UNIVERSITY OF PADOVA
Department of Physics and Astronomy ‘Galileo Galilei’
MSc in Physics of Data

# Binary Stellar Evolution - Dormant BH in binary system

**Gabriele Bertinelli – Martina Cacciola- Yelyzaveta Pervysheva - Wageesha Widuranga Waththe Liyanage – Poornima Amali Wickramasinghe Wickramasinghe Mudiyanselage**

June 29th , 2023

# Introduction

- Study binaries with a compact object → **GAIA DR3** → large amount of such candidates

- In our project we are interested in **BH-MS binaries**

# Introduction
# SEVN

Difficult to observe these systems → deal with Binary System Evolution (BSE) simulations ➔ **SEVN** (Stellar EVolution for N-body)

- Rapid binary population synthesis code
- Initial conditions → evolution → interpolating PARSEC stellar tracks
- Analytic and semi-analytic prescriptions
- Flexible, low computation time → easy to change or update it

# Goal of the project

- Understand what kind of processes these systems are likely to experience during their lifetime → DNN and XGBoost
  - understand the **importance** that each feature has had in the evolution of systems
  - **classify** the GAIA candidates
  - retrieve the guessed full **evolution history** of candidates

# Simulated data
## Selection

- Systems with **metallicities** of 0.02 (solar-like systems), 0.001 (metal-poor), and 0.0001 (very metal-poor)
- Three different **α values**: 0.1 (indicating low efficiency), 1 (intermediate efficiency), and 5 (high efficiency).

- We utilized `Dask` to deal with the large amount of data

# Simulated data
## Selection

- Select **BH+MS** systems

```
mask = (
          ((outputs['RemnantType_0'] == 6) & (outputs['RemnantType_1'] == 0) &
          (outputs['Phase_1'] == 1)) | ((outputs['RemnantType_1'] == 6) &
          (outputs['RemnantType_0'] == 0) & (outputs['Phase_0'] == 1))
        )
bh_star_df = outputs[mask]
```

# Simulated data
## Selection

- Dividing between interacting and non-interacting systems → **Roche Lobe** radius

$$\frac{r1}{a} = \frac{0.49\, q^{2/3}}{0.6\, q^{2/3} + \ln\left(1+q^{1/3}\right)}$$

- If the radius of the star or the BH in the system is greater than the Roche Lobe radius r1, then the binary is classified as interacting. Otherwise, for r≤r1 it is non-interacting.

# Simulated data
## Selection

- Add **period** and converted in log

```
period = (2*m.pi*np.sqrt( semimaj**3 / (const.G * mass_tot) )) / s_to_day

int_binsys['logP'] = np.log10(period)
```

# Simulated data
## Elapsed BWorld time

- For **visualization purposes** only
- Restricting to non-interacting → retain only the first row → initial time of the BH+MS phase
- From initial dataset → systems (bound and non-interacting) composed of a BH and a star no more in MS phase (`Phase_* > 1`) → retain only the first row → end time of the BH+MS phase
- Calculate the difference

# Simulated data
## Evolution channel selection

- SEVN records each phase the two objects underwent

- Restricting to the selection to:
  - Non-Interacting (**NI**)
  - (stable) Mass Transfer (**MT**)
  - (at least one) Common Envelope (**CE**)

- Create series of masks → apply them in succession

# Simulated data

## Evolution channel selection

- Selection based on `BEvent` column

```
# Selecting systems that are in the evolutionary phase of two stars
mask_3 = (((outputs['RemnantType_0'] == 0)) & (outputs['RemnantType_1'] == 0))
```

```
# Creating a mask to select systems that went through at least 1 Common Envelope
mask_CE = msms_df['BEvent'].isin([7, 9, 11, 12, 14, 15])
```

```
# Creating a mask to select systems that went through a stable Mass Transfer AND NOT a CE
mask_MT1 = (out_masked['BEvent'].isin([4, 5, 6, 8, 10, 13, 16, 17, 18, 19, 20]))
```

```
# Filling the remaining rows (NaN) with the Non Interacting flag
out_masked_1['type'] = out_masked_1['type'].fillna('NI')
```

# Simulated data
## Evolution channel selection

- **Merge** the new dataframes

  - non-interacting BH+MS systems (`non_int_binsys`)

  - evolution types (`type_df`)

  - elapsed time (`df_bwt`)

| | ID | name | Mass_BH | Mass_1 | Eccentricity | logP | z | alpha | BEvent | type | elapsed_bwt | tot_mass |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 80 | 0_201673565337120 | 17.66923 | 28.74325 | 0.212435 | 4.221513 | 0.0200 | 0.5 | -1 | NI | 1.621285 | 46.41248 |
| 1 | 87 | 0_929528790266714 | 23.50375 | 44.82448 | 0.752931 | 5.343398 | 0.0200 | 0.5 | -1 | NI | 0.180697 | 68.32823 |
| 2 | 109 | 0_889165204871227 | 21.41187 | 34.20442 | 0.227290 | 4.468927 | 0.0200 | 0.5 | -1 | NI | 0.026888 | 55.61629 |
| 3 | 122 | 0_111281974618981 | 17.45101 | 19.13807 | 0.326498 | 5.457667 | 0.0200 | 0.5 | -1 | NI | 4.488835 | 36.58908 |
| 4 | 226 | 0_363554648261798 | 16.08790 | 28.28335 | 0.112235 | 3.496335 | 0.0200 | 0.5 | -1 | NI | 1.450930 | 44.37125 |

# Confirmed BH-MS systems

| | ID | name | Mass_BH | dMass_BH | Mass_1 | dMass_1 | Period | dPeriod | logP | dlogP | Eccentricity | dEccentricity | loc | system | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | R00 | BH1 | 9.6 | 0.18 | 0.93 | 0.05 | 185.59000 | 0.05000 | 2.268555 | -1.301030 | 0.450 | 0.005 | MW | disc-field | solar |
| 1 | R01 | VFTS 243 | 10.1 | 2.00 | 25.00 | 2.30 | 10.40310 | 0.00040 | 1.017163 | -3.397940 | 0.017 | 0.012 | LMC | OC | sub-solar |
| 2 | R02 | HD130298 | 7.7 | 1.50 | 24.20 | 3.80 | 14.62959 | 0.00085 | 1.165232 | -3.070581 | 0.457 | 0.007 | MW | runway-field | solar |
| 3 | R03 | BH2* | 8.9 | 0.30 | 1.07 | 0.19 | 1300.00000 | 26.00000 | 3.113943 | 1.414973 | 0.515 | 0.010 | MW | field | solar |

# Data distribution

# Data distribution

# Classification models

- **Combine** dataset created before
- **Shuffle** to add more randomness → removes patterns or biases
- **Balance** the dataframe → select equal number of systems for each label

```
data_tot = pd.concat([data_a05, data_a1, data_a5])

NI = shuffle(data_tot[data_tot['type']=='NI']).head(430242)
MT = shuffle(data_tot[data_tot['type']=='MT']).head(430242)
CE = shuffle(data_tot[data_tot['type']=='CE'])

balanced_data = pd.concat([NI, MT,CE])

shuffle_data = shuffle(balanced_data)
```

# Deep Neural Network (DNN)

- Type of **artificial neural network**
- Particularly effective for supervised learning tasks → classification
- Composed of:
  - Multiple **layers** of interconnected nodes called **neurons**
  - Neurons receive input from the previous layer and produce an output
  - Connections → assigned **weights** → strength/importance of the connection
- DNN learn by adjusting weights during training → optimization algorithms → e.g., gradient descent → ADAM

# Deep Neural Network (DNN)

- "**Deep**" → multiple hidden layers → capture intricate patterns and relationships
- For classification tasks → output layer consists same number of neurons of the classes you want to classify
- DNN assign  probability to each class based on the learned weights and biases
- The class with the highest probability is chosen as the predicted class

# Deep Neural Network (DNN)
## Our model

- Several tests and “gridsearches” → optimal structure
- **4** hidden layers, **20** neurons each
- Input layer → n° neurons = n° of input features
- Output layer → n° neurons = 3 (n° of classes)

- Activation function → **ReLU**
- Optimizer → **Adam**
- Loss function → **categorical cross-entropy**

# Deep Neural Network (DNN)
## Our model

- **Epochs** → 100-200 → 150
- **Batch size** → 450-750 → 750

➔ balance between model performance and computational efficiency

- Training → all the **features**
  → `Mass_BH`, `Mass_1`, `Eccentricity`, `logP`, `z`, `alpha`
  → scaling of the first four features

# Deep Neural Network (DNN)
## Results

- Performance metrics → **accuracy** and **loss**

# Deep Neural Network (DNN)
## Results

- Confusion matrix

# Deep Neural Network (DNN)
## Results

- Classification report
  - **Precision** → TP/(TP+FP)
  - **Recall** → TP/(TP+FN)
  - **F1 score** → harmonic mean of precision and recall
  - **Accuracy** → (TP+TN)/(TP+FP+TN+FN)

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| CE | 0.97 | 0.97 | 0.97 | 86412 |
| MT | 0.96 | 0.96 | 0.96 | 85576 |
| NI | 0.99 | 0.99 | 0.99 | 86158 |
| accuracy | | | 0.97 | 258146 |

# Deep Neural Network (DNN)
## Classification

- Same structure as before
- Added a **dropout layer** after each hidden layer → prevent overfitting
- **Epochs** → 100-200 → 150
- **Batch size** → 1000

- Training → **features** present in the GAIA dataset
  → `Mass_BH, Mass_1, logP`
  → scaling of the first four features

# Deep Neural Network (DNN)
## Classification

- Classification report

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| CE | 0.89 | 0.88 | 0.88 | 86051 |
| MT | 0.87 | 0.88 | 0.87 | 86181 |
| NI | 0.97 | 0.98 | 0.98 | 85914 |
| accuracy | | | 0.91 | 258146 |

# Different classifier

- Conducted experiments using various classification algorithms
- Evaluated and ranked based on their performance metrics

# Different classifier

- `RandomForestClassifier`, `BaggingClassifier`, `ExtraTreesClassifier`, `DecisionTreeClassifier`, and `XGBClassifier` demonstrated similar performances
- We've chosen to utilize **`XGBoost`**
  - **Sequential Ensemble Building** → each subsequent model rectifies the errors made by the preceding models
  - **Gradient Boosting Framework** → iteratively optimizing a objective function to enhance performances
  - **High Predictive Power and Efficiency** → optimize computational efficiency → handle large datasets and complex problems

# XGboost

- Algorithm used for **supervised** learning tasks
- Ensemble learning method → combine weak decision trees → create strong model

- Works by iteratively training and adding decision trees to the ensemble
- Identifies areas where the previous trees have made errors → correct them
- Minimize loss function by learning from gradients errors
- Final prediction → aggregating the predictions from all the trees

# XGBoost
## Our model

- Gridsearch analysis → default model is the best (when considering all features)

## Results

```
Accuracy: 0.979038993437822
Precision: 0.9790641302747958
Recall: 0.979038993437822
F1-Score: 0.9790338071672348
```

```
Classification Report:
              precision    recall  f1-score   support

          CE       0.98      0.97      0.97     85954
          MT       0.97      0.97      0.97     86443
          NI       0.99      0.99      0.99     85749

    accuracy                           0.98    258146
```

# XGBoost
## Classification

- Gridsearch analysis:
  - **Learning rate** → 1
  - **n° estimators** → 300
  - **Max depth** → 7

## Results

```
Accuracy: 0.9226019384379383
Precision: 0.9225351950821429
Recall: 0.9226019384379383
F1-Score: 0.9225555799194637
```

```
Classification Report:
              precision    recall  f1-score   support

          CE       0.90      0.89      0.89     86117
          MT       0.89      0.89      0.89     86190
          NI       0.98      0.99      0.98     85839

    accuracy                           0.92    258146
```

# Naïve Bayesian Classifier

- Simple **probabilistic** ML algorithm → Bayes' theorem
- Strong assumption on **independence** among features
  → presence or absence of one feature does not affect the model
- Calculates the probability of each class → selects the class with the highest probability
- → **unsuitable** → because of features independence
- Suboptimal results → accuracy < 80%

# SHAP

- SHapley Additive exPlanations → explain the predictions of ML models

- Core idea → SHAP values → distributing the payout of a cooperative game among the players

- Based on cooperative game theory
  → allocate the contribution of each feature in a prediction model to the final prediction
  → quantify the impact of each feature on the prediction → considering all combinations

# SHAP

Several advantages:

1. Unified framework for feature importance measurement

2. Consistent → if a feature is removed or added, the SHAP values change accordingly

3. Locally/Globally accurate → explain individual predictions/overview of features importance

# SHAP
# DNN

- Mean absolute SHAP value → magnitude of each feature's contribution towards the predicted label

# SHAP
# DNN

- Violin plot → most useful type
- Representation of the distribution of SHAP values for different features

# SHAP
# DNN

- **Vertical axis**: features in descending order of importance
- **Horizontal axis**: range of SHAP values
- **Violin shape**: density of the SHAP values for a specific feature
- **Color**: original value

**DNN**

Class 2 - NI
logP
Mass_BH
z
Eccentricity
Mass_1
alpha
−0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8
SHAP values (impact on output model)
High
Feature value
Low
Class 1 - MT
logP
Mass_1
Mass_BH
z
Eccentricity
alpha
−0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
SHAP values (impact on output model)
High
Feature value
Low
Class 0 - CE
Mass_BH
Mass_1
logP
z
Eccentricity
alpha
−1.25 −1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50
SHAP values (impact on output model)
High
Feature value
Low

**XGB**

Class 2 - NI
logP
Mass_BH
z
Eccentricity
Mass_1
alpha
−10 −5 0 5 10
SHAP values (impact on output model)
High
Feature value
Low
Class 1 - MT
logP
Mass_1
Mass_BH
Eccentricity
z
alpha
−6 −4 −2 0 2 4 6
SHAP values (impact on output model)
High
Feature value
Low
Class 0 - CE
logP
Mass_1
z
Mass_BH
Eccentricity
alpha
−8 −6 −4 −2 0 2 4
SHAP values (impact on output model)
High
Feature value
Low

# Results interpretation (DNN)

## NI

- **logP**: High values of the period contribute greatly

- **Mass_BH**: intermediate values → no interaction → higher stellar mass → higher mass of the BH

- **Z**: based on stellar tracks → higher z values → higher radius during He-b → exceeding RL
  ➔ preference for low-medium values for NI
  → lacks definitive evidence

- **Eccentricity**: slight contribution
- **Mass_1**: seems not contribute
- **Alpha**: seems not contribute

# Results interpretation (DNN)

## MT

- **logP**: Low period values contribute significantly
- **Mass_1 & Mass_BH**: during stars evolution, mass ratio < QC → this info have influenced the mass values in the BH-MS phase
  → it's not crystal clear if the model had interpreted correctly these features
- **Z**: high metallicity values contribute
- **Eccentricity**: Slight contribution
- **Alpha**: seems not contribute

# Results interpretation (DNN)

## CE

- **Mass_1 & Mass_BH**: during stars evolution, mass ratio > QC → this info have influenced the mass values in the BH-MS phase
  → it's not crystal clear if the model had interpreted correctly these features

- **LogP & Eccentricity**: low values contribute greatly

- **Z**: inconsistent from an astrophysical point of view

- **Alpha**: seems not contribute

# Results interpretation (XGB)

## NI

- **Z**: inconsistent from an astrophysical point of view
- **Eccentricity**: inconsistent from an astrophysical point of view
- **Mass_1**: not clear what is the contribution

# Results interpretation (XGB)

## MT

- **Eccentricity**: slight contribution given by low to medium values of eccentricity
- **Z**: no great contribution. Inconsistent from an astrophysical point of view

# Results interpretation (XGB)

## CE

Same as for the DNN

# Results interpretation
## Conclusions

- Enhance the model's understanding and recognition of the threshold (QC)
- Inclusion of additional training data → information regarding the initial mass ratio or the mass ration in each phase experienced by the donor star.
- In our training experiments →including the mass ratio of the BH-MS → notable decrease in the model's accuracy.
  → we excluded this feature

# Candidates Classification

- Classify GAIA DR3 candidates → trained models
- For statistical analysis → **sampled** 10.000 gaussian samples for each system
- The classification generates a dataset → percentage for each label

| pred_type | CE | MT | NI | total_count | CE_percentage | MT_percentage | NI_percentage |
|---|---|---|---|---|---|---|---|
| ID | | | | | | | |
| 9720614798216320 | 7255.0 | 2561.0 | 184.0 | 10000.0 | 72.55 | 25.61 | 1.84 |
| 9881104840594176 | 1152.0 | 3607.0 | 5241.0 | 10000.0 | 11.52 | 36.07 | 52.41 |
| 14686108092825984 | 695.0 | 262.0 | 9043.0 | 10000.0 | 6.95 | 2.62 | 90.43 |
| 30994481168584576 | 0.0 | 109.0 | 9891.0 | 10000.0 | 0.00 | 1.09 | 98.91 |

# Candidates Classification

- Setting a threshold based on statistical significance, such as 1, 2, or 3 standard deviations

**DNN**

| Label | $\sigma \geq 1$ | $\sigma \geq 2$ | $\sigma \geq 3$ |
|---|---|---|---|
| NI | 1301 (28.0%) | 762 (16.4%) | 418 (9.0%) |
| CE | 527 (11.34%) | 61 (1.31%) | 17 (0.37%) |
| MT | 679 (14.61%) | 2 (0.04%) | 2 (0.04%) |

**XGB**

| Label | $\sigma \geq 1$ | $\sigma \geq 2$ | $\sigma \geq 3$ |
|---|---|---|---|
| NI | 1271 (27.36%) | 648 (13.95%) | 231 (4.97%) |
| CE | 393 (8.46%) | 45 (0.97%) | 6 (0.13%) |
| MT | 543 (11.69%) | 2 (0.04%) | 2 (0.04%) |

# Candidates Classification

- Comparing simulated data with candidate systems
- Narrow down to "**gold**" plating → more reliable
- From simulated data → **z=0.02** → systems observed by GAIA exhibit solar-like Z
- Consistent agreement between the two models
- **Goal** → assess how well SEVN model capture the characteristics of real binary systems

Common Envelope - CE
Common Envelope - MT
Common Envelope - NI
Mass_BH [M_Sun]
Mass_1 [M_Sun]
logP [log(day)]
Simulated data distribution

Common Envelope - CE
Common Envelope - MT
Common Envelope - NI
Mass_BH [M_Sun]
Mass_1 [M_Sun]
logP [log(day)]
Simulated data distribution

# Candidates Classification

- Close correspondence in the logP-Mass_BH phase space
- When considering Mass_1 (mass of the MSs) → significant discrepancy for values below 10 Msun
- Challenge in providing a definitive explanation → current models do not anticipate so lightweight MSs

# Candidates Classification

Possible explanations:

- Reevaluating **mass transfer mechanisms**: mass transfer is more significant
- Limitations of current **stellar tracks**: min MZAMS = 2.2 Msun → absence of tracks for the objects observed by GAIA

# Similar systems

- Find similar systems to the real candidates of BH-MS binary systems and predicting the history of the evolution

$$d_{Sevn,Real} = \sqrt{\left(\frac{Sec_M_{Sevn} - Sec_M_R}{S_M_{upper} - S_M_{lower}}\right)^2 + \left(\frac{MS_M_{Sevn} - MS_M_R}{MS_M_{upper} - MS_M_{lower}}\right)^2 + \left(\frac{P_{Sevn} - P_R}{P_{upper} - P_{lower}}\right)^2}$$

- Selected most similar 10 systems and 1000 systems for the analysis

| Index | Source Id | BH Mass | MS Mass | Period |
|---|---|---|---|---|
| 44 | 6.70435e+18 | 10.34443 | 0.9273589 | 0.04849124 |

| | Name | Mass_BH | Mass_1 | P | z | alpha | type | value |
|---|---|---|---|---|---|---|---|---|
| 0 | 0_420895437325839 | 8.291450 | 2.310986 | 0.044914 | 0.001 | 1.0 | CE | 0.622455 |
| 1 | 0_617684874163902 | 7.746563 | 2.305923 | 0.052374 | 0.020 | 0.5 | CE | 0.637278 |
| 2 | 0_478118807791845 | 7.422293 | 2.278359 | 0.043486 | 0.020 | 0.5 | CE | 0.641886 |

NI
MT
CE
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
P
3.5
3.0
2.5
2.0
1.5
1.0
Mass_1
6
7
8
9
10
11
12
13
Mass_BH

NI
MT
CE
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
P
13
12
11
10
9
8
7
6
Mass_BH
3.5
3.0
2.5
2.0
1.5
1.0
Mass_1

The graph of Mass_1 vs Mass_BH
Mass_1
Mass_BH
3.5
3.0
2.5
2.0
1.5
1.0
6
7
8
9
10
11
12
13

The graph of P vs Mass_BH
P
Mass_BH
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
6
7
8
9
10
11
12
13

The graph of P vs Mass_1
P
Mass_1
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
1.0
1.5
2.0
2.5
3.0
3.5

## Percentage possibility of history type

## Predicted Evolution History

| | Mass_0 | Mass_1 | Semimajor | Eccentricity | logP |
|---|---|---|---|---|---|
| **0** | 22.343182 | 2.333571 | 2538.700000 | 2.439900e-01 | 5.224156 |
| **1** | 21.729472 | 2.333557 | 2633.953640 | 2.439899e-01 | 5.246389 |
| **2** | 21.725657 | 2.333557 | 2634.618130 | 2.439899e-01 | 5.246531 |
| **3** | 21.686325 | 2.333608 | 2345.928790 | 2.237674e-01 | 5.211639 |
| **4** | 17.379476 | 2.339919 | 1089.713271 | 8.820789e-02 | 4.422595 |
| **5** | 16.291458 | 2.339936 | 877.833962 | 8.819750e-02 | 4.206295 |
| **6** | 8.447135 | 2.339950 | 59.604493 | 0.000000e+00 | 3.094087 |
| **7** | 7.918371 | 2.339953 | 62.834400 | 4.882281e-02 | 3.139410 |
| **8** | 7.918571 | 2.337774 | 62.843623 | 4.882171e-02 | 3.139548 |
| **9** | 7.918572 | 2.337769 | 62.843637 | 4.882171e-02 | 3.139548 |
| **10** | 7.918663 | 2.337587 | 59.773876 | 0.000000e+00 | 3.106891 |
| **11** | 8.011951 | 1.814928 | 92.650760 | 0.000000e+00 | 3.396434 |

| Index | Source Id | BH Mass | MS Mass | Period |
|---|---|---|---|---|
| 15 | 1.84975e+18 | 12.92211 | 1.008930 | 0.05640179 |

| | Name | Mass_BH | Mass_1 | P | z | alpha | type | value |
|---|---|---|---|---|---|---|---|---|
| 0 | 0_941788181138103 | 8.131221 | 2.235984 | 0.059276 | 0.001 | 1.0 | CE | 0.573940 |
| 1 | 0_308481399850096 | 8.652071 | 2.331987 | 0.063940 | 0.001 | 1.0 | CE | 0.600773 |
| 2 | 0_251362050400014 | 7.595382 | 2.261225 | 0.057378 | 0.020 | 0.5 | CE | 0.601288 |

NI
MT
CE
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
P
6
8
10
12
14
Mass_BH
3.5
3.0
2.5
2.0
1.5
1.0
Mass_1

NI
MT
CE
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
P
14
12
10
8
6
Mass_BH
3.5
3.0
2.5
2.0
1.5
1.0
Mass_1

The graph of Mass_1 vs Mass_BH
Mass_1
Mass_BH
3.5
3.0
2.5
2.0
1.5
1.0
6
8
10
12
14

The graph of P vs Mass_BH
P
Mass_BH
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
6
8
10
12
14

The graph of P vs Mass_1
P
Mass_1
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
1.0
1.5
2.0
2.5
3.0
3.5

## Percentage possibility of history type

## Predicted Evolution History

| | Mass_0 | Mass_1 | Semimajor | Eccentricity | logP |
|---|---|---|---|---|---|
| **0** | 22.777832 | 2.340671 | 2843.500000 | 2.328800e-01 | 5.338222 |
| **1** | 21.979220 | 2.340658 | 2962.345380 | 2.328799e-01 | 5.366404 |
| **2** | 21.973112 | 2.340658 | 2963.307780 | 2.328799e-01 | 5.366625 |
| **3** | 21.920654 | 2.340729 | 2677.593580 | 2.151050e-01 | 5.333032 |
| **4** | 16.240294 | 2.349304 | 1075.302656 | 8.352793e-02 | 4.331777 |
| **5** | 15.142083 | 2.349322 | 863.492979 | 8.351754e-02 | 4.116152 |
| **6** | 8.614823 | 2.349335 | 65.565025 | 0.000000e+00 | 3.152544 |
| **7** | 8.085015 | 2.349338 | 69.056809 | 4.798366e-02 | 3.197147 |
| **8** | 8.085213 | 2.347084 | 69.067184 | 4.798240e-02 | 3.197287 |
| **9** | 8.085214 | 2.347079 | 69.067200 | 4.798239e-02 | 3.197287 |
| **10** | 8.085324 | 2.346866 | 65.837928 | 0.000000e+00 | 3.165996 |
| **11** | 8.166855 | 1.900922 | 94.922475 | 0.000000e+00 | 3.406405 |

# The percentage possibility of type for Real candidates : Gold

Why in almost all cases, the Mass_1 of the real candidate is significantly lower than the masses of similar systems in the SEVN simulation ?

# Similar systems
## Limitations

- **Initial Stellar Mass:** The model imposes a lower limit of **2.2 solar masses** for the initial stellar mass in binary systems

- **Individual System Variability**: Each binary system has its own unique characteristics and evolutionary path
- **External and Missing Factors**

- **Model Assumptions**: The methodology relies on the assumption that the selected similar systems from the SEVN simulation accurately represent the range of possible evolutionary paths for the real candidate system.
- **Unique Nature of Real Candidates**
- **Evolution History**: The lack of complete evolution history for real binary system

# Similar systems
## Future improvements

- **Enhanced Simulation Models →** better capture the complexities and variations of BSE
- **Use more parameters**