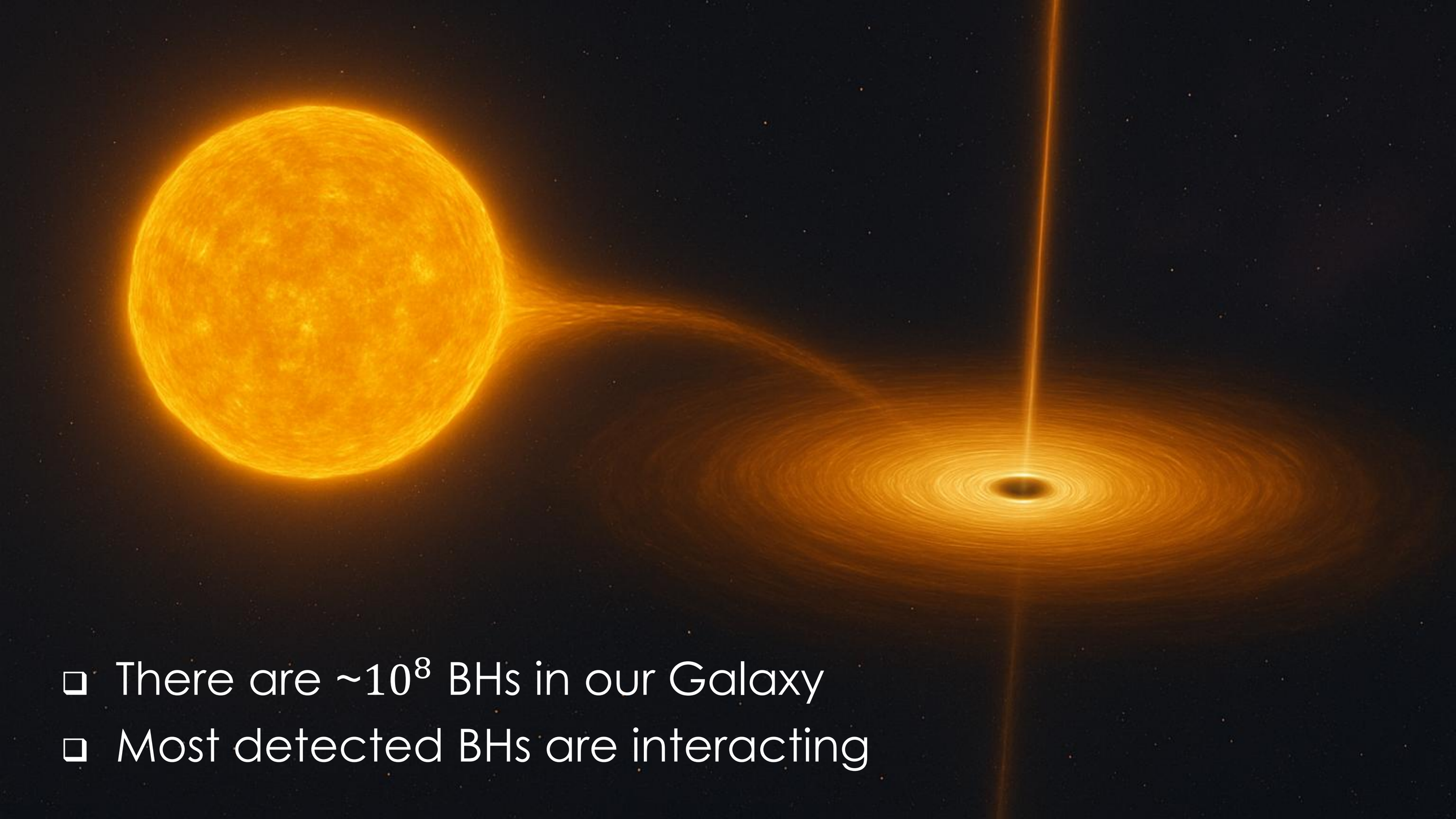
Toward the detection of massive dark companions in photometric surveys using GPR and PCA

Milan Pešta

CCAPP Fellows Symposium
Sep 18, 2025

FUNDED BY
GAČR
CZECH SCIENCE FOUNDATION



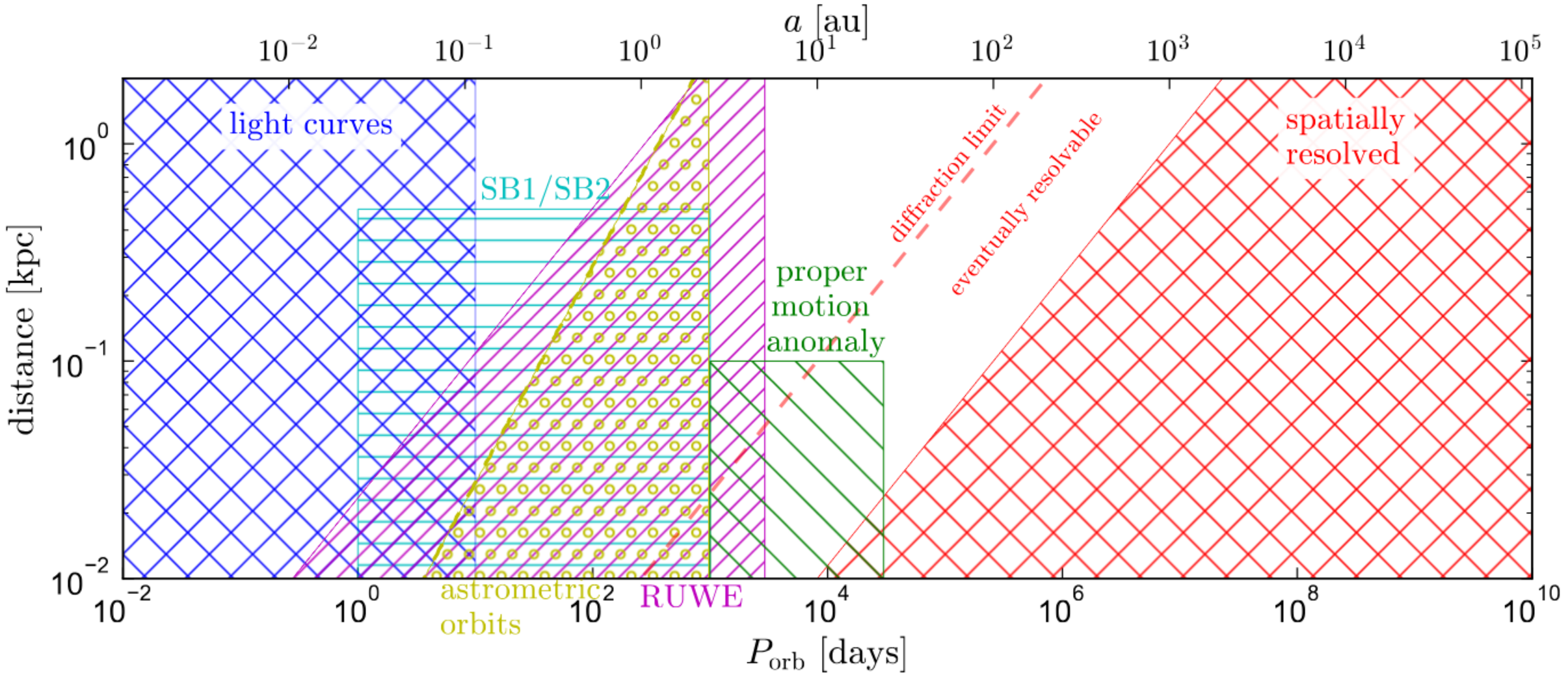
- There are ~10^8 BHs in our Galaxy
- Most detected BHs are interacting

Large population dark companion binaries

- There are ~10^8 BHs in our Galaxy
- Most detected BHs are interacting

Methods of detection

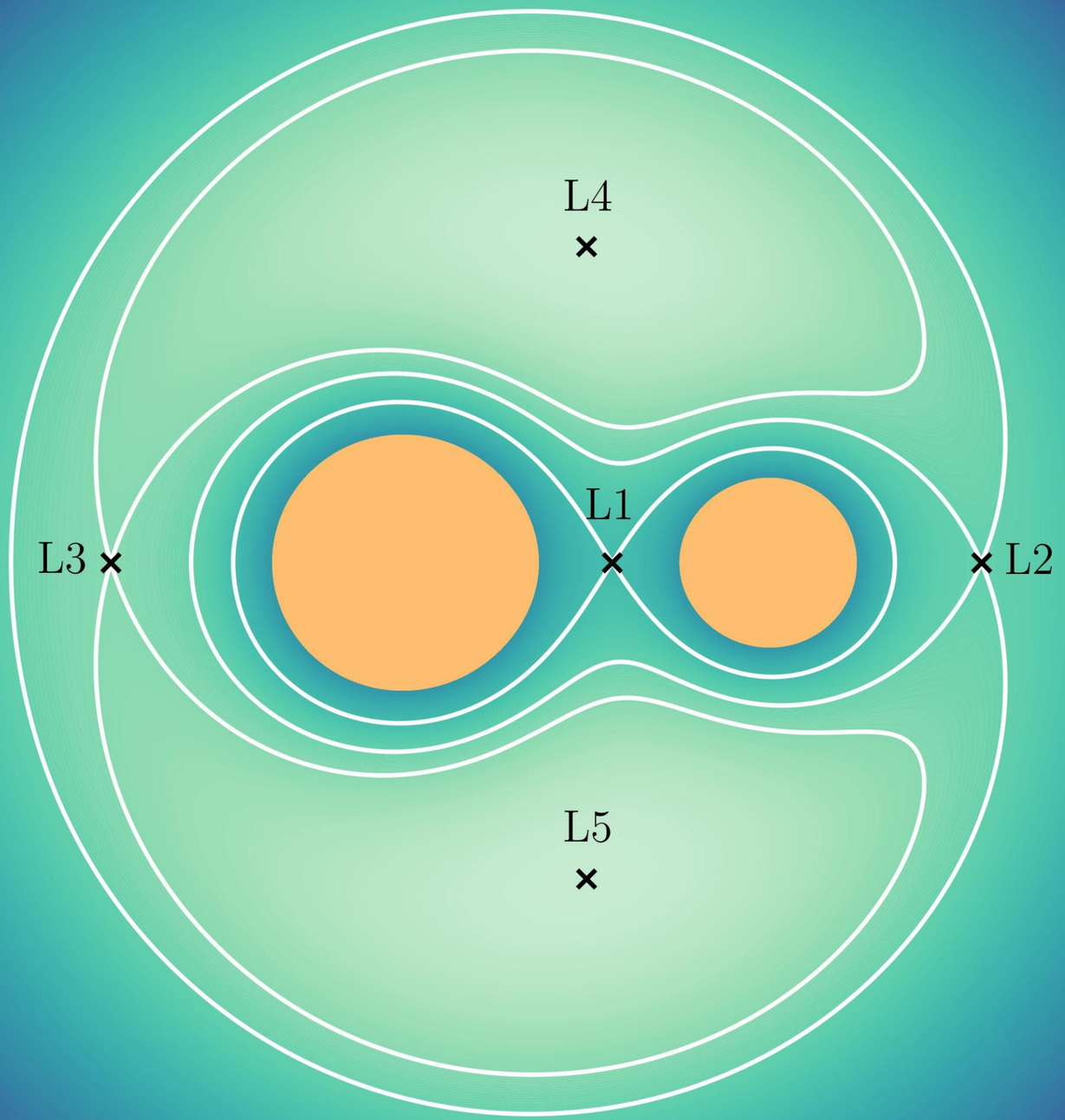
- High radial accelerations and binary mass functions (Thompson et al. 2019)
- Large wobbles in astrometric binary solutions (El-Badry et al. 2023, El-Badry et al. 2024)


a [au]
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
distance [kpc]
10^{0}
10^{-1}
10^{-2}
light curves
SB1/SB2
astrometric orbits
RUWE
proper motion anomaly
diffraction limit
eventually resolvable
spatially resolved
10^{-2}
10^{0}
10^{2}
10^{4}
10^{6}
10^{8}
10^{10}
$P_{\rm orb}$ [days]

El-Badry24

Methods of detection

- High radial accelerations and binary mass functions (Thompson et al. 2019)
- Large wobbles in astrometric binary solutions (El-Badry et al. 2023, El-Badry et al. 2024)
- Ellipsoidal variations (Gomel et al. 2021, Gomel et al. 2022)

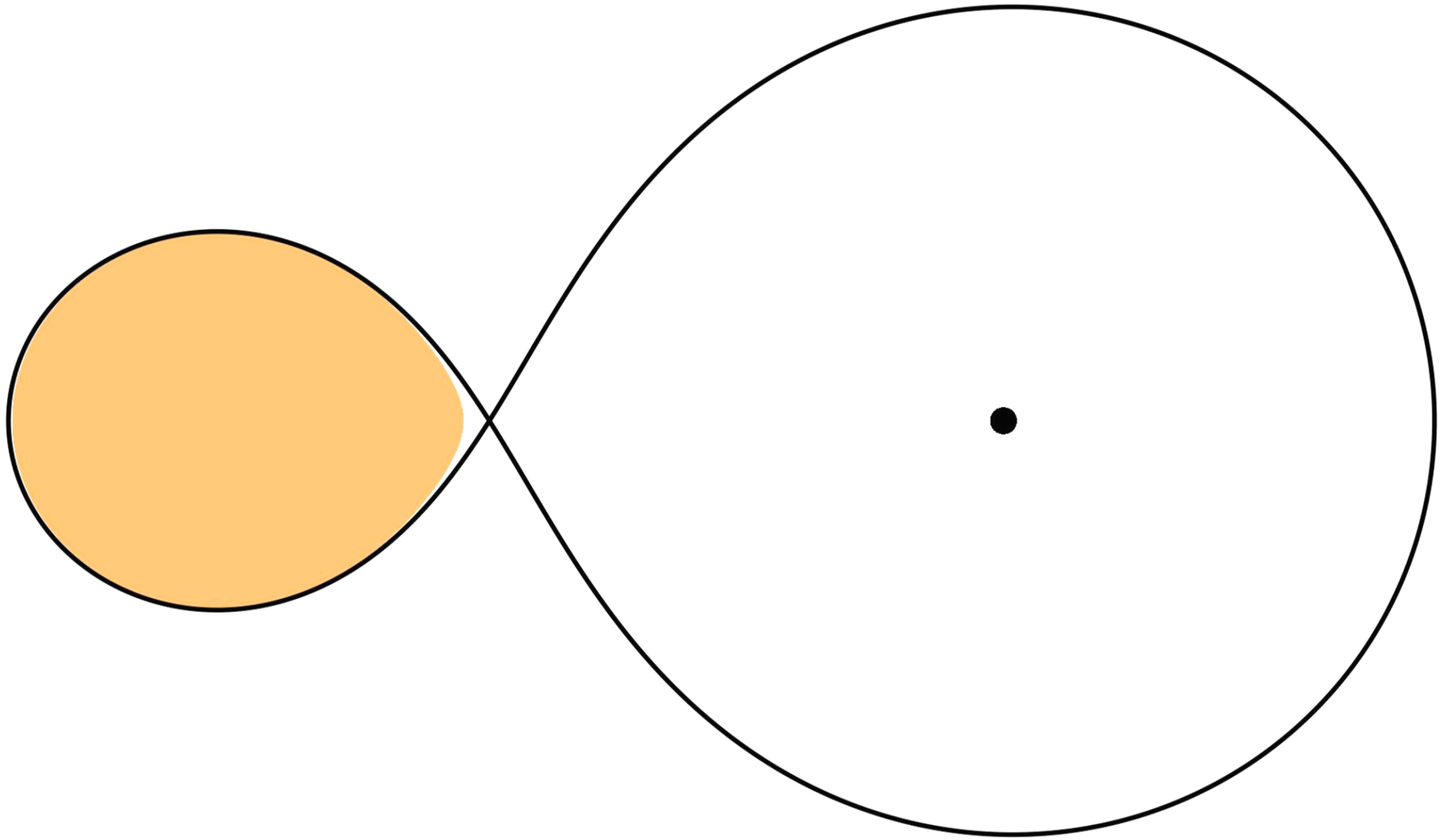


Tidally distorted stars

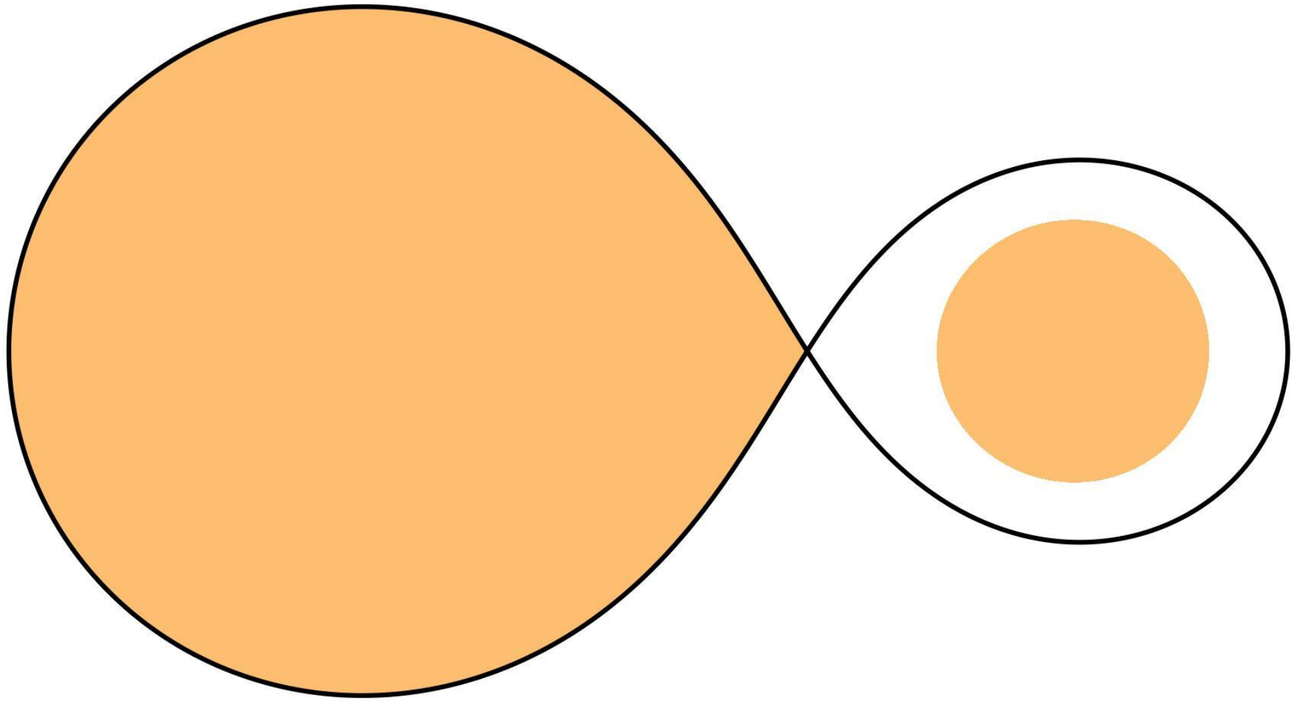
Visible surface area changes in time

Ellipsoidal variations

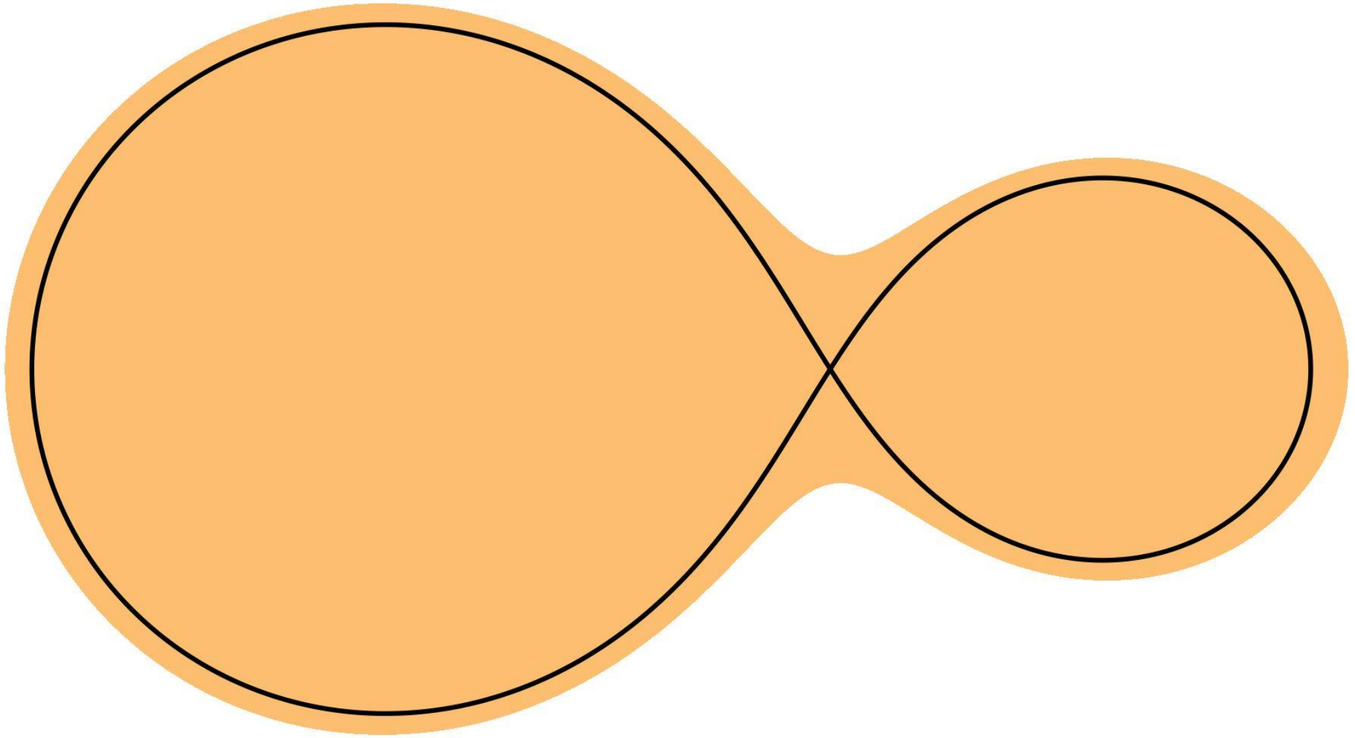
Star + black hole/neutron star

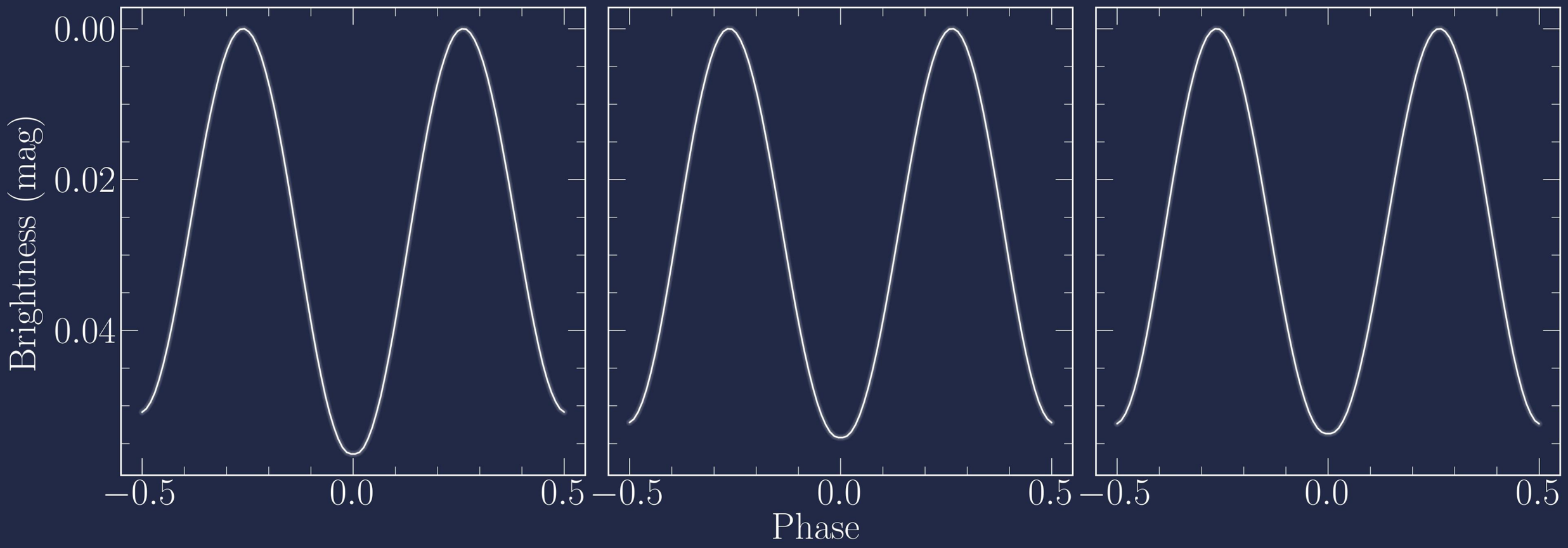


Semidetached configuration



Contact configuration

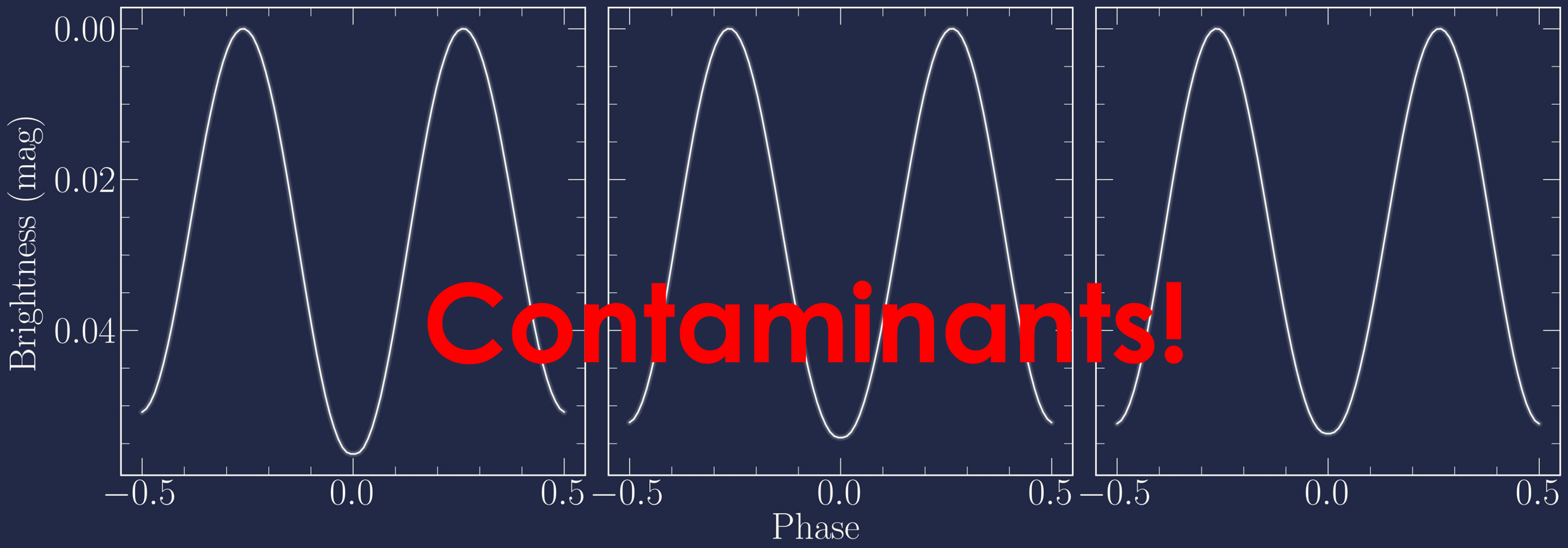




Contact binary
$q = 0.5$
$i = 25°$

Star + BH/NS
$q = 2$
$i = 45°$

Semidetached binary
$q = 0.5$
$i = 35°$



Contaminants!

Contact binary
$q = 0.5$
$i = 25°$

Star + BH/NS
$q = 2$
$i = 45°$

Semidetached binary
$q = 0.5$
$i = 35°$

How do we get rid of contaminants?

How do we get rid of contaminants?

- Machine learning

How do we get rid of contaminants?

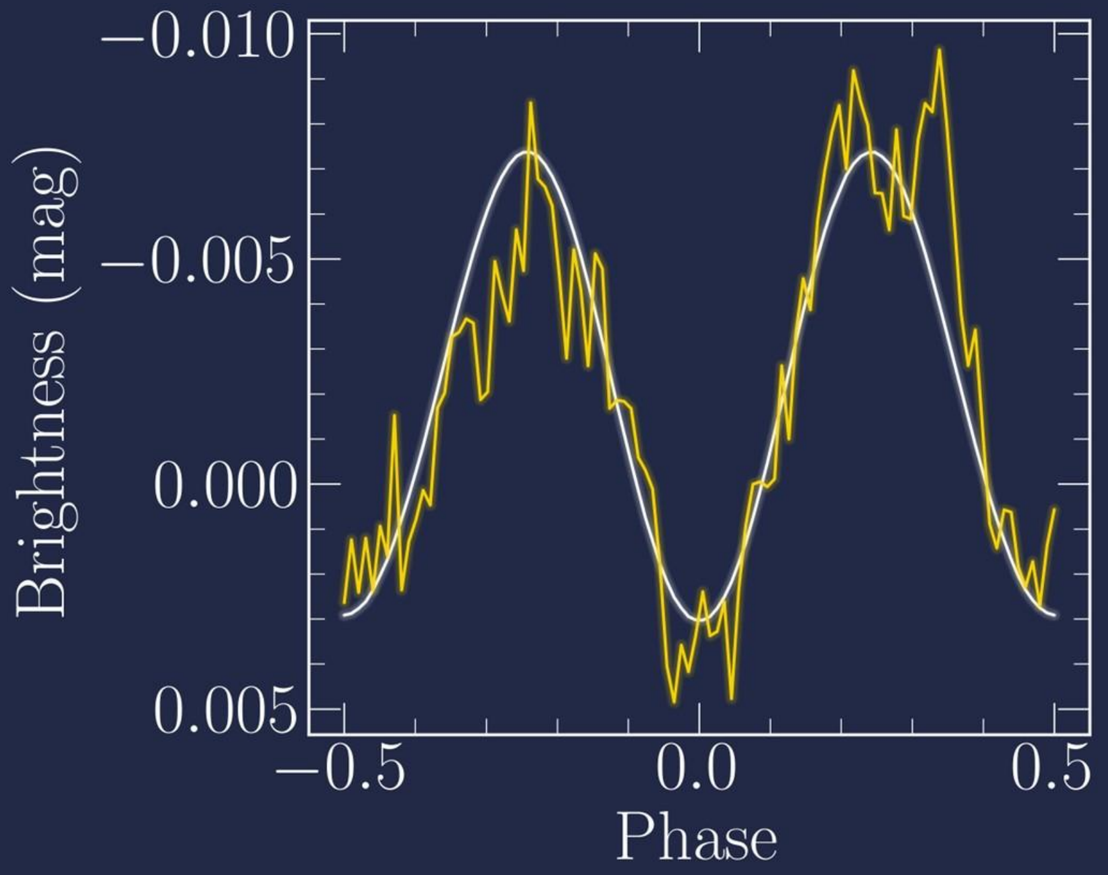
- Machine learning
- Problem: no training data

How do we get rid of contaminants?

- Machine learning
- Problem: no training data
- Solution: synthetic data

Synthetic light curves

- PHOEBE (Prša et al. 2016, Conroy et al. 2020)
- Dark companion binaries, contact binaries, semidetached binaries
- Wide range of physical and orbital parameters
- Correlated and uncorrelated Gaussian noise

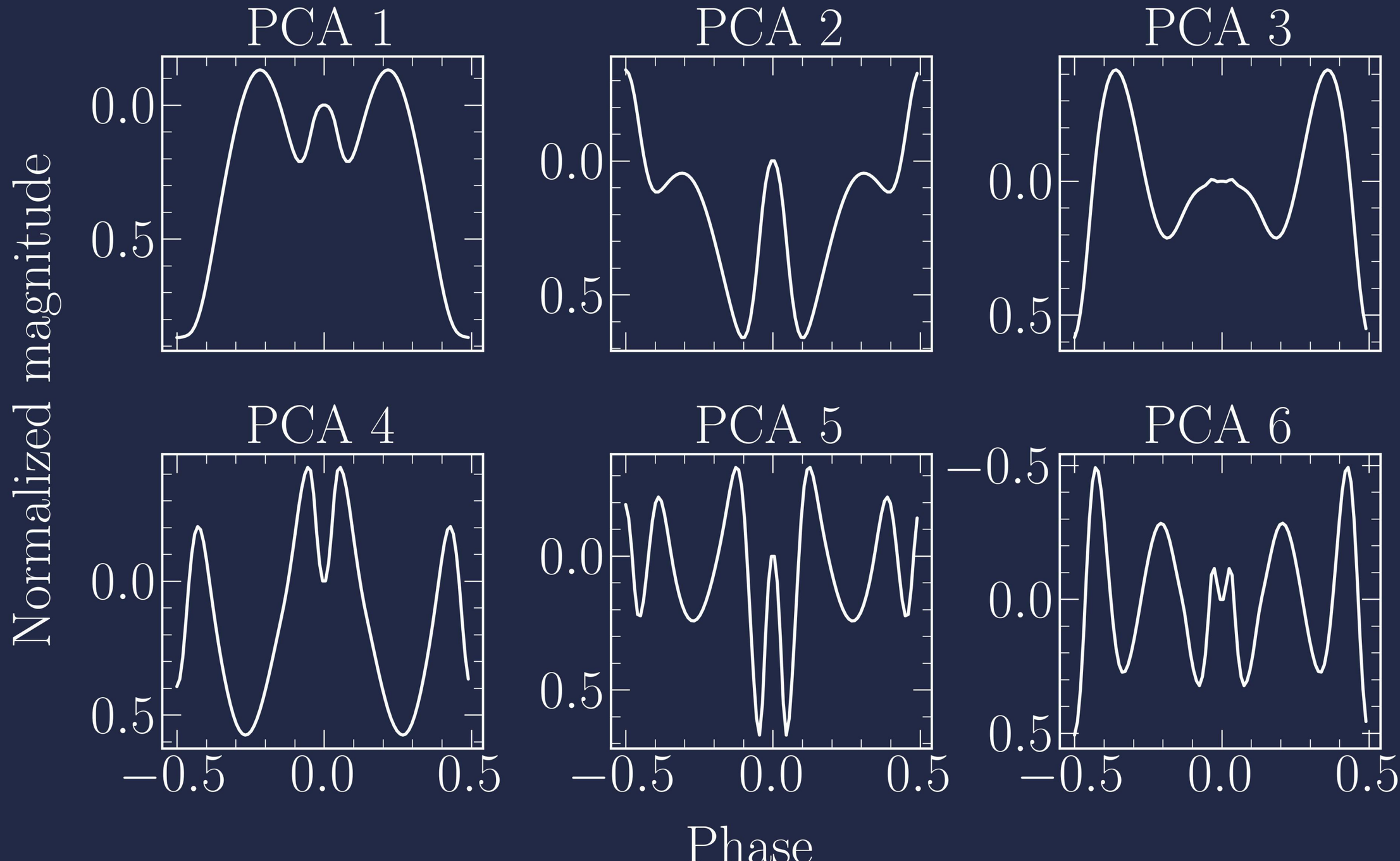


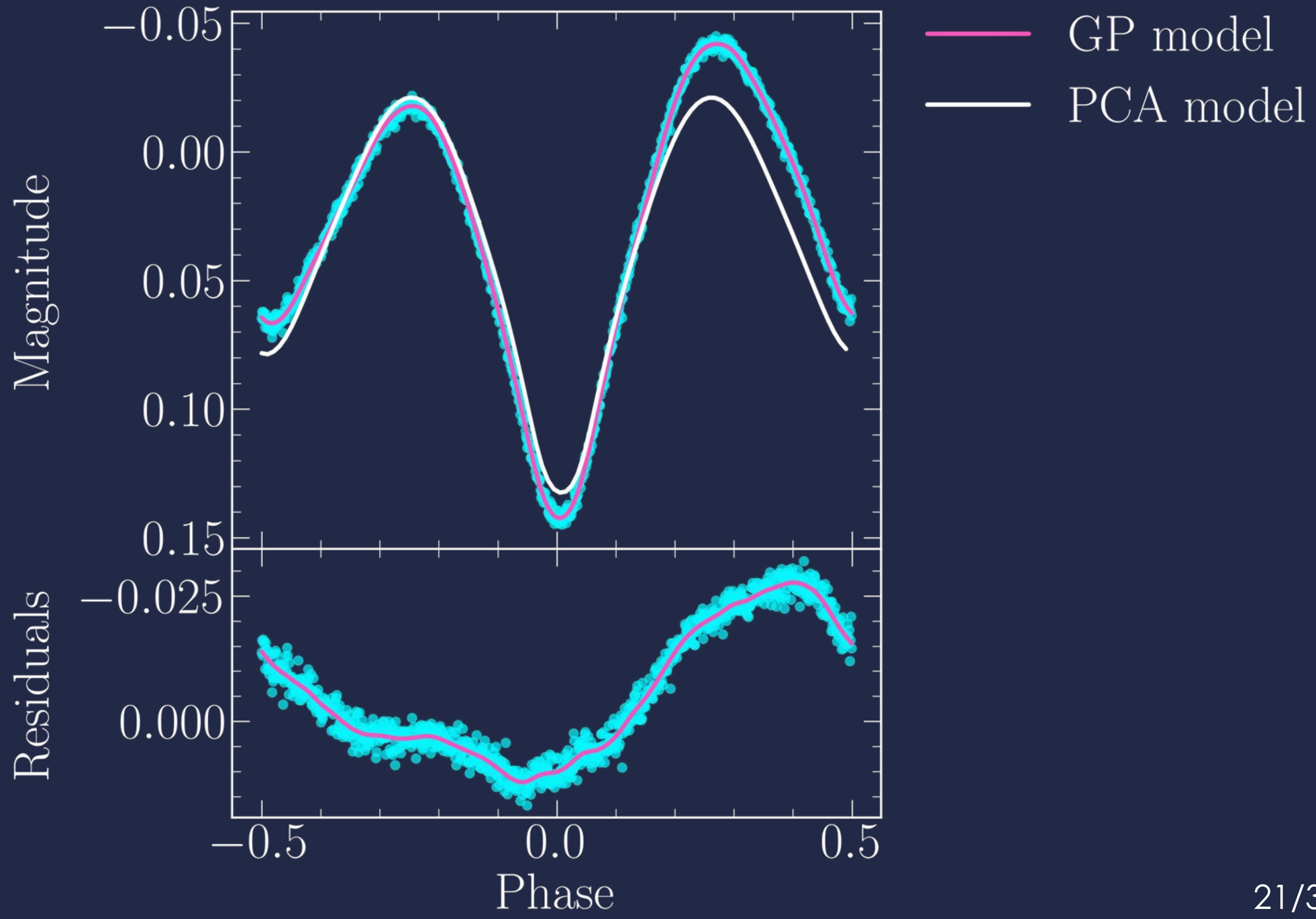
Dimensionality reduction

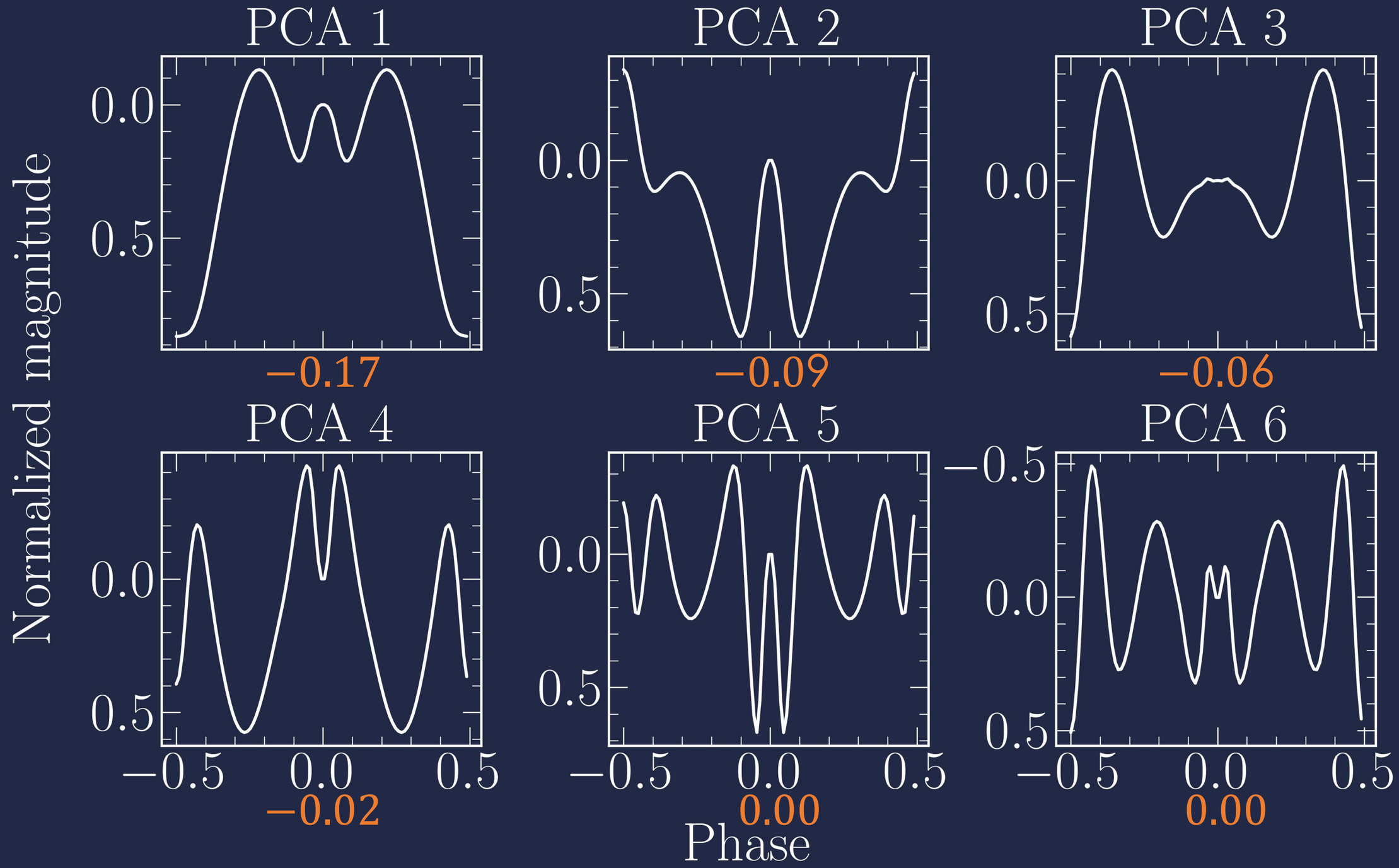
- Easier to visualize and interpret
- Can reveal underlying structure and patterns
- Effective at mitigating the curse of dimensionality
- Naturally applicable to time series

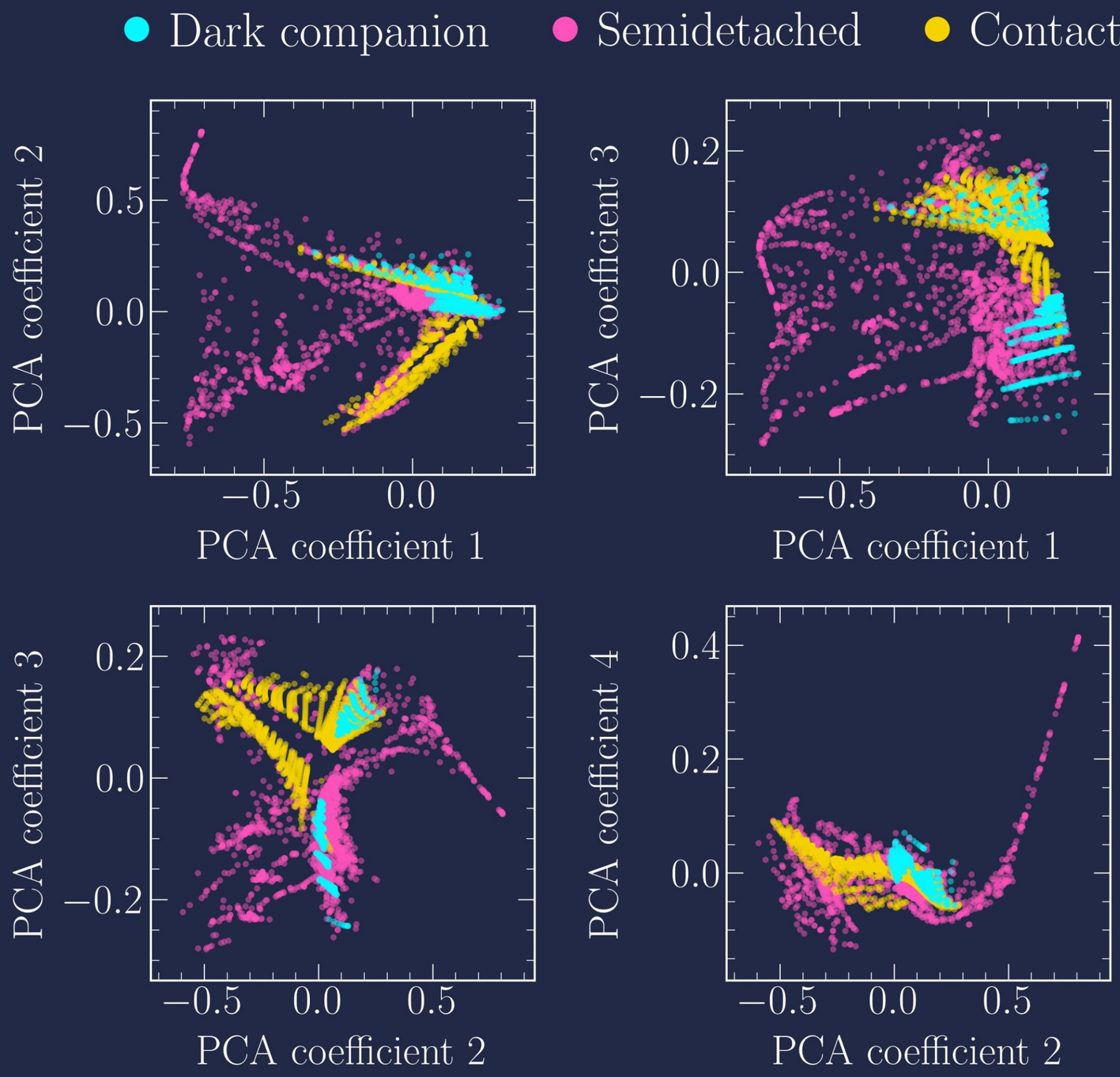
Gaussian processes + principal component analysis

- PCA trained on synthetic light curves
- Mean function = linear combination of PCA components
- PCA: ellipsoidal signal
- GP: noise + non-ellipsoidal signal (including spots)
- Variable #components → Bayesian information criterion (BIC)


PCA 1
PCA 2
PCA 3
PCA 4
PCA 5
PCA 6
Normalized magnitude
Phase
0.0
0.5
−0.5
0.0
0.5

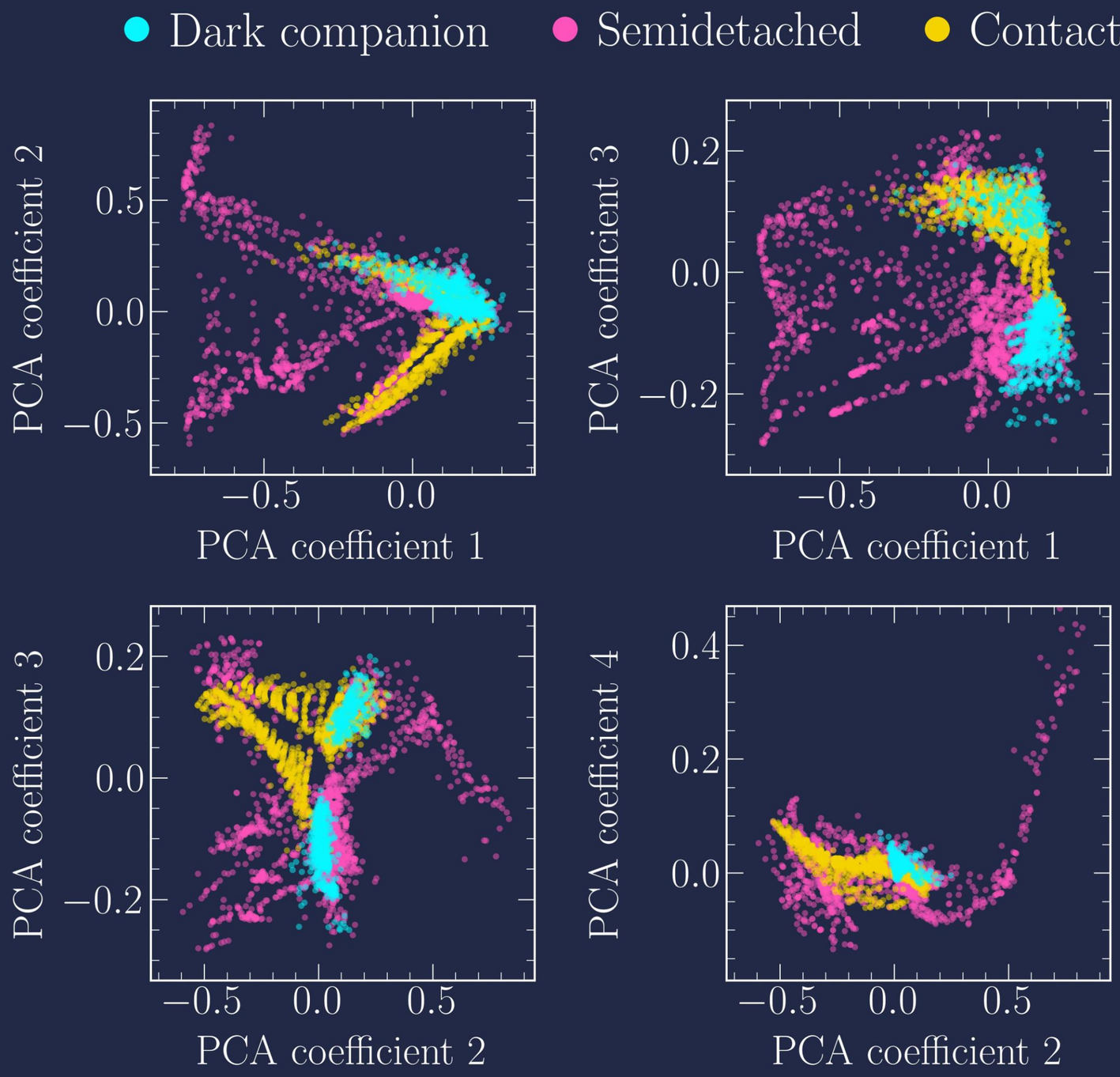

GP model
PCA model
Magnitude
−0.05
0.00
0.05
0.10
0.15
Residuals
−0.025
0.000
−0.5
0.0
0.5
Phase


PCA 1
PCA 2
PCA 3
PCA 4
PCA 5
PCA 6
Normalized magnitude
0.0
0.5
−0.17
−0.09
−0.06
−0.02
0.00
0.00
−0.5
0.0
0.5
Phase

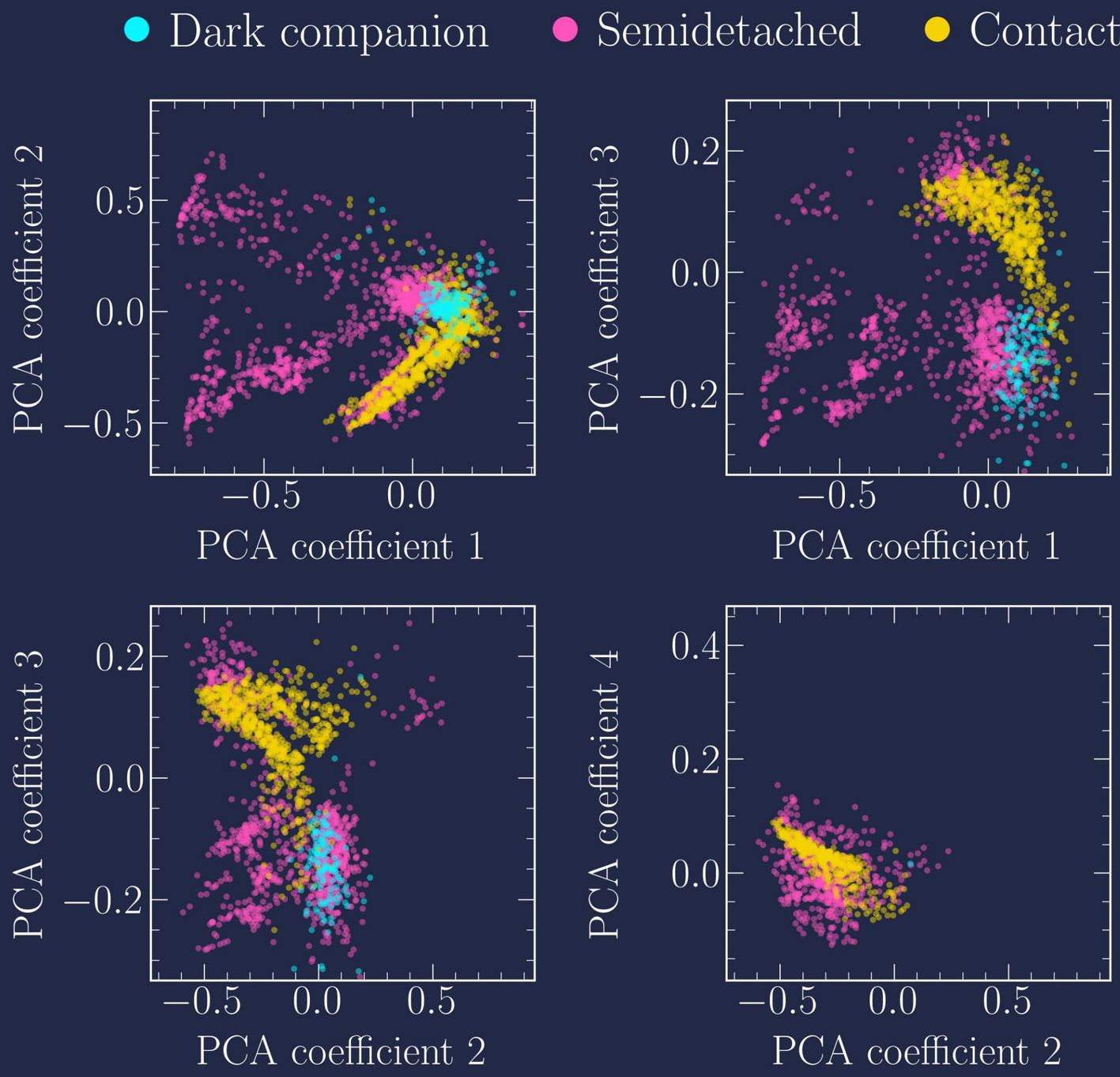


$\sigma = 0.000$ mag

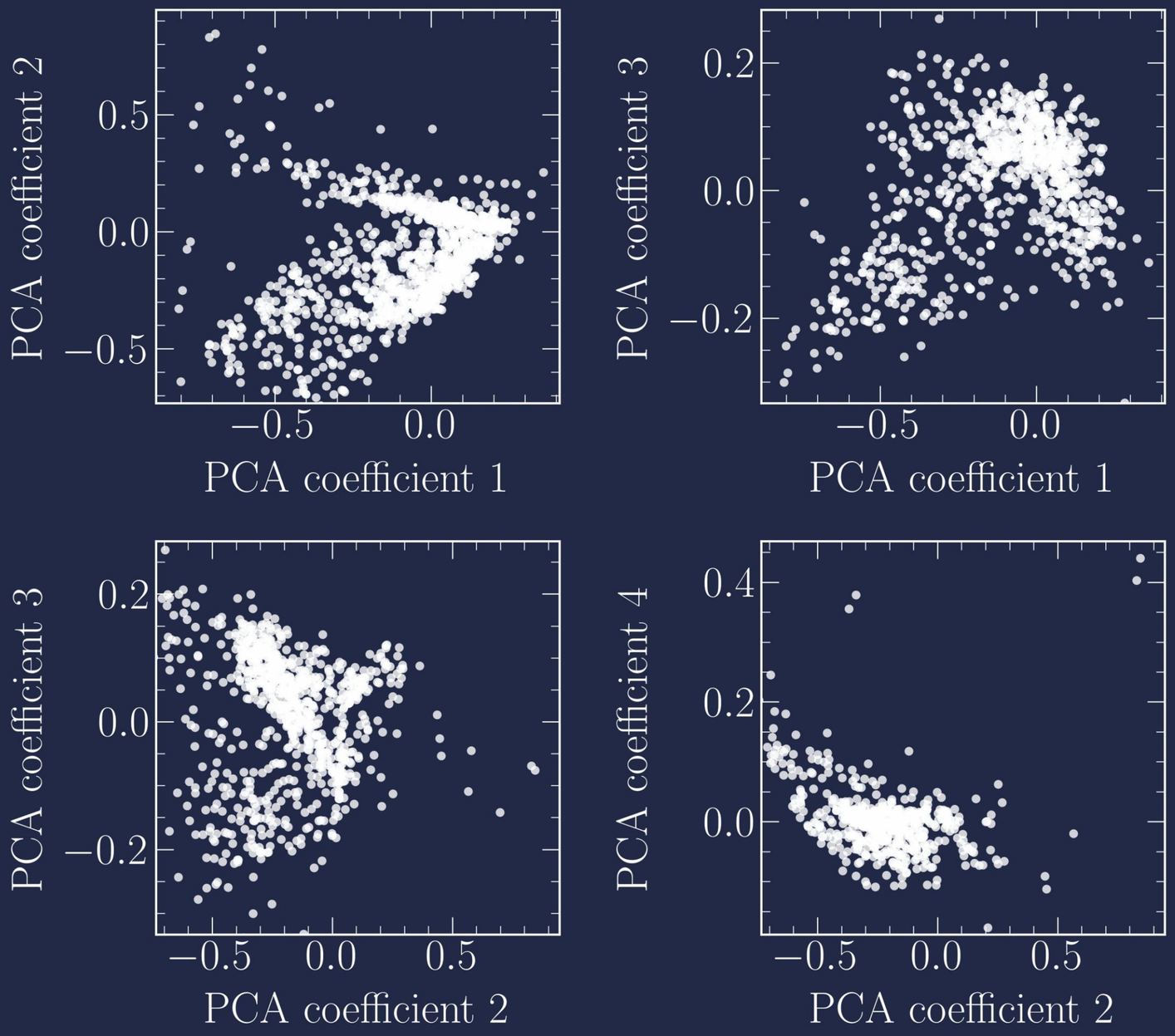
$\sigma = 0.001$ mag



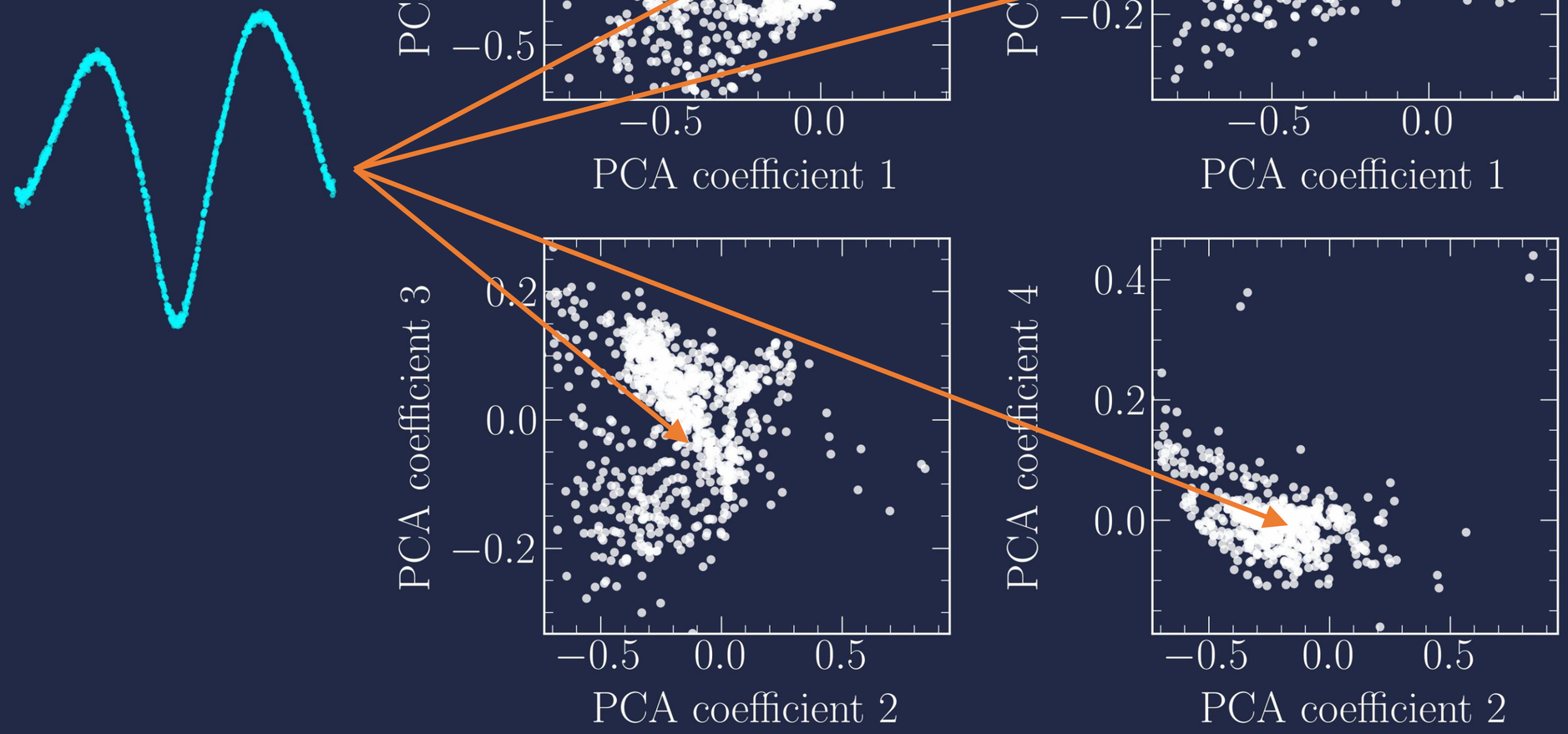
$\sigma = 0.010$ mag



- Kepler Eclipsing Binary Catalog



- Kepler Eclipsing Binary Catalog



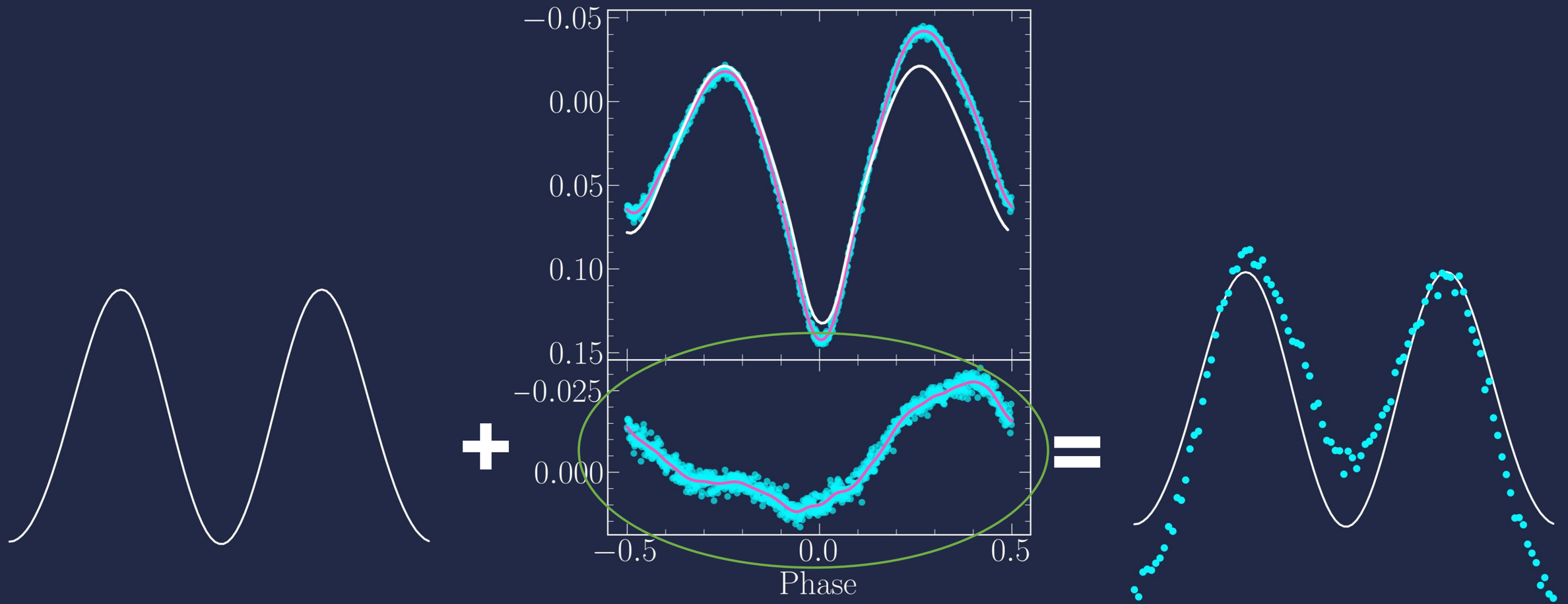
How do we train a classifier?

How do we train a classifier?

- Problem: synthetic gap

How do we train a classifier?

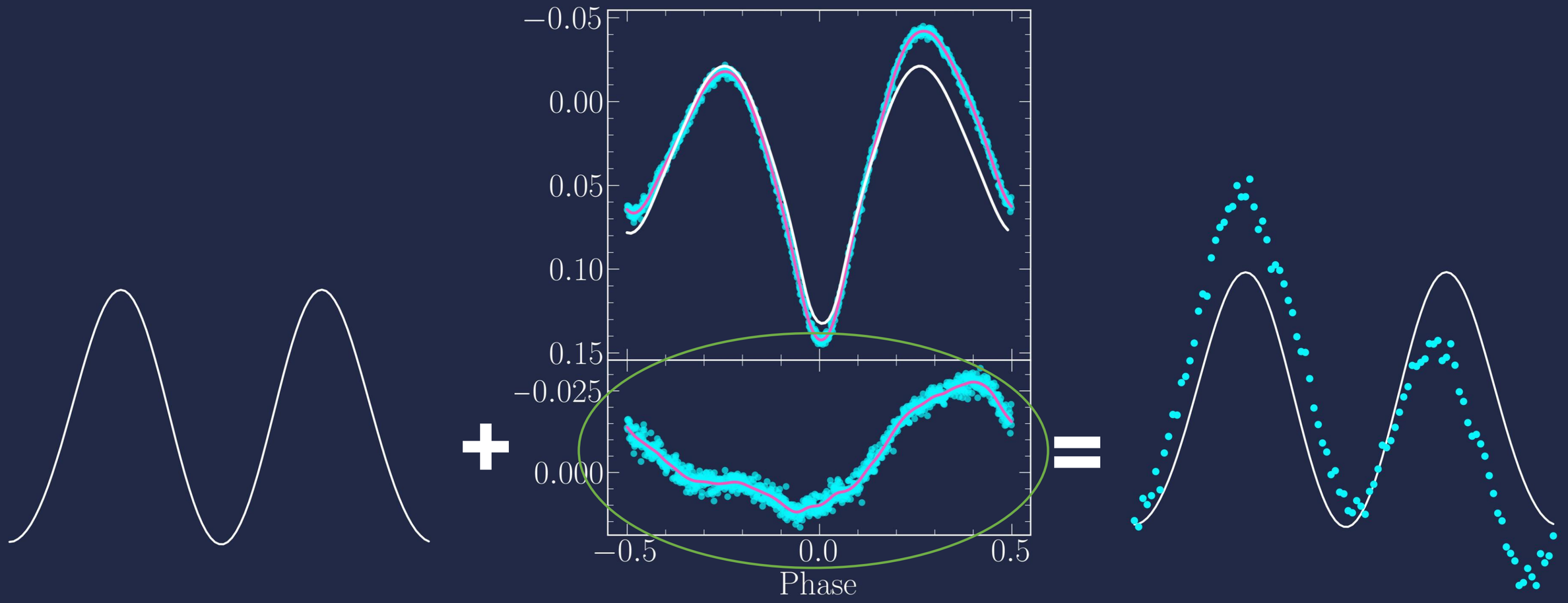
- Problem: synthetic gap
- Solution: noise transfer



Noiseless
synthetic light curve

Residuals

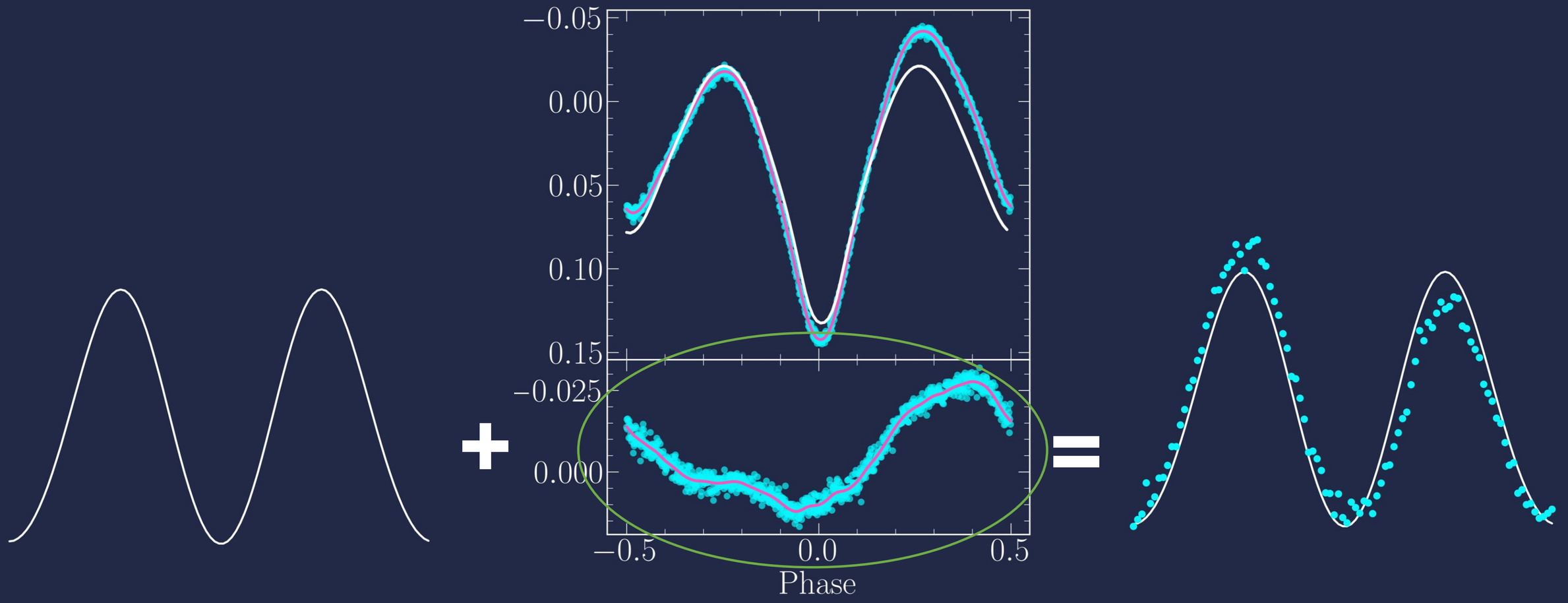
Synthetic light curve
with realistic noise



Noiseless
synthetic light curve

Residuals

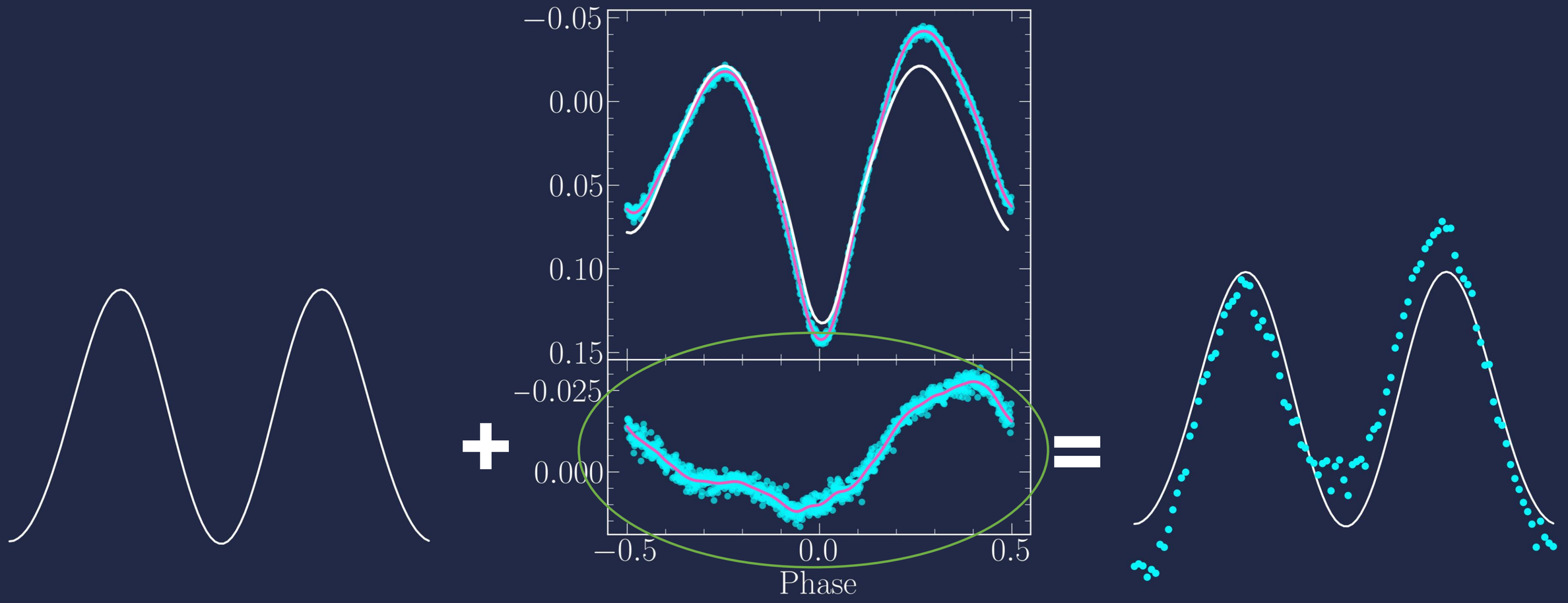
Synthetic light curve
with realistic noise



Noiseless
synthetic light curve

Residuals

Synthetic light curve
with realistic noise

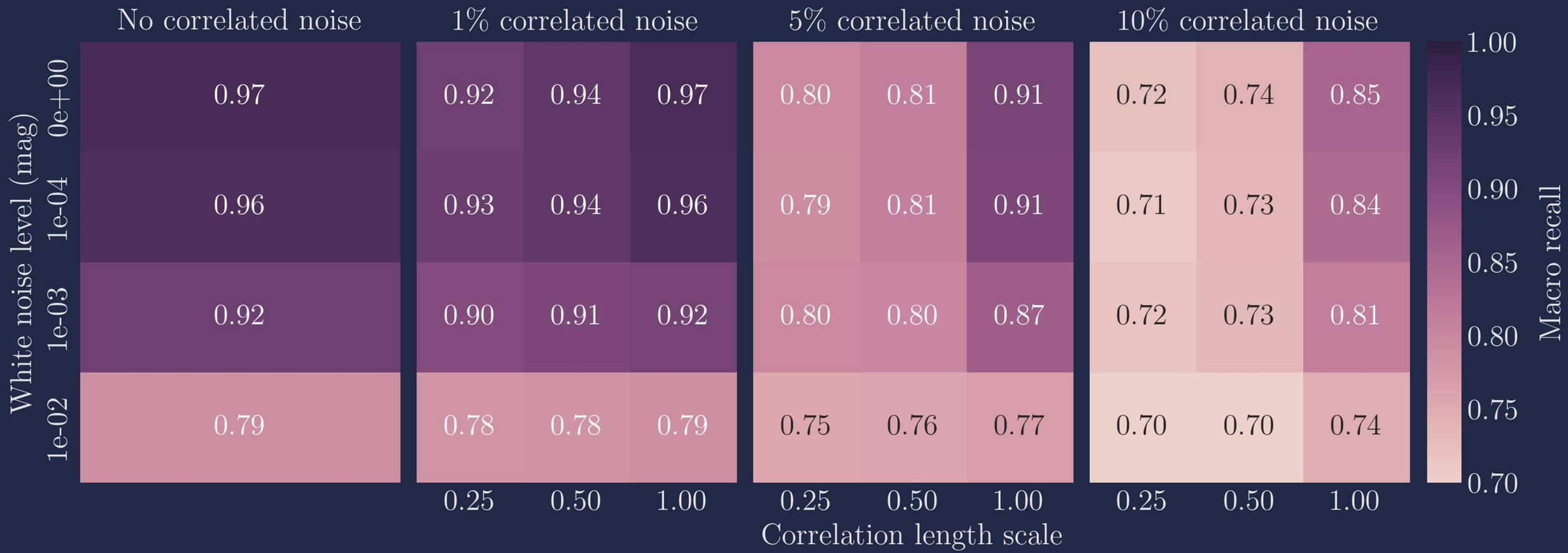


Noiseless
synthetic light curve

Residuals

Synthetic light curve
with realistic noise

Completeness



Pešta+25

Conclusions

- Novel method for detection of dark companion binaries using principal component analysis, and Gaussian processes
- Synthetic gap ↔ noise transfer
- Applicable to any sample of ellipsoidal variables

Backup slides

Detached configuration

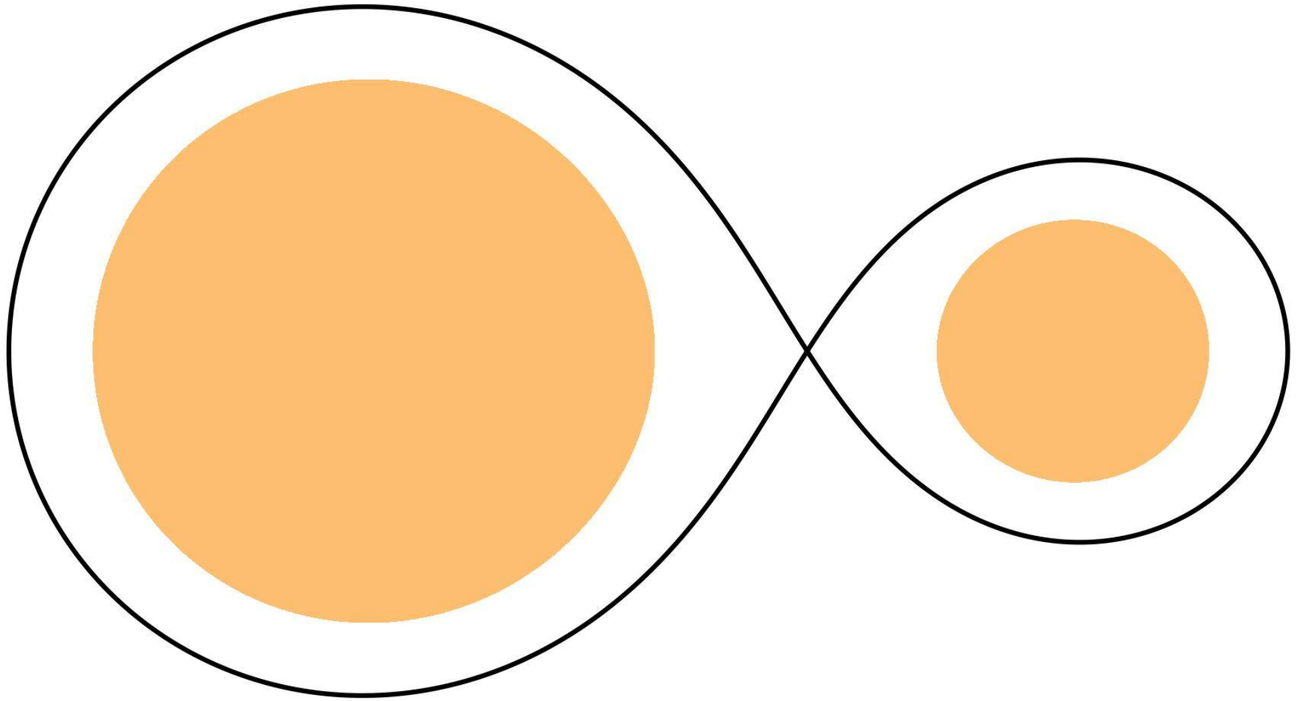
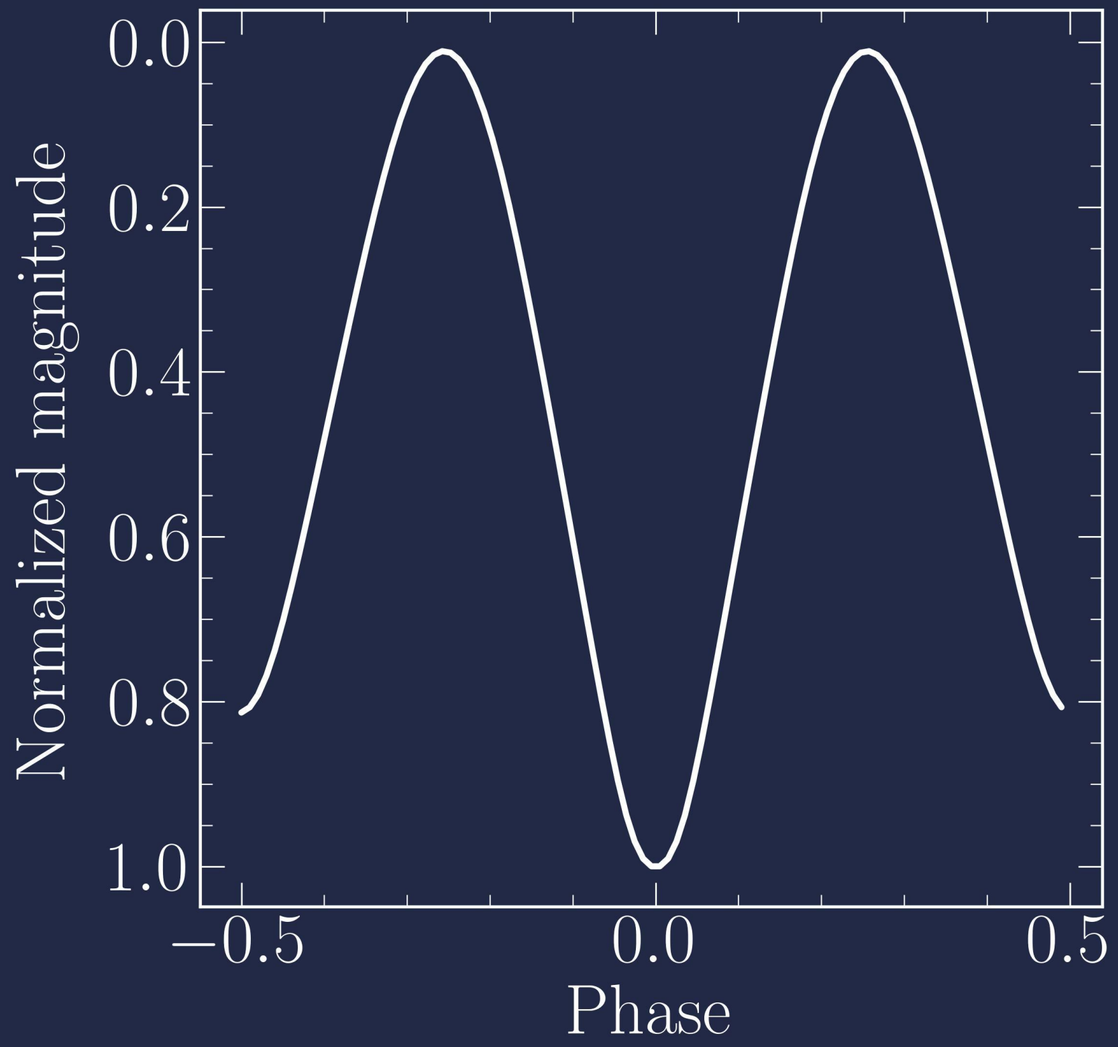


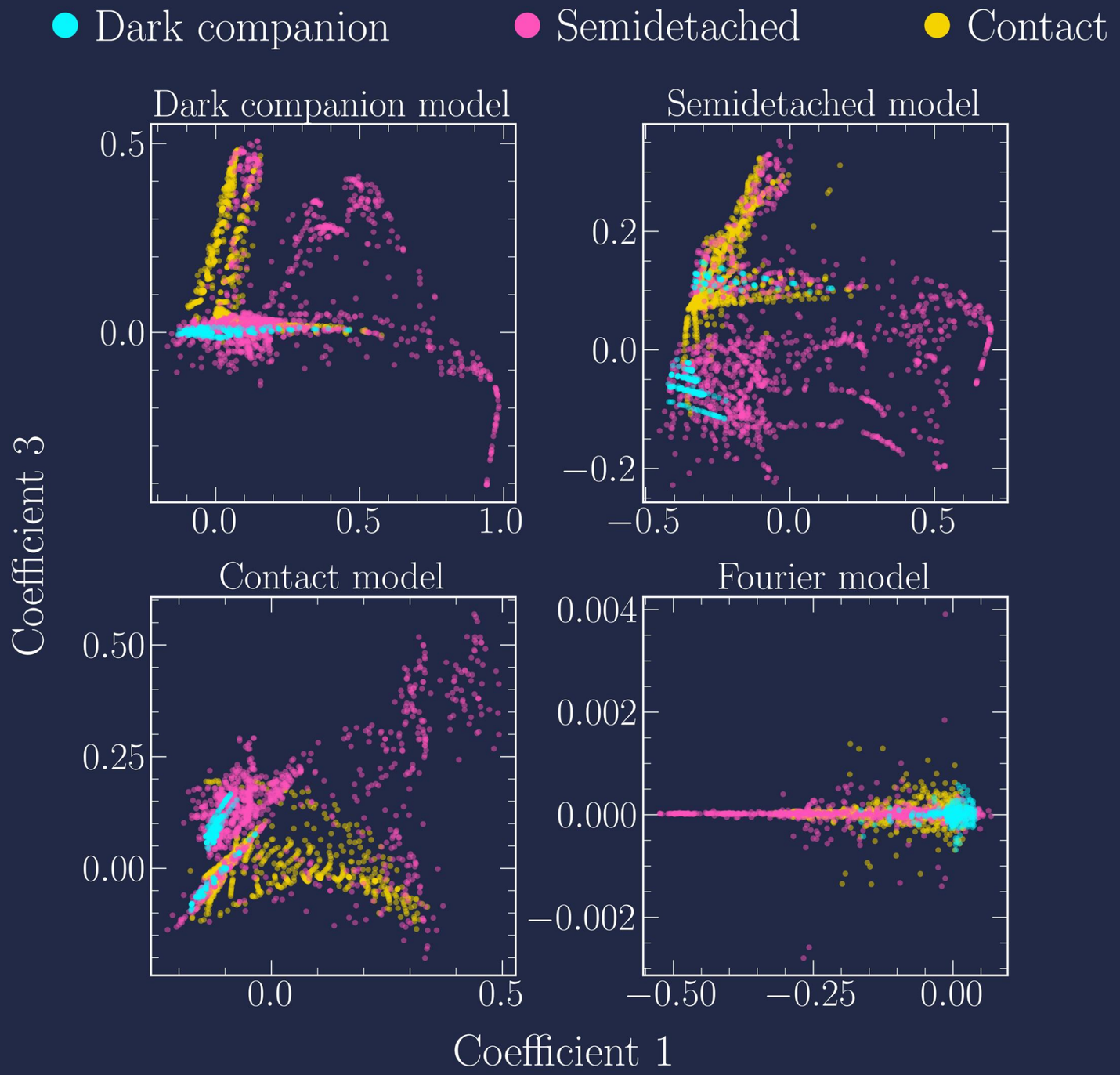
Table 1. Ranges of physical and orbital parameters used to generate synthetic light curves of dark companion, semidetached, and contact binary systems.

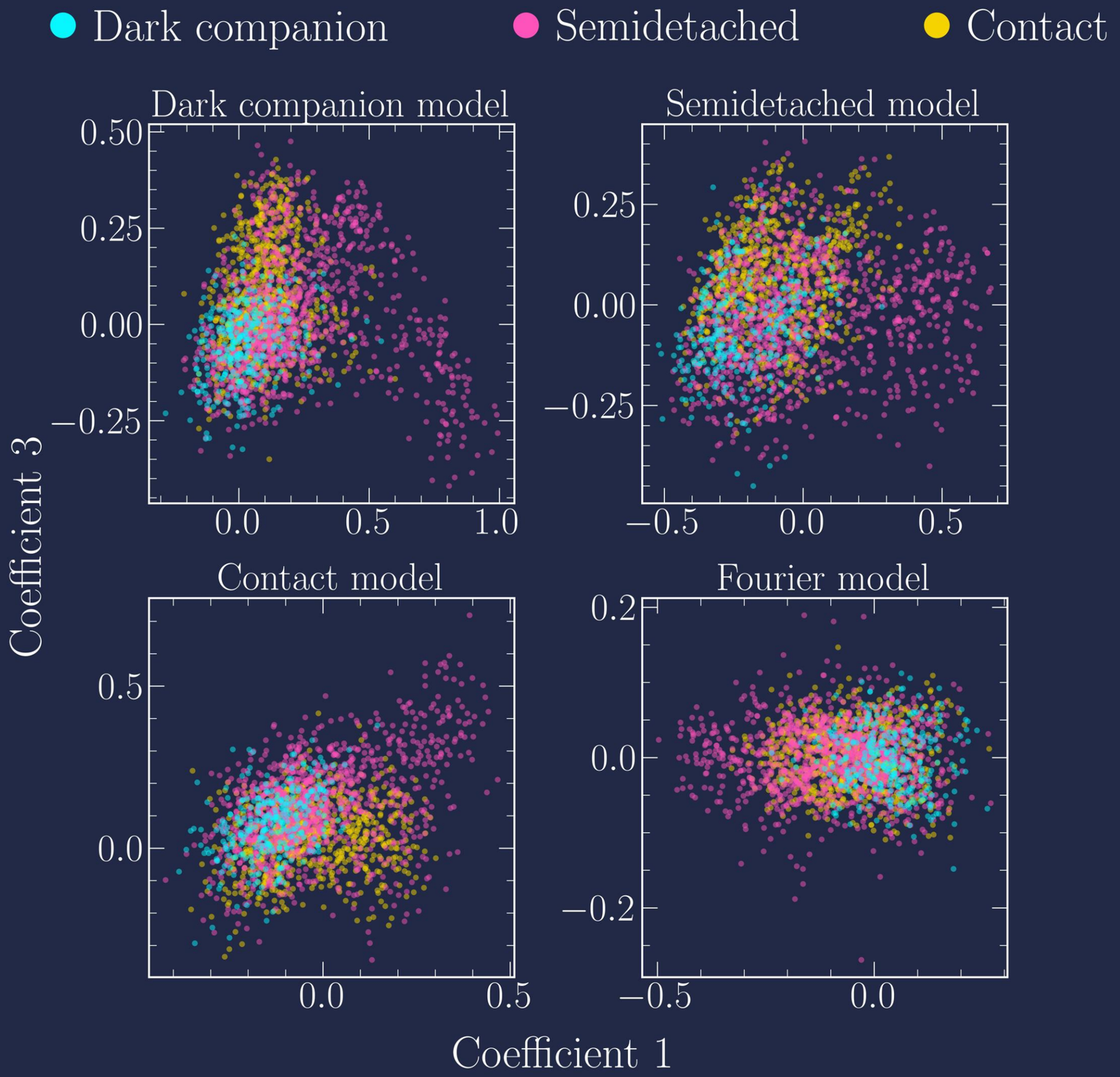
Parameter	Dark companion binaries	Semidetached binaries	Contact binaries
M_1 ($M_\odot$)	1	1	1
q	0.05–1 ($\Delta q = 0.05$) + 1–10 ($\Delta q = 1$)	0.05–1 ($\Delta q = 0.05$) + 1–10 ($\Delta q = 1$)	0.05–1 ($\Delta q = 0.05$)
$T_{\rm eff,1}$ (K)	6000	6000, 8000, 15 000	6000
$T_{\rm eff,2}/T_{\rm eff,1}$	–	0.5, 1, 2	1
R_1 ($R_\odot$)	1	–	–
R_2/R_1	–	0.1, 0.5, 2, 5	–
i (°)	5–90 ($\Delta i = 5°$)	5–90 ($\Delta i = 5°$)	5–90 ($\Delta i = 5°$)
$\log a/R_\odot$	0–1 ($\Delta \log a/R_\odot = 0.11$)	0–1 ($\Delta \log a/R_\odot = 0.11$)	–
f	–	–	0.15, 0.25, 0.5, 0.75

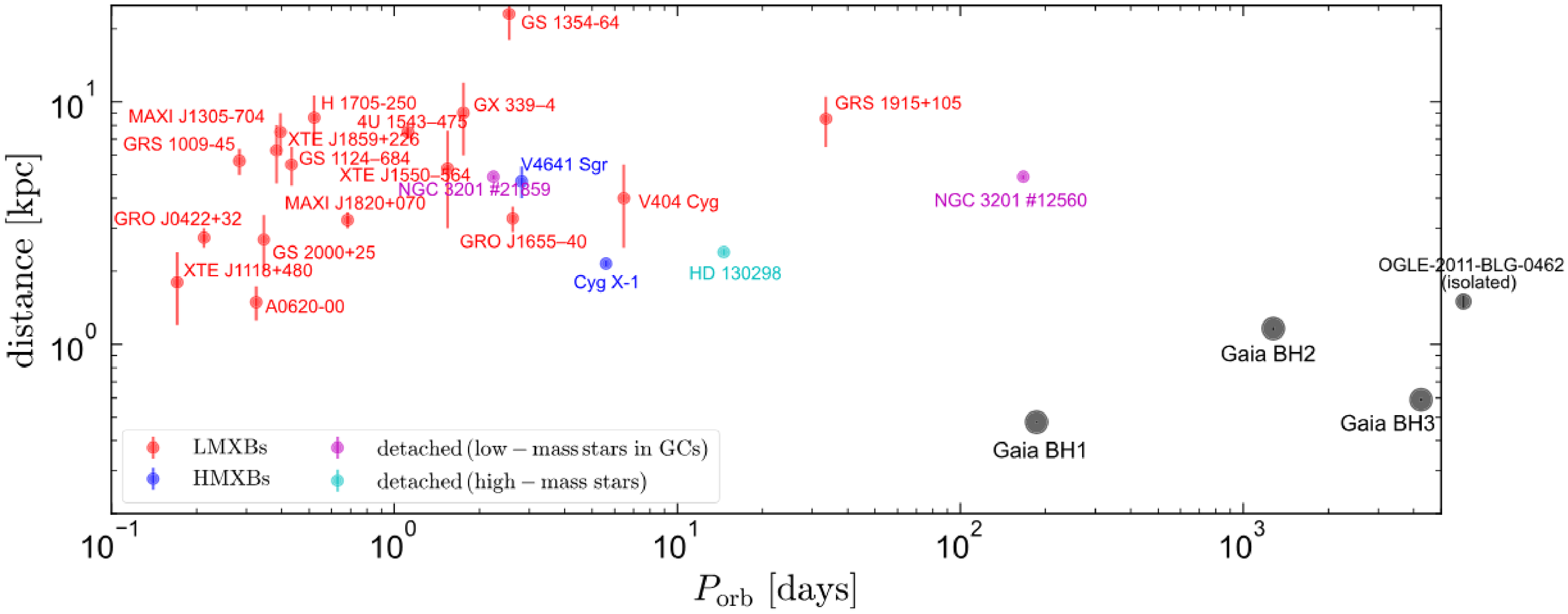
Notes. Parameters denoted with subscripts 1 and 2 refer to the primary and secondary components, respectively.

Mean light curve




Dark companion
Semidetached
Contact
Dark companion model
Semidetached model
Contact model
Fourier model
Coefficient 3
Coefficient 1
0.5
0.0
1.0
0.2
−0.2
−0.5
0.50
0.25
0.00
0.004
0.002
0.000
−0.002
−0.50
−0.25


Dark companion
Semidetached
Contact
Dark companion model
Semidetached model
Contact model
Fourier model
Coefficient 3
Coefficient 1
0.50
0.25
0.00
−0.25
0.0
0.5
1.0
−0.5
0.2
−0.2


GS 1354-64
H 1705-250
GX 339–4
MAXI J1305-704
4U 1543–475
GRS 1915+105
XTE J1859+226
GRS 1009-45
GS 1124–684
V4641 Sgr
XTE J1550–564
NGC 3201 #21359
NGC 3201 #12560
V404 Cyg
MAXI J1820+070
GRO J0422+32
GRO J1655–40
GS 2000+25
XTE J1118+480
Cyg X-1
HD 130298
OGLE-2011-BLG-0462
(isolated)
A0620-00
Gaia BH2
Gaia BH3
Gaia BH1
LMXBs
HMXBs
detached (low – mass stars in GCs)
detached (high – mass stars)
distance [kpc]
$P_{\rm orb}$ [days]

El-Badry24