The Mass Distribution of Stellar-Mass Black Holes

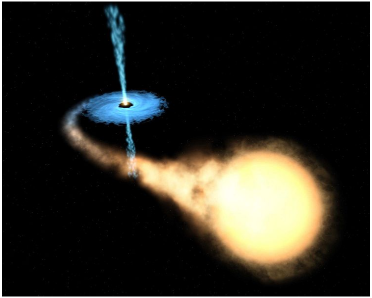
Robust Constraints Through Model Selection
arXiv:1011.1459

Will M. Farr Niharika Sravan Andrew Cantrell Laura Kreidberg Charles Bailyn Ilya Mandel Vicky Kalogera

GWPAW
January 29, 2011

Stellar Mass Black Holes

Black holes come in many varieties; we're interested in the smallest ones. Our sample comes from identified X-Ray Binaries:



Previous Work

- Bailyn, *et al*, 1998
  - Seven systems.
  - Black hole distribution flat between 5 and 10 $M_\odot$.
  - Mass gap between low-mass BH and high-mass NS.
- Ozel, *et al*, 2010
  - 16 systems.
  - Narrow (Gaussian) distribution with peak at 8 $M_\odot$, width of few.
  - Also mass gap.
  - Also considered theoretically favored models (Freyer and Kalogera 2001).
  - No model comparison.

Black Hole Masses

We want (accurate) estimates of the mass, so we only use binaries where we know the mass function:

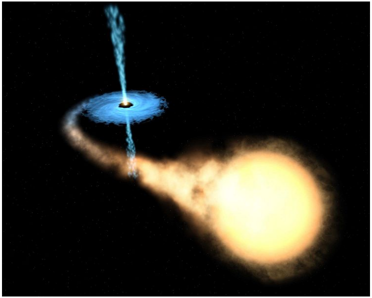
$$f(M) = \frac{PK^3}{2\pi G} = \frac{M \sin^3 i}{(1+q)^2},$$

- P is the period.
- K is the secondary's velocity semi-amplitude.
- i is the inclination.
- q is the mass ratio M_2/M.

Requires optical observations (spectra) and lightcurve fitting.

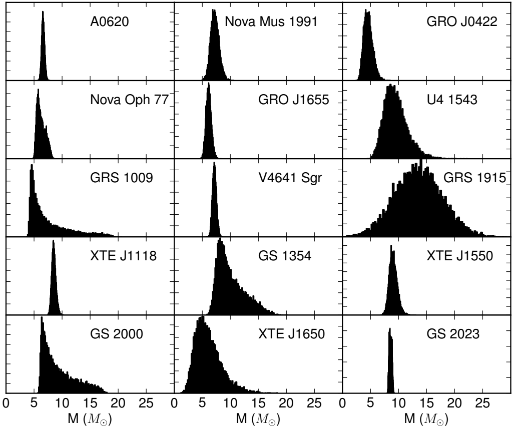
Our Sources

15 Lower-mass, Roche-Lobe overflow systems (e.g. A0620):

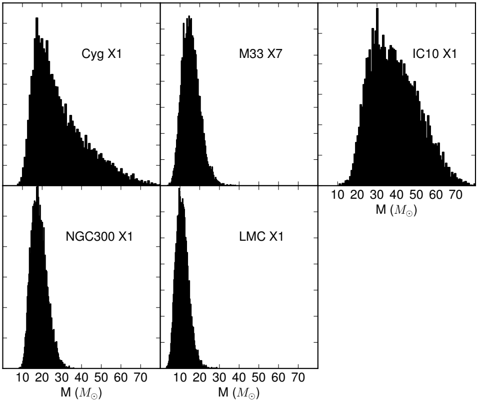


5 Higher-mass, wind-fed systems (e.g. M33-X7):

Low Mass Distributions



High Mass Distributions



Models for Mass Distribution

Model	$p(M\|\vec{\theta})$	Params $(\vec{\theta})$
Power Law	$p(M\|\vec{\theta}) \propto M^{\alpha},\ M_{\min} \leq M \leq M_{\max}$	$\alpha,\ M_{\min},\ M_{\max}$
Gaussian	$p(M\|\vec{\theta}) = N(\mu, \sigma, M)$	$\mu,\ \sigma$
Exponential	$p(M\|\vec{\theta}) \propto \exp\left(-\frac{M}{M_0}\right),\ M_{\min} < M$	$M_0,\ M_{\min}$
Log Normal	$p(M\|\vec{\theta}) \propto \frac{N(\mu,\sigma,\log M)}{M}$	$\mu,\ \sigma$
Two Gaussian	$p(M\|\vec{\theta}) \propto \alpha N_1 + (1-\alpha) N_2$	$\mu_1,\ \sigma_1,\ \mu_2,\ \sigma_2,\ \alpha$
Histogram	$p(M\|\vec{\theta}) \propto (\text{bin width})^{-1}$	bin locations

Models for Mass Distribution II

Each distribution tries to capture properties of $p(M|d)$:

- **Power Law** Why not? (Can capture M_{min} and M_{max}.)
- **Gaussian** A bump.
- **Exponential** Theoretically preferred.
- **Log Normal** Long tail.
- **Two Gaussian** Secondary bump.
- **Histogram** Non-parametric.

Likelihood

Assume:

- Observations independent.
- Errors on $f(M)$, q, and i independent.
- Ignore selection effects

$$p(d|\vec{\theta}) = \prod_{\text{systems}} \int dM\, p_{\text{obs}}(M) p(M|\vec{\theta})$$

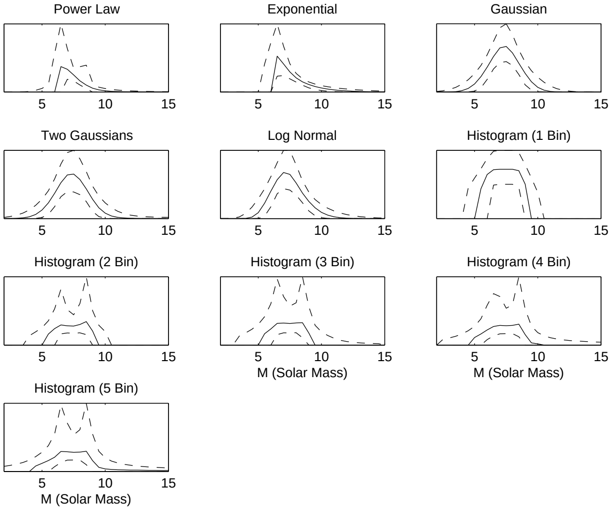
MCMC

Technique for sampling parameter space proportional to posterior probability:

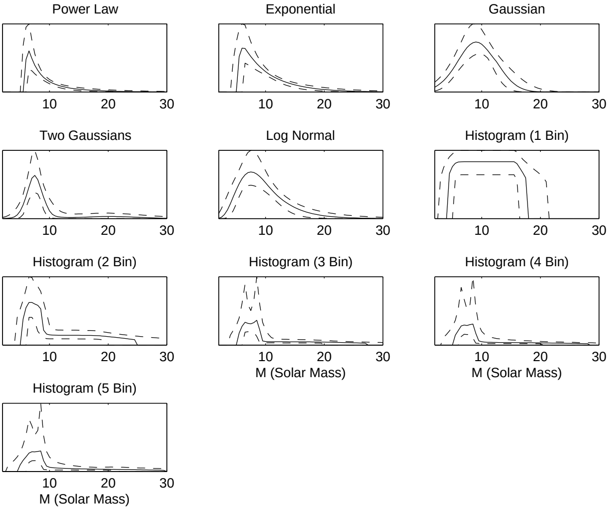
$$n(\vec{\theta}) \propto p(\vec{\theta}|d) = \frac{p(d|\vec{\theta})p(\vec{\theta})}{p(d)}$$

- Wander stochastically in parameter space.
- Propose local jumps, accept/reject probabilistically.
- Obtain a sequence of parameter values: $\{\vec{\theta}_i | i = 1, \ldots\}$
- Each sample represents one particular mass distribution favored by the data.
- Chain is a *distribution of distributions*.

Low Mass: Distributions of Distributions



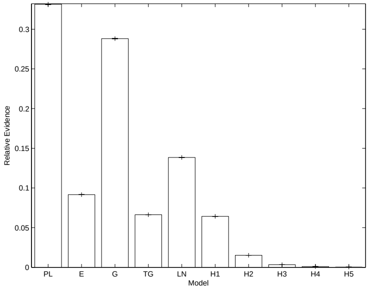
Complete Sample: Distributions of Distributions



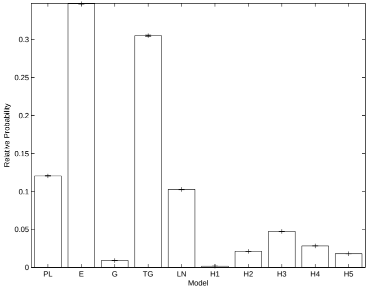
Model Selection

- Models look similar, but which ones are *the best*?
- Settle with another MCMC, this time jumping *between* models.
- Model counts give relative model probability.

Low-mass:



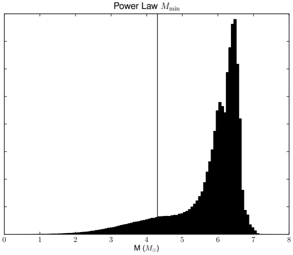
High-mass:



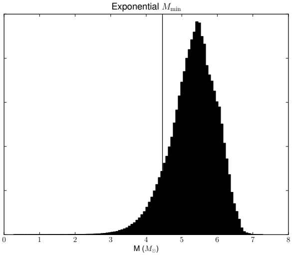
Minimum BH Mass

Strong evidence for a mass gap in both samples:

Low-mass systems:



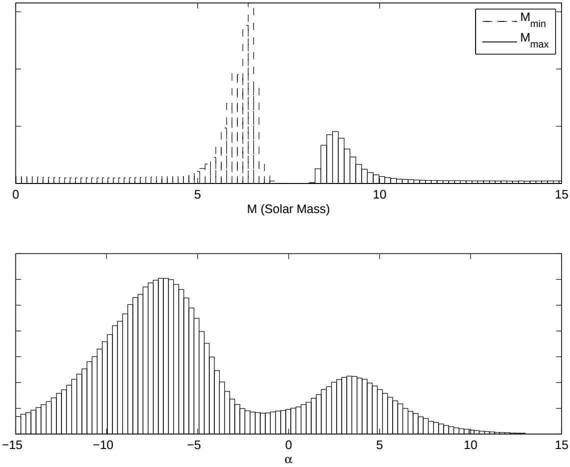
Combined sample:



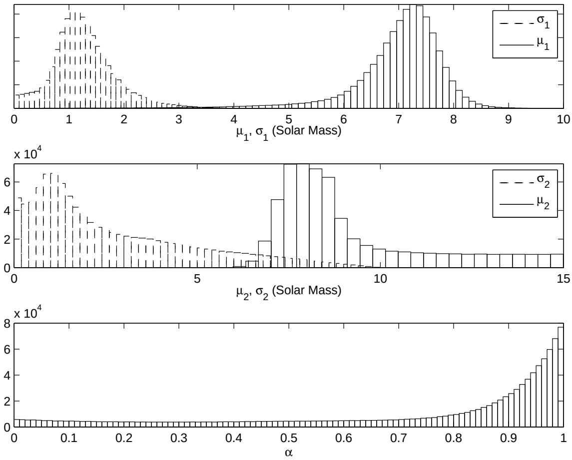
Conclusion

- Best models show large "mass gap": $M_{\text{min}} \sim 4.25 M_{\odot} > 3 M_{\odot}$
- Theoretically preferred model disfavored for low-mass systems, OK for complete sample.
- Difference between favored models for low-mass and complete sample implies different underlying distributions.
- Model selection is a powerful technique for extracting information about best-fitting models in the absence of prior knowledge.

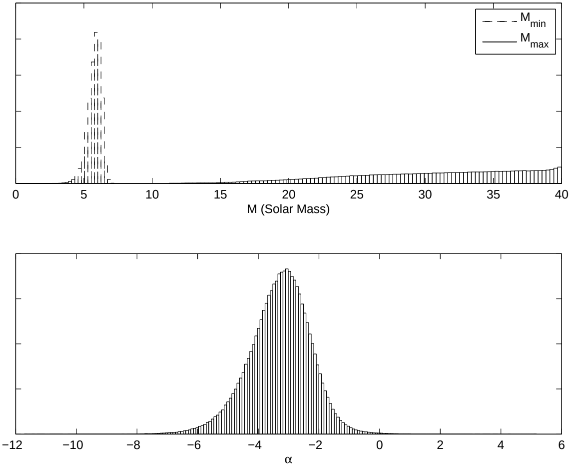
Power-Law Fitting to a Peak



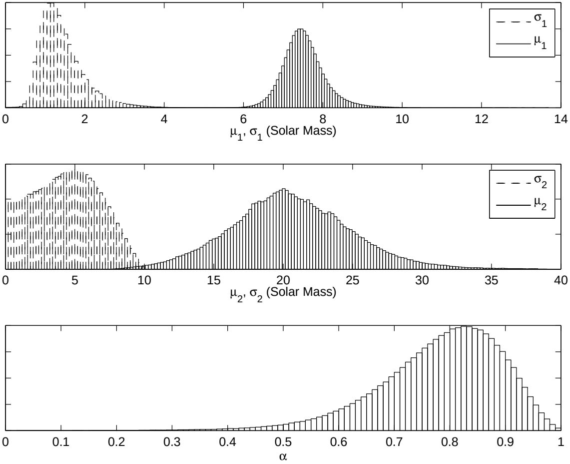
Two-Gaussian: No Second Bump



Tail Fixes Power-Law Degeneracy

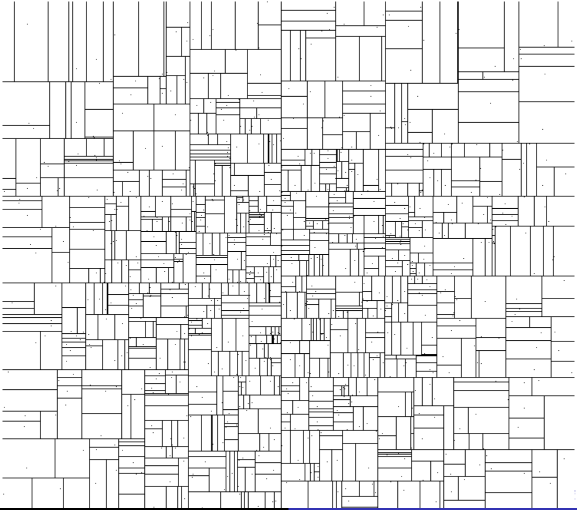


Two-Gaussian: Second Bump



Model Selection Techniques

Bayesian discussion on board:



Model Selection

Models

We assume a shape for the underlying mass distribution, and then “fit” to the observed masses.... (Bayesian math on board.)

Power Law

$$p(M|\vec{\theta}) = p(M|\{M_{\text{min}}, M_{\text{max}}, \alpha\}) = \begin{cases} AM^{\alpha} & M_{\text{min}} \le m \le M_{\text{max}} \\ 0 & \text{otherwise} \end{cases}$$

Decaying Exponential

(Preferred by Fryer & Kalogera (2001))

$$p(M|\vec{\theta}) = p(M|\{M_{\min}, M_0\}) = \begin{cases} \frac{e^{\frac{M_{\min}}{M_0}}}{M_0} \exp\left[-\frac{M}{M_0}\right] & M_{\min} \leq M \\ 0 & \text{otherwise} \end{cases} .$$

Gaussian and Two Gaussians

$$p(M|\vec{\theta}) = p(M|\{\mu, \sigma\}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{M-\mu}{\sqrt{2}\sigma}\right)^2\right].$$

and

$$p(M|\vec{\theta}) = p(M|\{\mu_1, \mu_2, \sigma_1, \sigma_2, \alpha\}) = \frac{\alpha}{\sigma_1\sqrt{2\pi}} \exp\left[-\left(\frac{M-\mu_1}{\sqrt{2}\sigma_1}\right)^2\right] + \frac{1-\alpha}{\sigma_2\sqrt{2\pi}} \exp\left[-\left(\frac{M-\mu_2}{\sqrt{2}\sigma_2}\right)^2\right].$$

Log Normal

$$p(M|\vec{\theta}) = p(M|\{\mu, \sigma\}) = \frac{1}{\sqrt{2\pi} M \sigma} \exp\left[-\frac{(\log M - \mu)^2}{2\sigma^2}\right].$$

Non-parametric Histograms

Histogram math on board.

Summary of Models

- Power Law
- Exponential
- Gaussian
- Two Gaussians
- Log Normal
- One-, two-, three-, four-, five-bin histograms