers
BH-NS mergers
NS-NS mergers
0.3% uncertainty
1% uncertainty
3% uncertainty
this study
traditional
gain: ×35
gain: ×53
gain: ×39
gain: ×45
exploratory phase
refinement phase
N_T
$N_{binaries}$ (× 10^6)

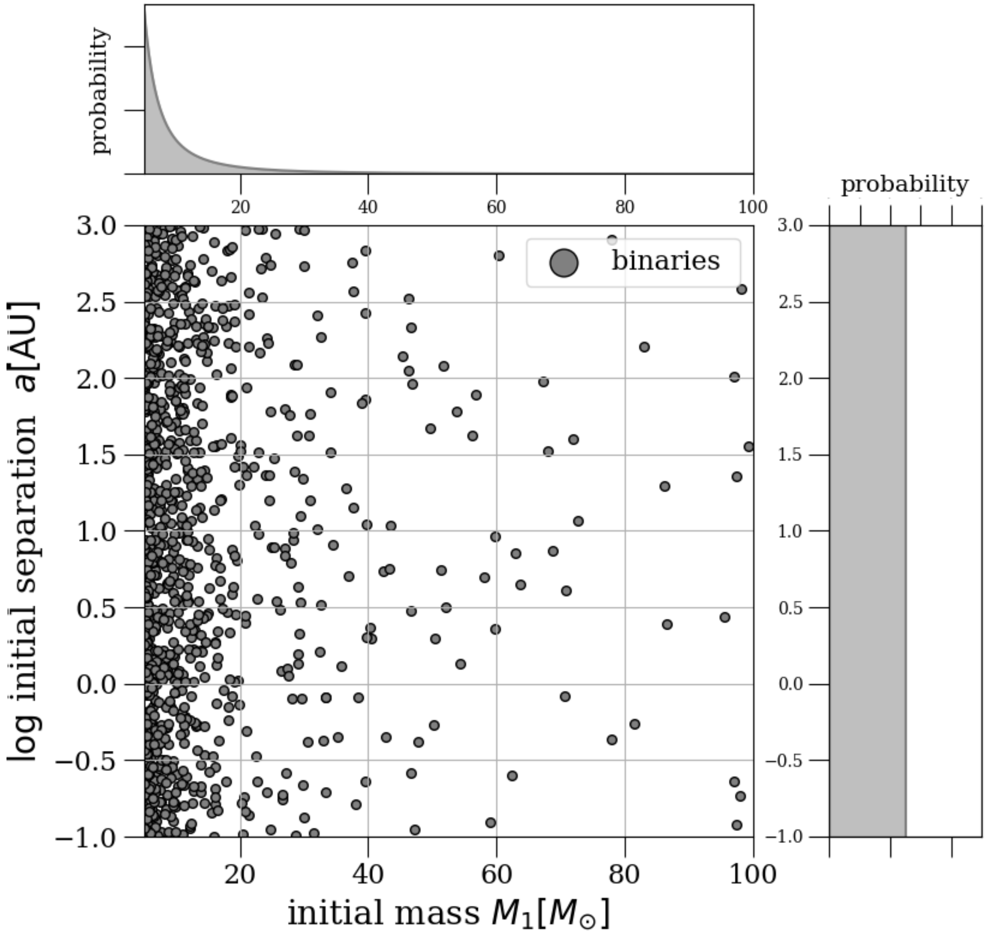

ALL mergers
this study (253883)
traditional (7678)
BH-BH mergers
this study (279107)
traditional (5329)
BH-NS mergers
this study (83857)
traditional (1534)
NS-NS mergers
this study (21294)
traditional (815)
5% uncertainty
10% uncertainty
25% uncertainty
N_x
$M_{tot}\ [M_\odot]$


ALL mergers
this study
traditional
BH-BH mergers
this study
traditional
BH-NS mergers
this study
traditional
NS-NS mergers
this study
traditional
$\mathrm{d}\mathcal{R}_\mathrm{T}/\mathrm{d}m_\mathrm{tot}\ [M_\odot]$
$m_\mathrm{tot}\ [M_\odot]$

Masses in the Stellar Graveyard



1 Populations:



E.g. Belczynski+02