# Deriving upper limits on event magnitude in case of a non-detection

Christian Röver[1], Chris Messenger[1,2], Reinhard Prix[1]

[1]Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Hannover, Germany

[2]School of Physics and Astronomy, Cardiff University, Cardiff, UK

PHYSTAT2011
January 18th, 2011
CERN

# Overview:

**1.** The signal detection problem

**2.** Frequentist approach

**3.** Bayesian approach

**4.** Toy example
- Upper limit approaches
- A simplification
- Comparison

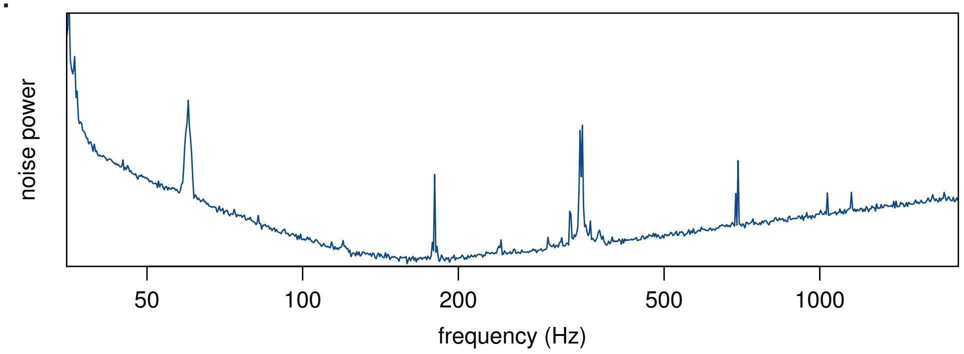
**5.** Conclusions

## Background: the measurement of gravitational waves

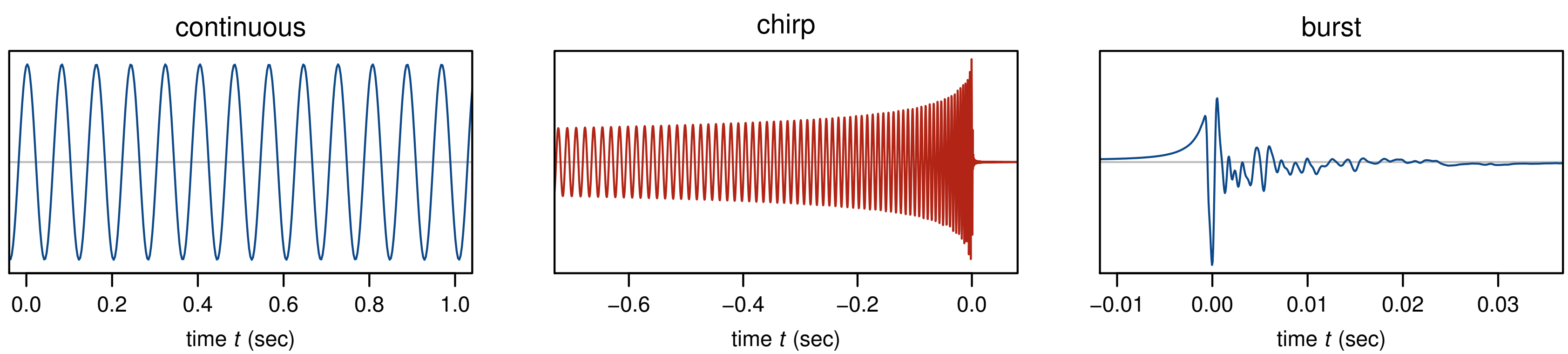
- Gravitational waves predicted by **general relativity** theory
- Large **interferometers** have been built to *measure* GWs

- output: **time series**

- Measurement **noise**:




- aimed for **signals**:




- time series analysis, signal detection

# Data model

- data $y(t)$ are a sum of **signal** and **noise**
- **signal** $s_\theta(t)$: function of time $t$; shape depending on parameters $\theta$
- **noise** $n(t)$: Gaussian with some power spectral density $S_n(f)$
- **Likelihood** function: *Whittle* approximation – Fourier-domain residual sum-of-squares

$$p(\theta \mid y) \;\propto\; \sum_j \frac{\left|\tilde{y}(f_j) - \tilde{s}_\theta(f_j)\right|^2}{S_n(f_j)}$$

# The frequentist approach

- "**Matched filter**" $\longrightarrow$ maximum likelihood filter / **generalized Neyman-Pearson** test:
  - compute likelihood under "*no signal*" hypothesis
  - maximize likelihood under "*signal*" hypothesis
  - likelihood ratio is **detection statistic**
  - determine threshold, claim detection if exceeded
- likelihood maximization **partly brute-force**/grid maximization, **partly analytical**

## The "loudest-event upper limit"

- detection procedure was ML / generalized Neyman-Pearson approach; detection/significance statement (based on $\mathrm{P}(D \mid H_0)$):

  *"If the data were only noise, a detection statistic value $\geq d$ would have been observed with probability $p$."*

  ($d$: observed detection statistic, $p$: p-value)

- in case of no detection (large p-value), derive an upper limit. upper limit statement (based on $\mathrm{P}(D \mid A, H_1)$):

  *"Had the signal amplitude been $\geq A^\star$, a larger detection statistic value ($> d$) would have been observed with probability $\geq \alpha$."*

  ($d$: observed detection statistic, $A^\star$: upper limit, $\alpha$: confidence level)

- ("**loudest event**" refers to (ML) maximization)

# The Bayesian approach

- detection and parameter estimation more separate problems
- detection: observe data $y$,
  then compute probabilities $\mathrm{P}(\text{“signal present”} \,|\, y)$ vs. $\mathrm{P}(\text{“no signal present”} \,|\, y)$
  detection/significance statement:

  *“(Given the observed data $y$,)*
  *the probability for the presence of a signal is $p$.”*
- parameter estimation: derive signal parameters' posterior distribution $\mathrm{P}(\theta \,|\, y, \text{“signal present”})$, marginalize to get (e.g.) marginal posterior of amplitude only.

  *“(Given the observed data $y$ and the presence of a signal,)*
  *the amplitude is less than $A^\star$ with probability $\alpha$.”*
- requires/allows prior specification for signal parameters and -hypotheses

# Two kinds of upper limit

- Frequentist limit:

$$\mathrm{P}\big(D \geq d \,\big|\, A \geq A^{\star}\big) \;\geq\; 90\%$$

- Posterior limit:

$$\mathrm{P}(A < A^{\star} \,|\, y, H_1) \;=\; 90\%$$

- **Questions**:
  - difference?
  - effect of **integration vs. maximization**?
  - effect of **parameter space size**?

# Example

- **noise**: white, Gaussian

$$n(t) \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma^2)$$

- **signal**: sinusoid

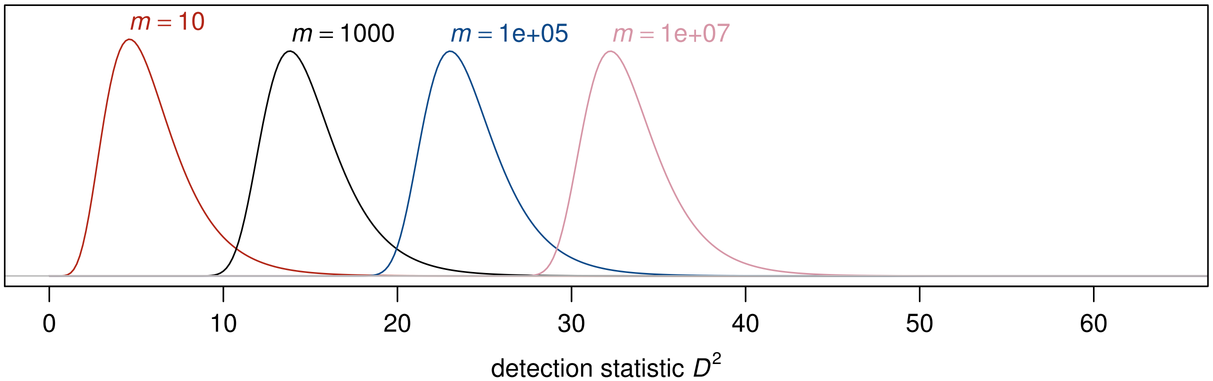
$$s_\theta(t) = A \sin(2\pi f t + \phi)$$

- signal **parameters** $\vec{\theta}$:
  - amplitude $A \geq 0$
  - frequency $f \in \{f_1, \ldots, f_m\}$
  - phase $\phi \in [0, 2\pi]$
- simple example exhibiting **common features**:
  - phase is a nuisance parameter
  - amplitude determines signal-to-noise ratio (SNR)
  - frequency requires numerical, "brute-force" search / optimization
    $m$: number of (Fourier) frequency bins ($\longrightarrow$ parameter space size, "**trials factor**")

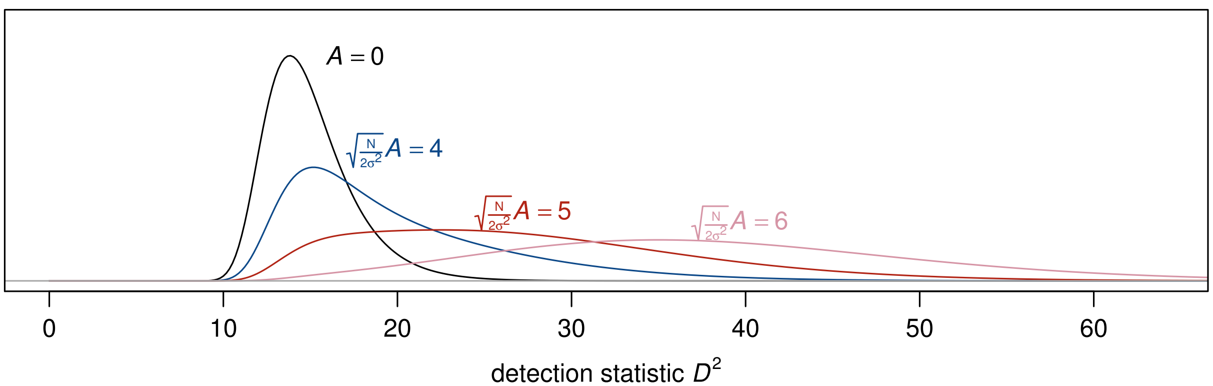
# Frequentist detection

- signal hypothesis $H_1$: 3D parameter space
  Likelihood (-ratio) maximization = maximization of **periodogram**;
  **Detection statistic**:
  $$d^2 \;:=\; \max_{j=1,\ldots,m} \tfrac{2}{N\sigma^2}\,|\tilde{y}_j|^2$$
  where $\tilde{y}$ is DFT of data $y$.
  (phase, amplitude: analytical; frequency: numerical)
- **No signal**: $\mathrm{P}(D^2 \,|\, H_0)$
  $D^2$ maximum of $m$ independent $\chi^2_2$
- **Signal**: $\mathrm{P}(D^2 \,|\, H_1, A)$
  $D^2$ maximum of $(m-1)$ independent $\chi^2_2$
  and $1$ noncentral-$\chi^2_2(\lambda = \frac{N}{2\sigma^2}A^2)$

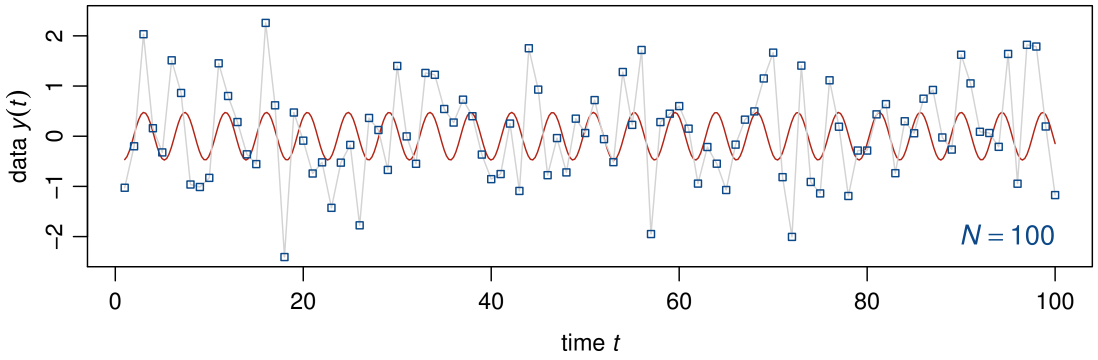
  (noncentrality parameter $\lambda = \frac{N}{2\sigma^2}A^2 \longrightarrow$ "signal to noise ratio (SNR)") distribution under $H_0$



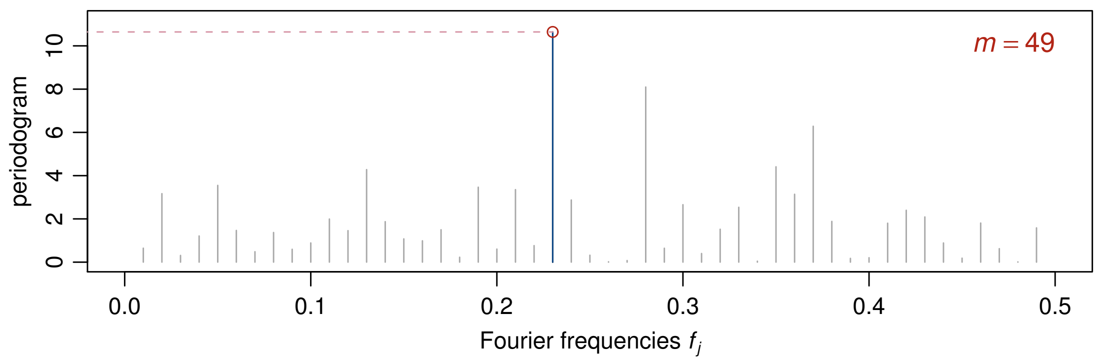

distribution under $H_1$ ($m = 1000$)

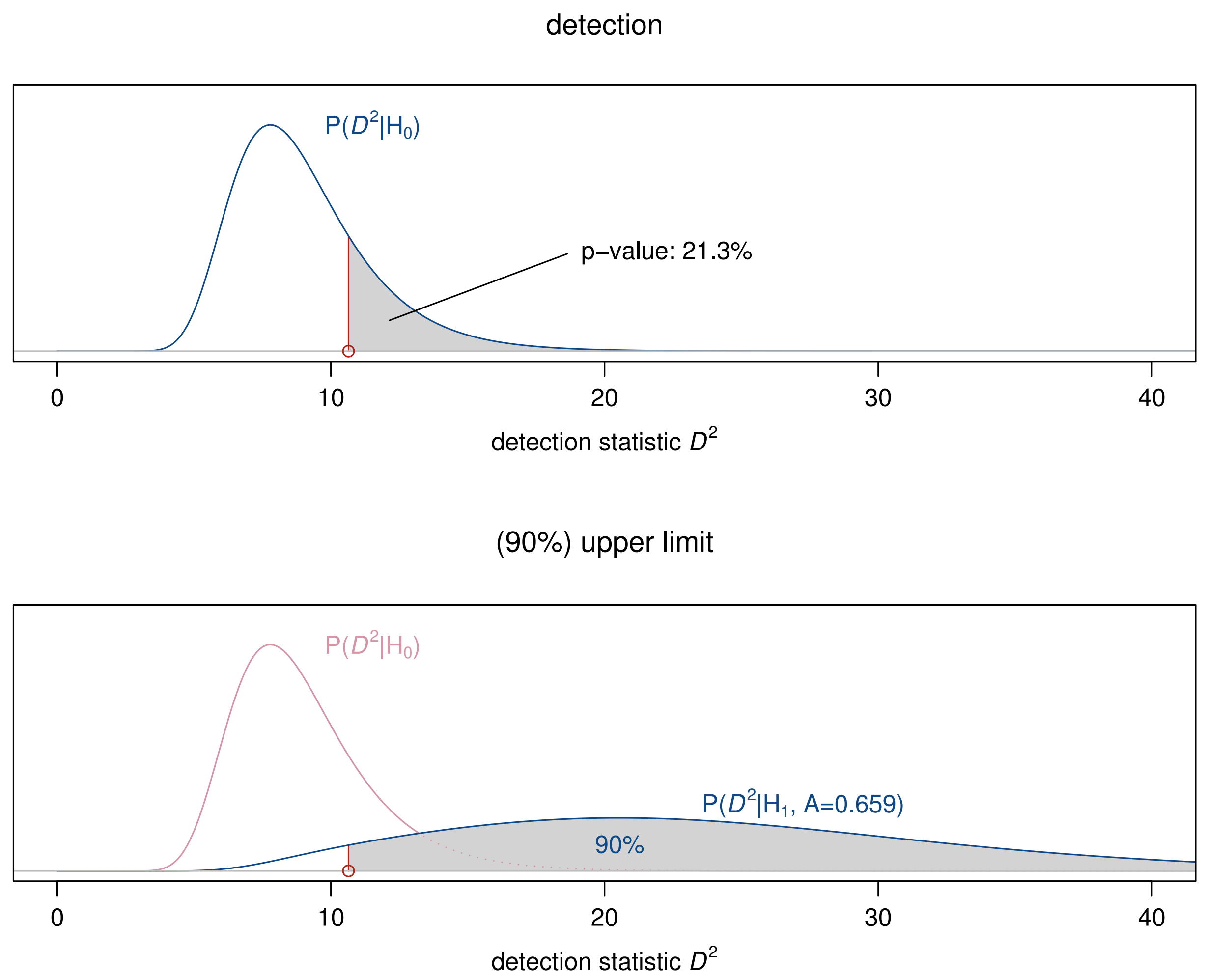


take data,...



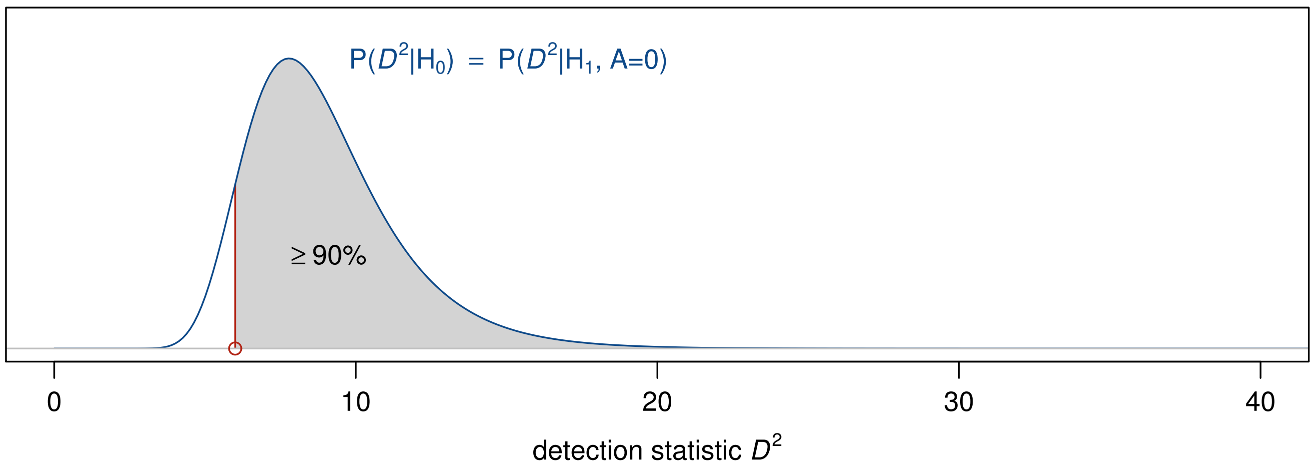

...compute periodogram, determine maximum $D^2$.




detection
P(D²|H₀)
p–value: 21.3%
0
10
20
30
40
detection statistic D²
(90%) upper limit
P(D²|H₀)
P(D²|H₁, A=0.659)
90%
0
10
20
30
40
detection statistic D²

# Note: Zero upper limits

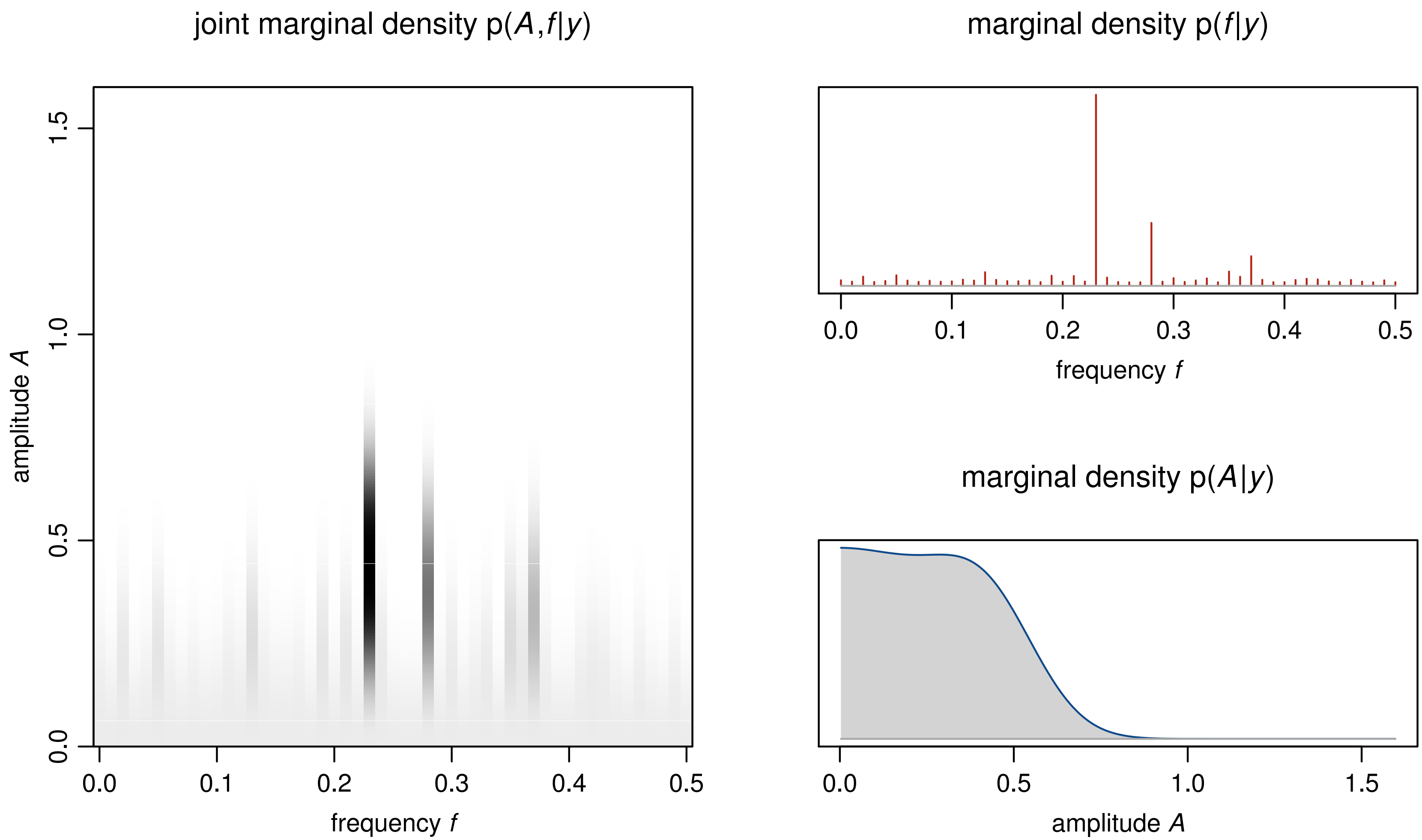
- upper limit is **zero** if $D^2$ falls in lower tail (under $H_0$)




- $\longrightarrow$ zero 90% upper limit 10% of times (under $H_0$)

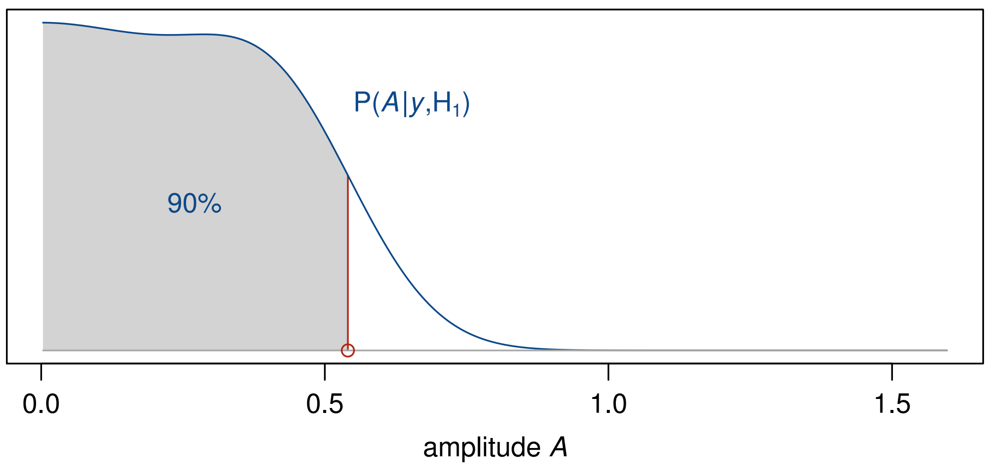
# Bayesian detection

- specify **priors**: uniform in phase, frequency, amplitude
- (**detection**: compare $\mathrm{P}(H_0 \,|\, y)$ vs. $\mathrm{P}(H_1 \,|\, y)$ — not of concern here)
- **upper limit**:
  - determine **posterior** $\mathrm{P}(A, f, \phi \,|\, y, H_1)$
  - **marginalize** to get $\mathrm{P}(A \,|\, y, H_1)$
  - amplitude's 90% limit is posterior's 90% **quantile**

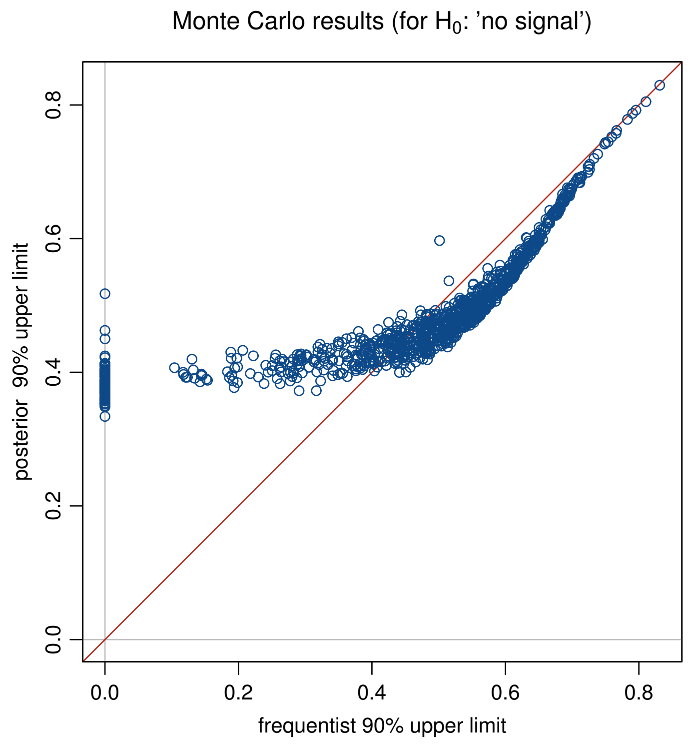

joint marginal density p(A,f|y)
amplitude A
0.0
0.5
1.0
1.5
0.0
0.1
0.2
0.3
0.4
0.5
frequency f
marginal density p(f|y)
0.0
0.1
0.2
0.3
0.4
0.5
frequency f
marginal density p(A|y)
0.0
0.5
1.0
1.5
amplitude A

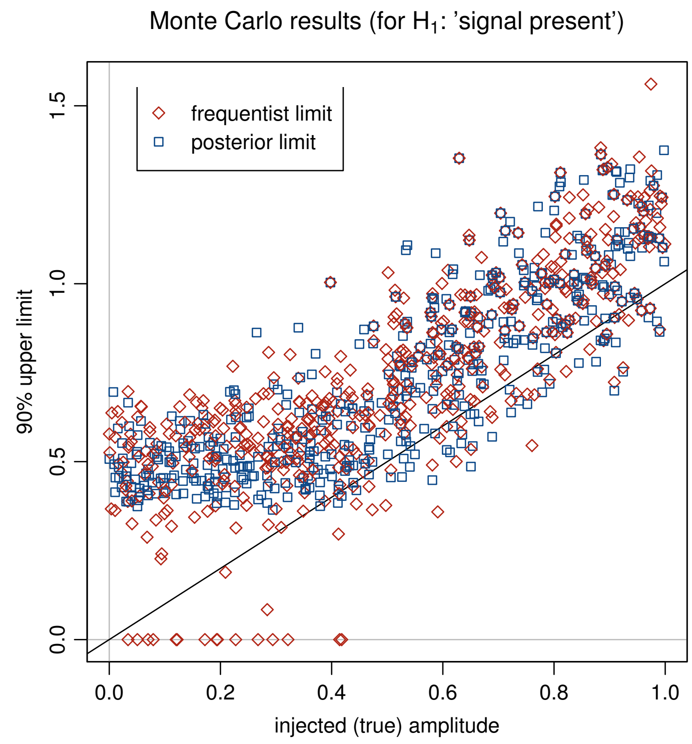
# Posterior upper limit

- **upper limit**



# Monte Carlo comparison: frequentist & posterior



# Integration vs. maximization

- frequentist limit based on **maximum** of periodogram
  (consequence of *generalized Neyman-Pearson* approach)
- posterior limit based on **integration** across frequencies

$$p(A|y) = \sum_j p(A|f_j, y) \times p(f_j|y)$$

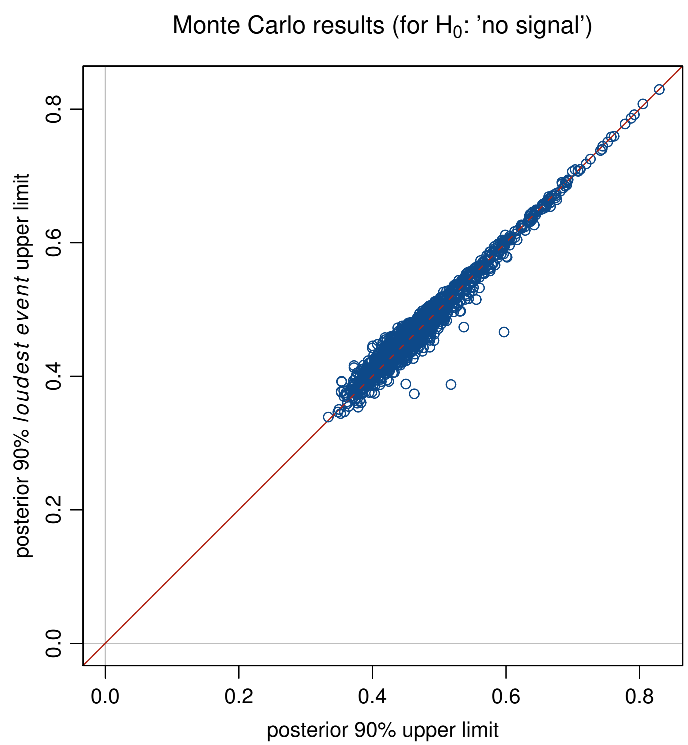
**But**: integral also dominated by "loudest frequency bin"

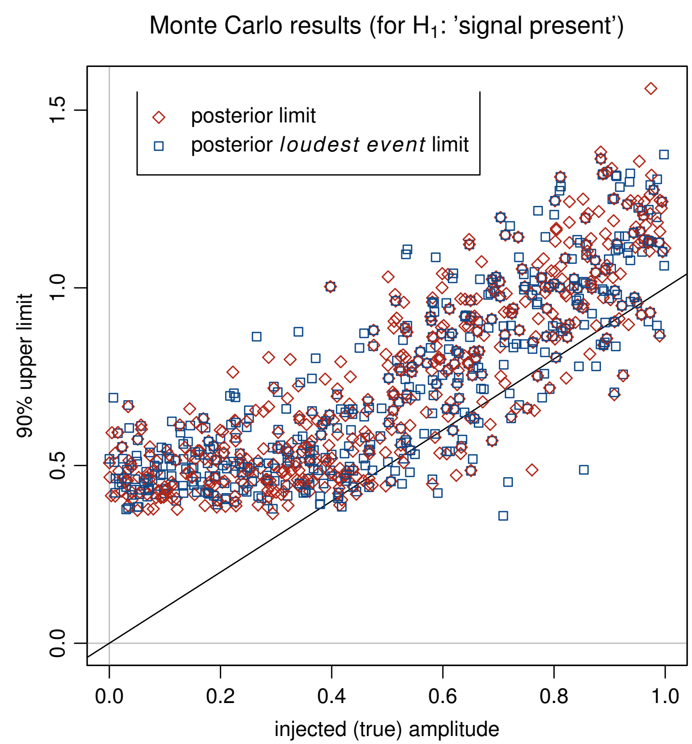
- especially for loud signals, also for noise only
- (makes sense: what you can rule out is determined by loudest observation)
- idea: check information loss via Bayesian upper limit based on loudest frequency bin

## Posterior "loudest event" limit

- slight modification:
  instead of $\mathrm{P}(A \,|\, y)$ consider $\mathrm{P}\big(A \,\big|\, \max_j |\tilde{y}_j|^2\big)$
- likelihood $\mathrm{P}\big(\max_j |\tilde{y}_j|^2 \,\big|\, A\big)$ already derived for frequentist limit

# Monte Carlo comparison: posterior & posterior loudest event

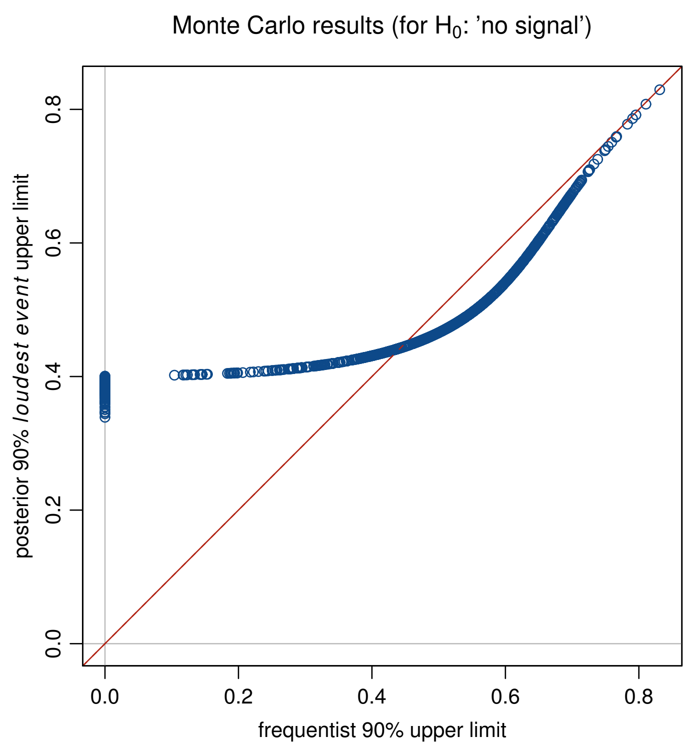


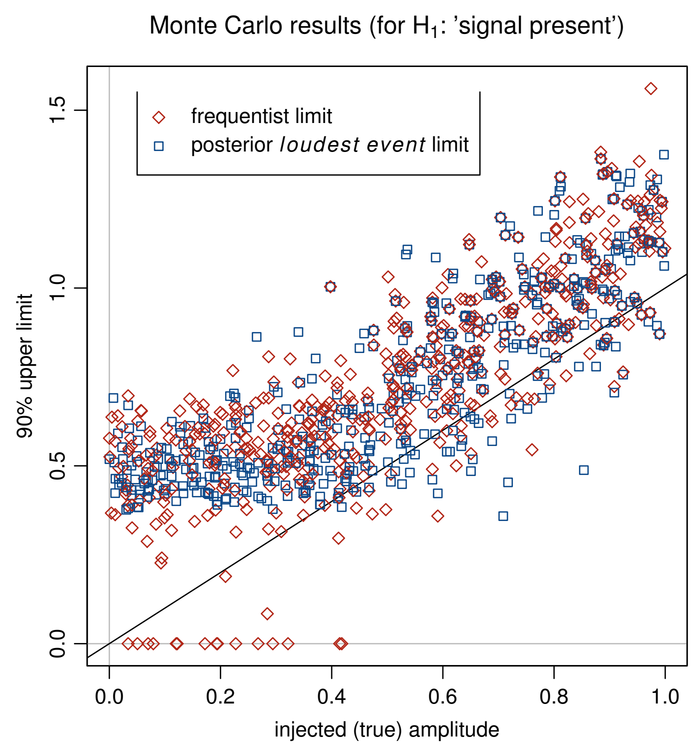



# Posterior "loudest event" limit

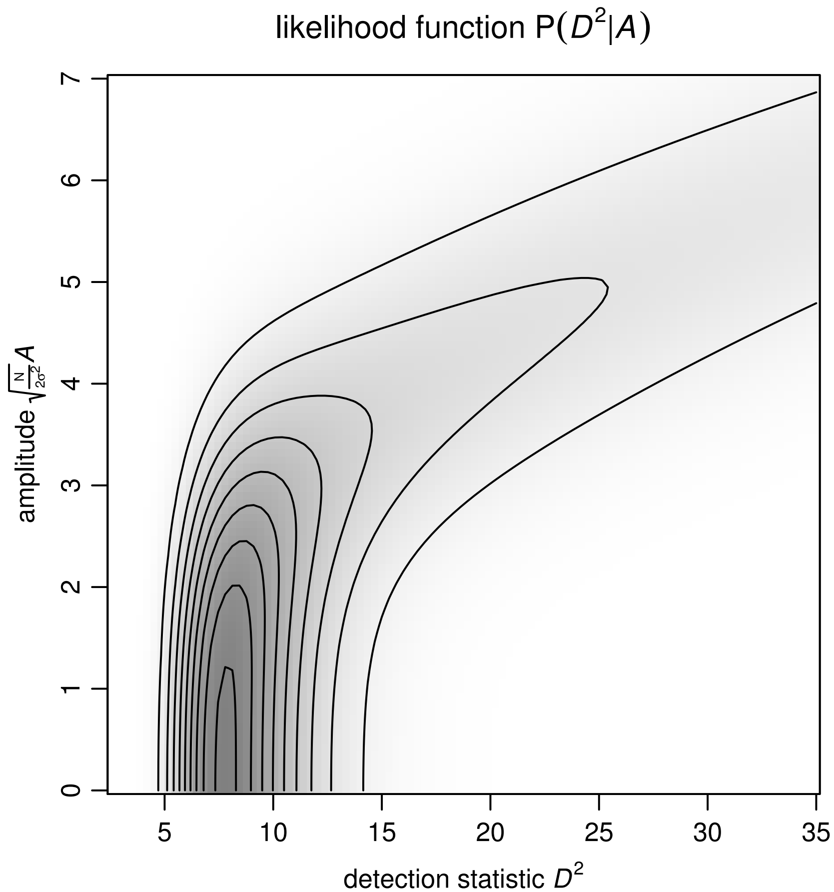
- small loss of information
  (especially for large $m$ (!))
- allows better 1-to-1 comparison with frequentist limit

# Monte Carlo comparison: loudest event limits

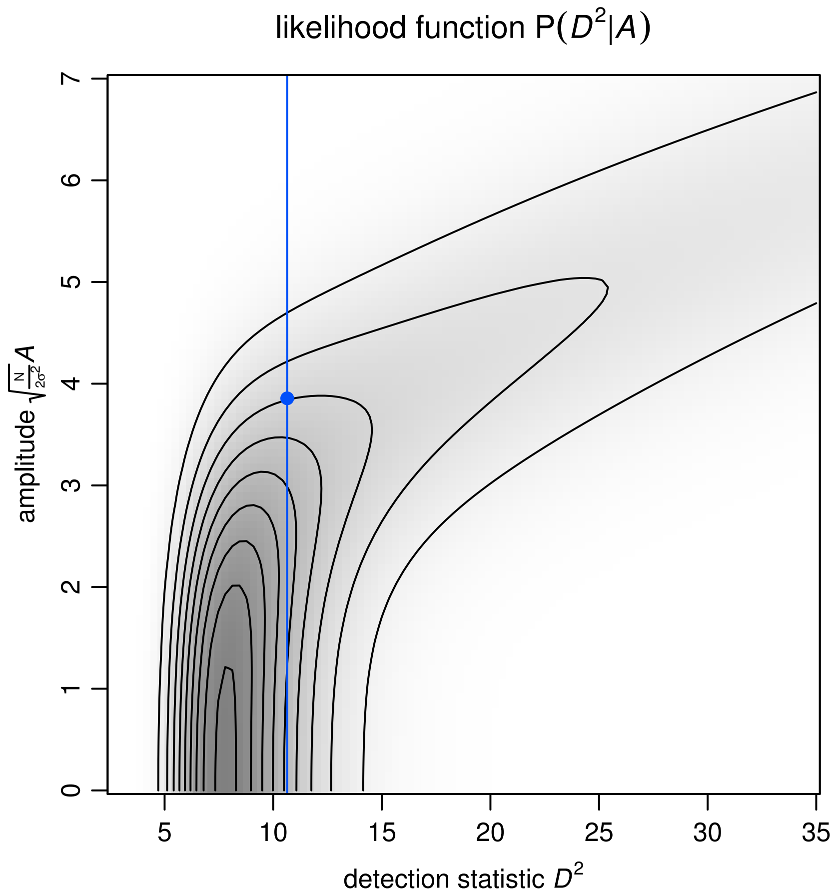


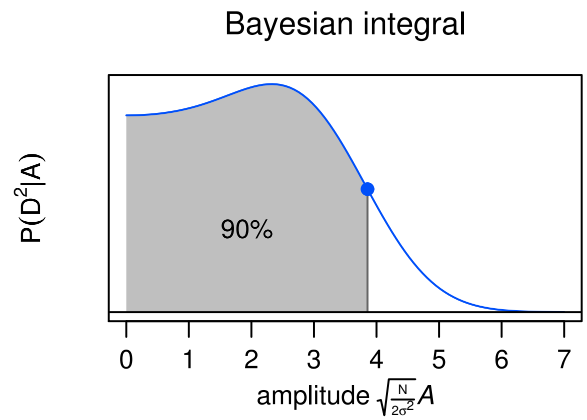



# Upper limit computation

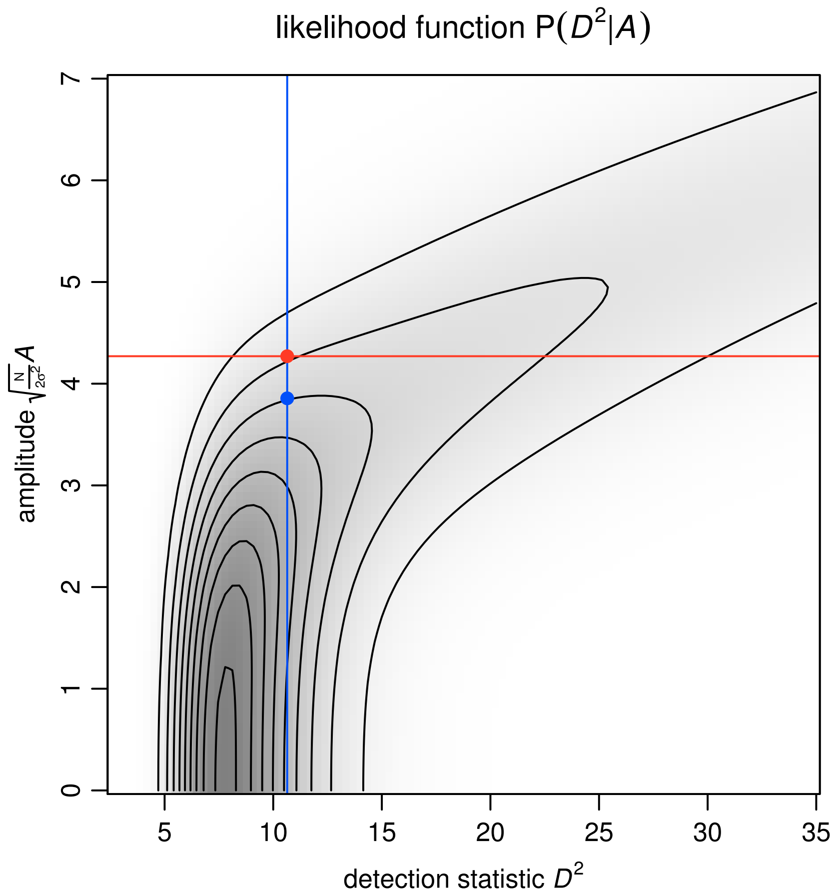


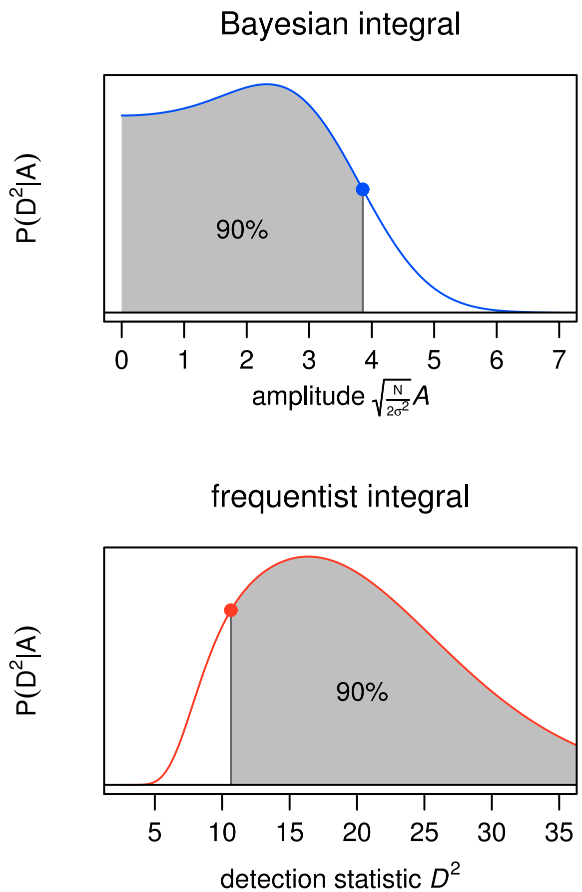
# Upper limit computation



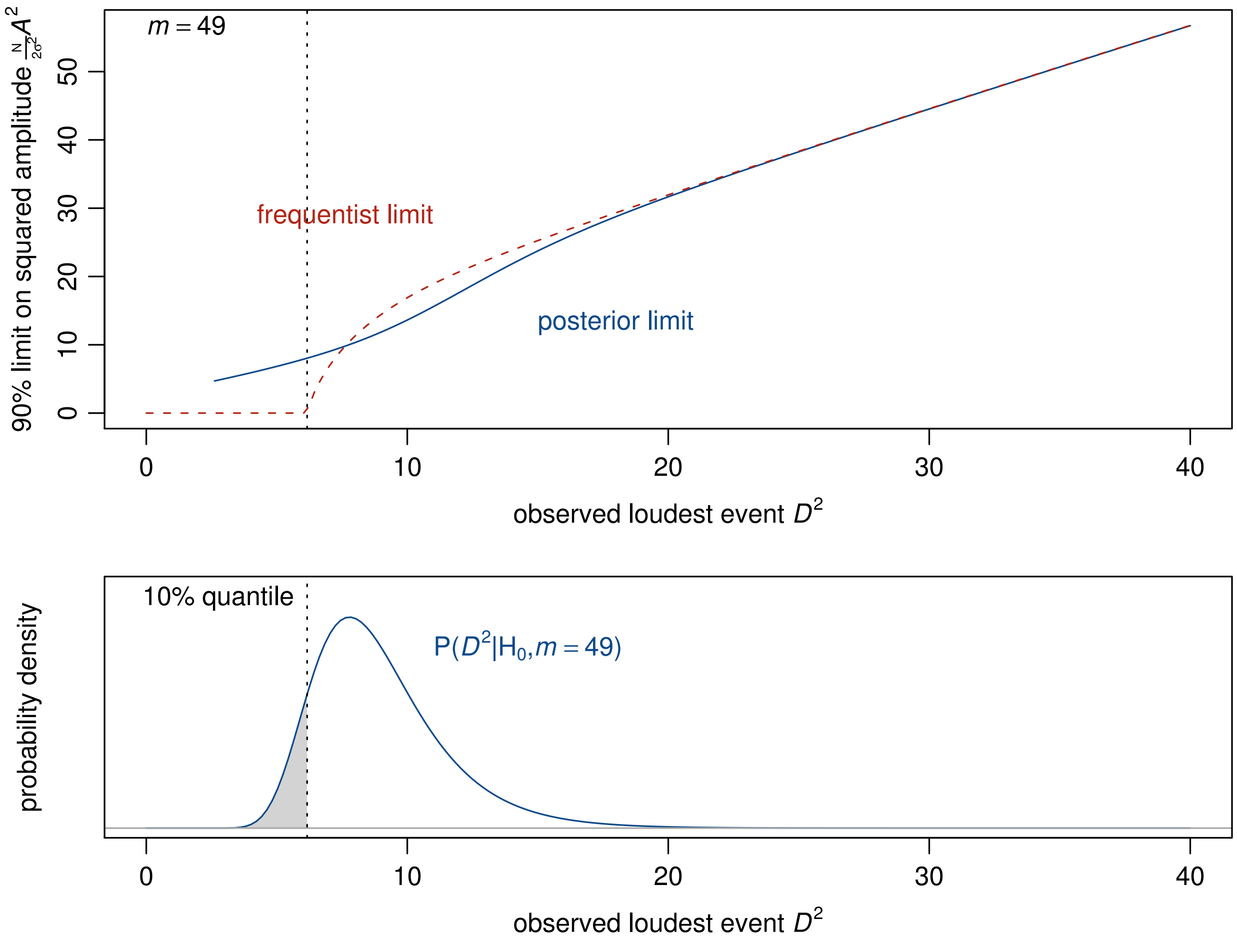



# Upper limit computation

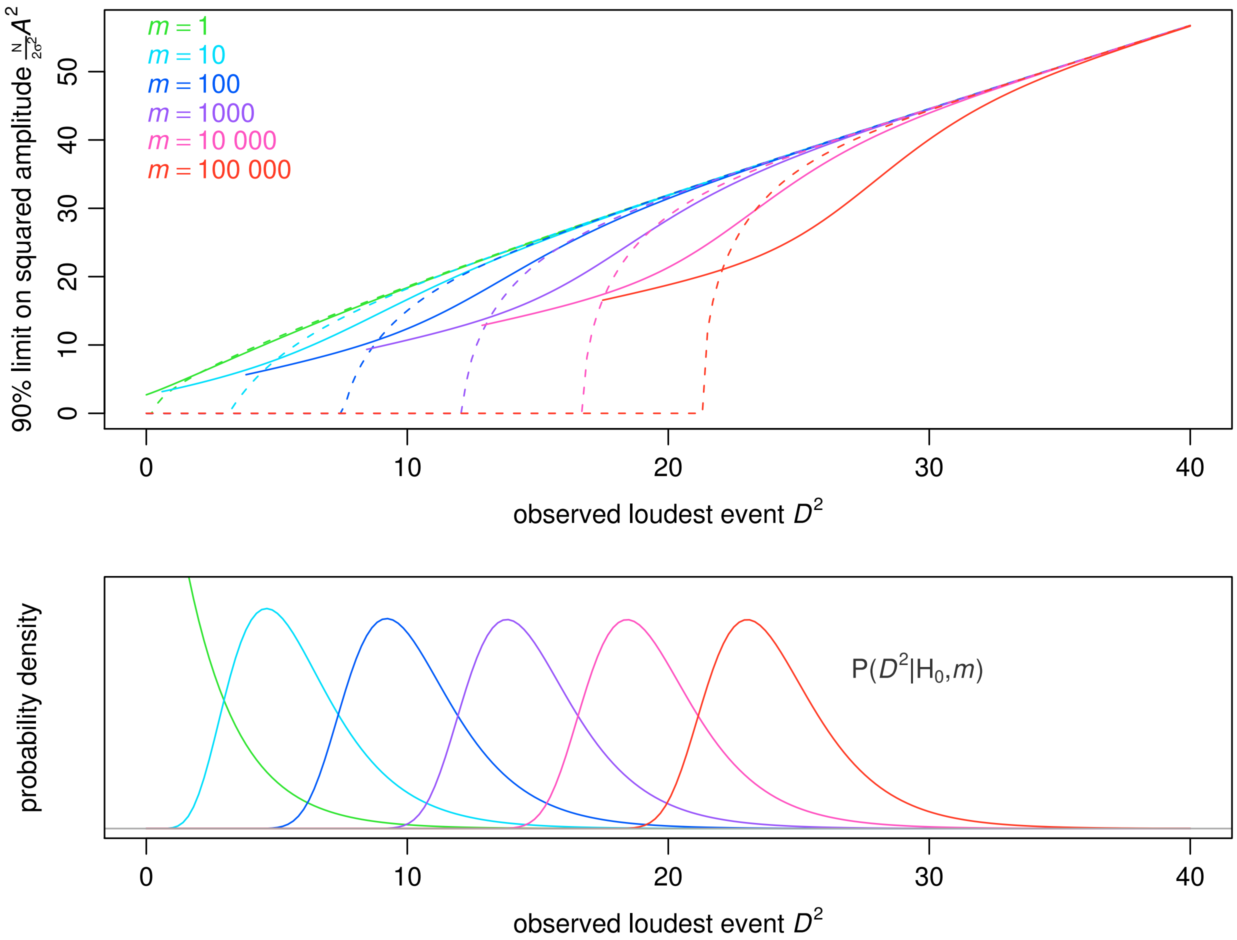


# 1-to-1 mapping to upper limit



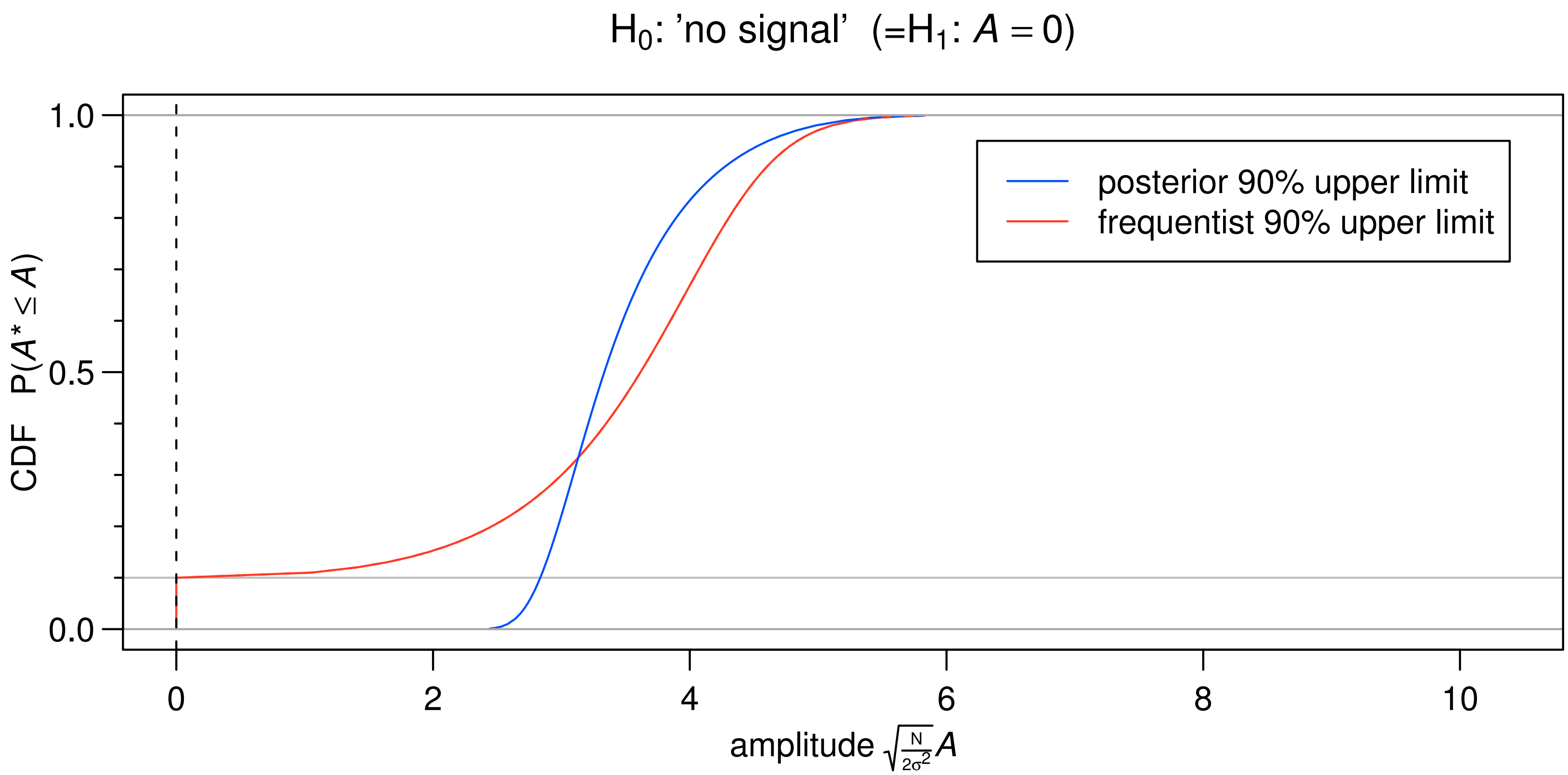
# 1-to-1 mapping to upper limit



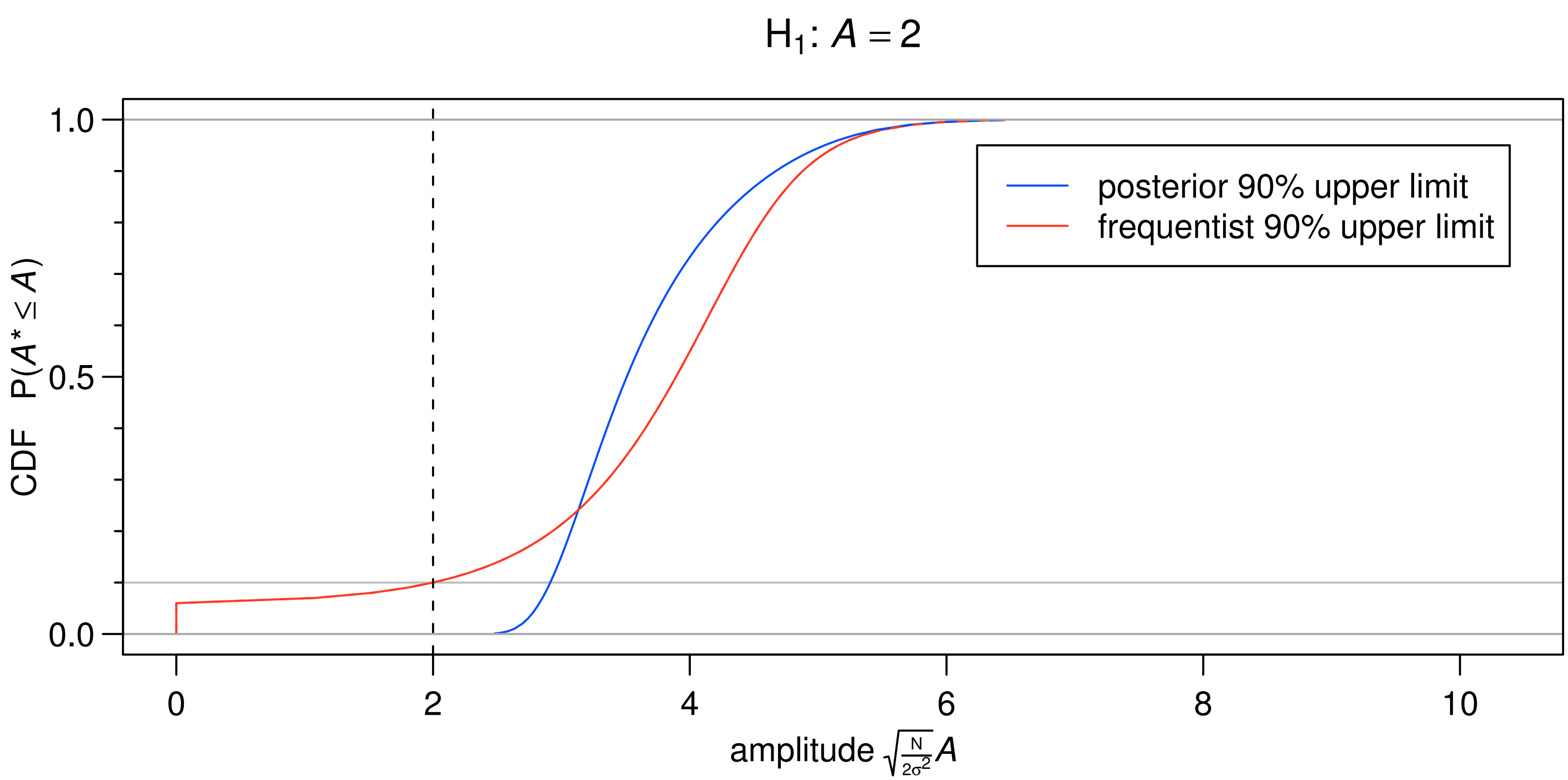
# Upper limits' distributions

- have:
  - distribution of $D^2$ (under $H_0$, $H_1$)
  - mapping $D^2 \longrightarrow 90\%$ limit
- derive: distribution of upper limit

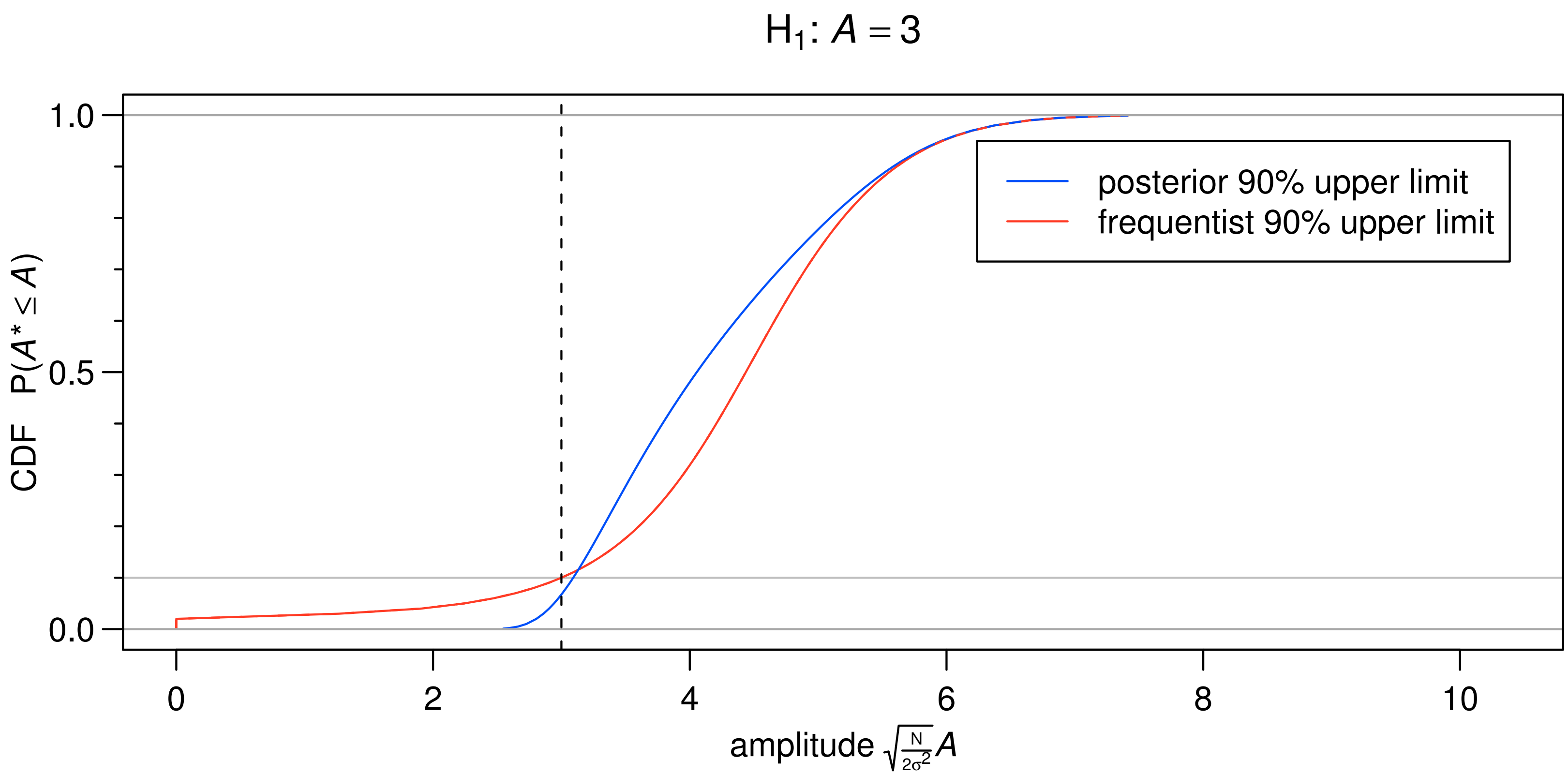
# Upper limits' distributions



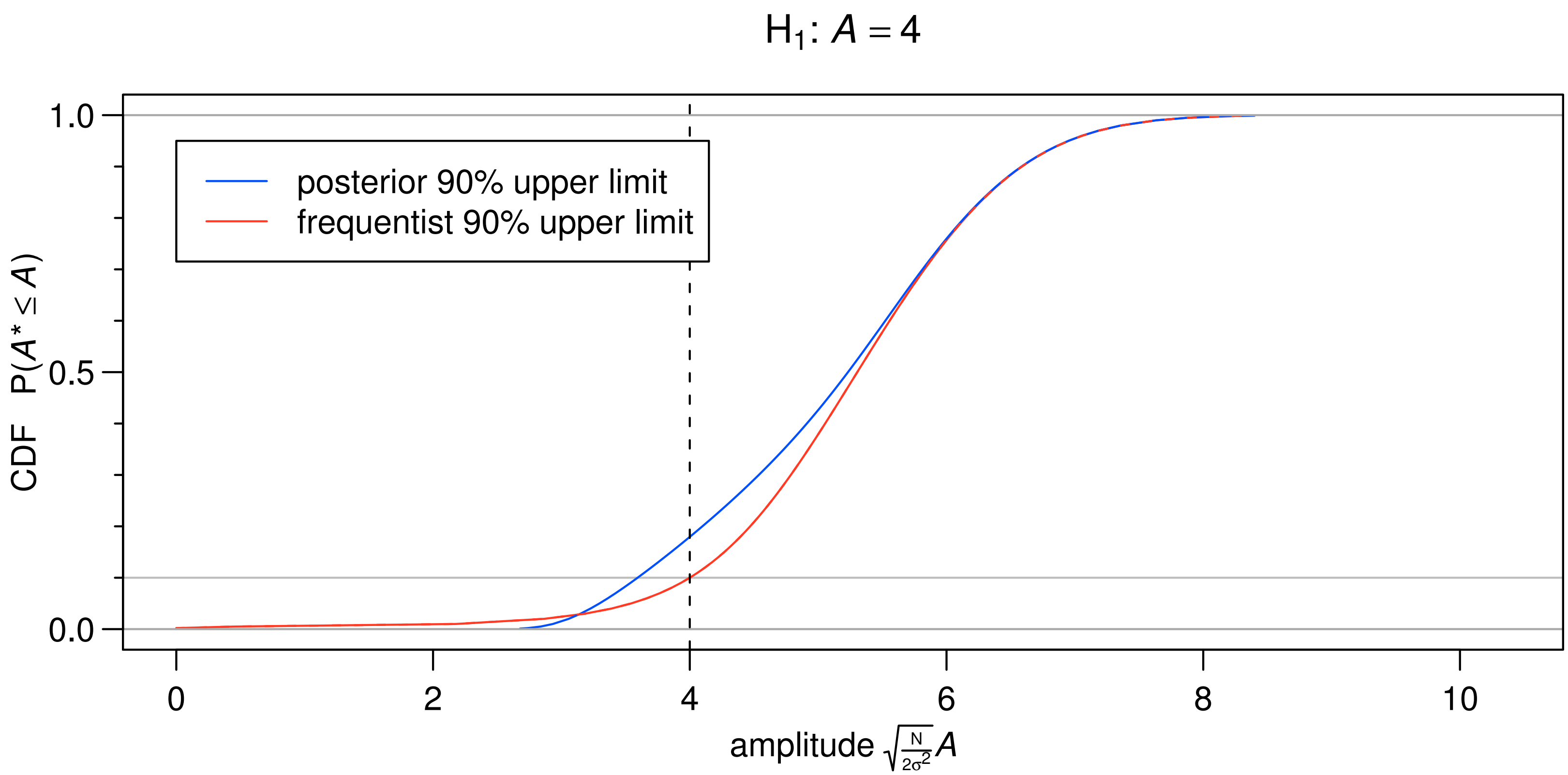
# Upper limits' distributions



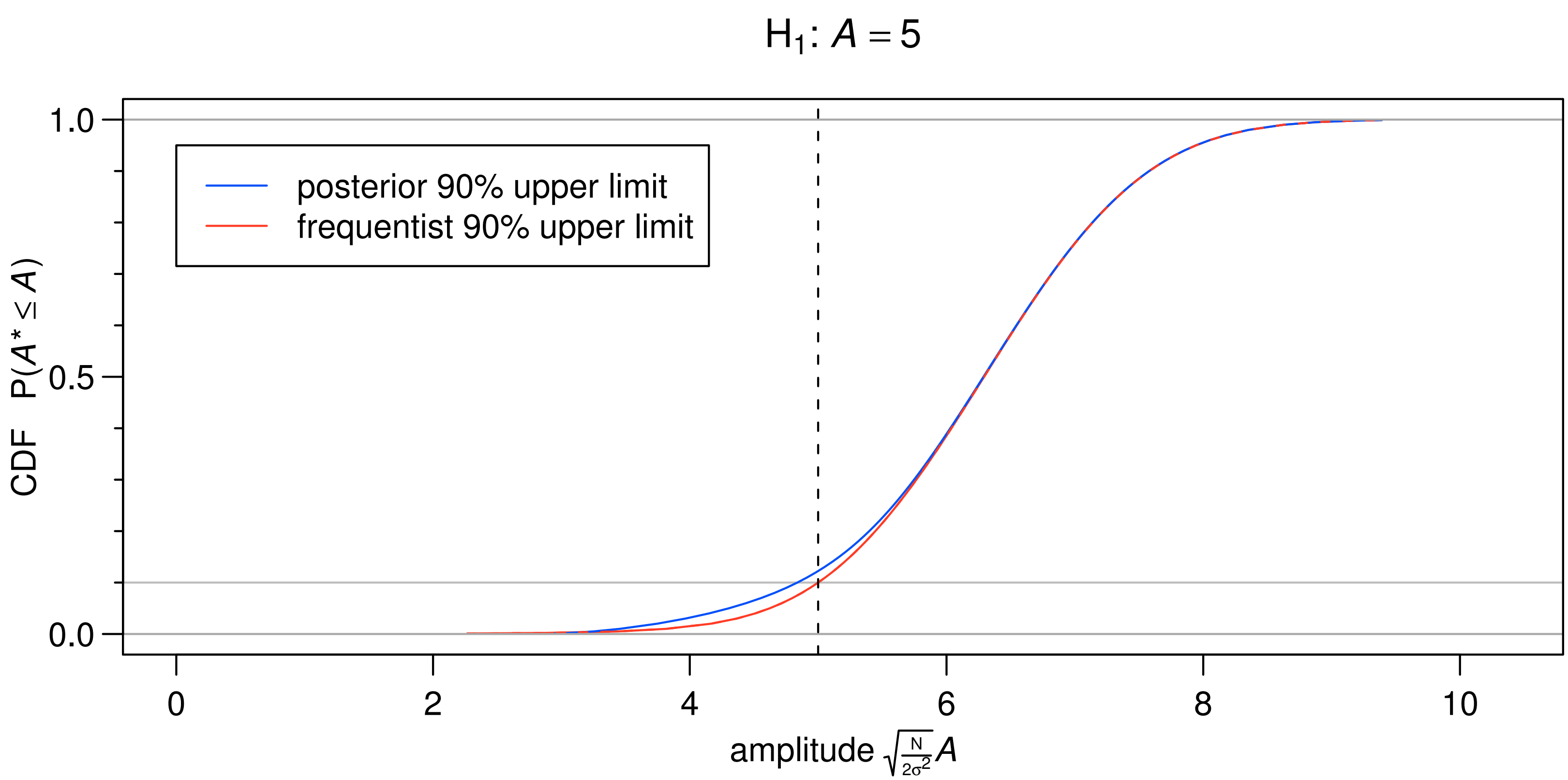
# Upper limits' distributions



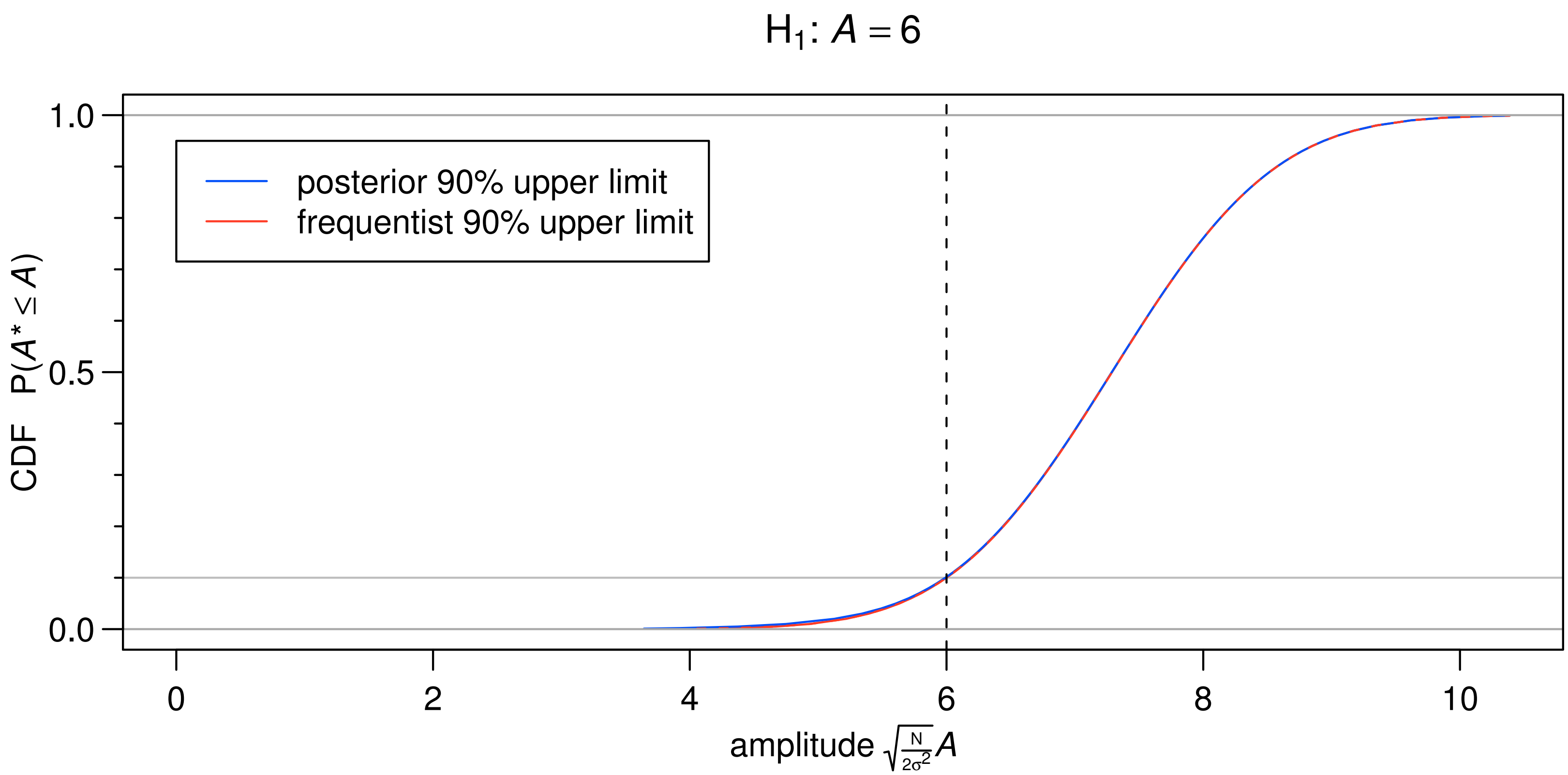
# Upper limits' distributions



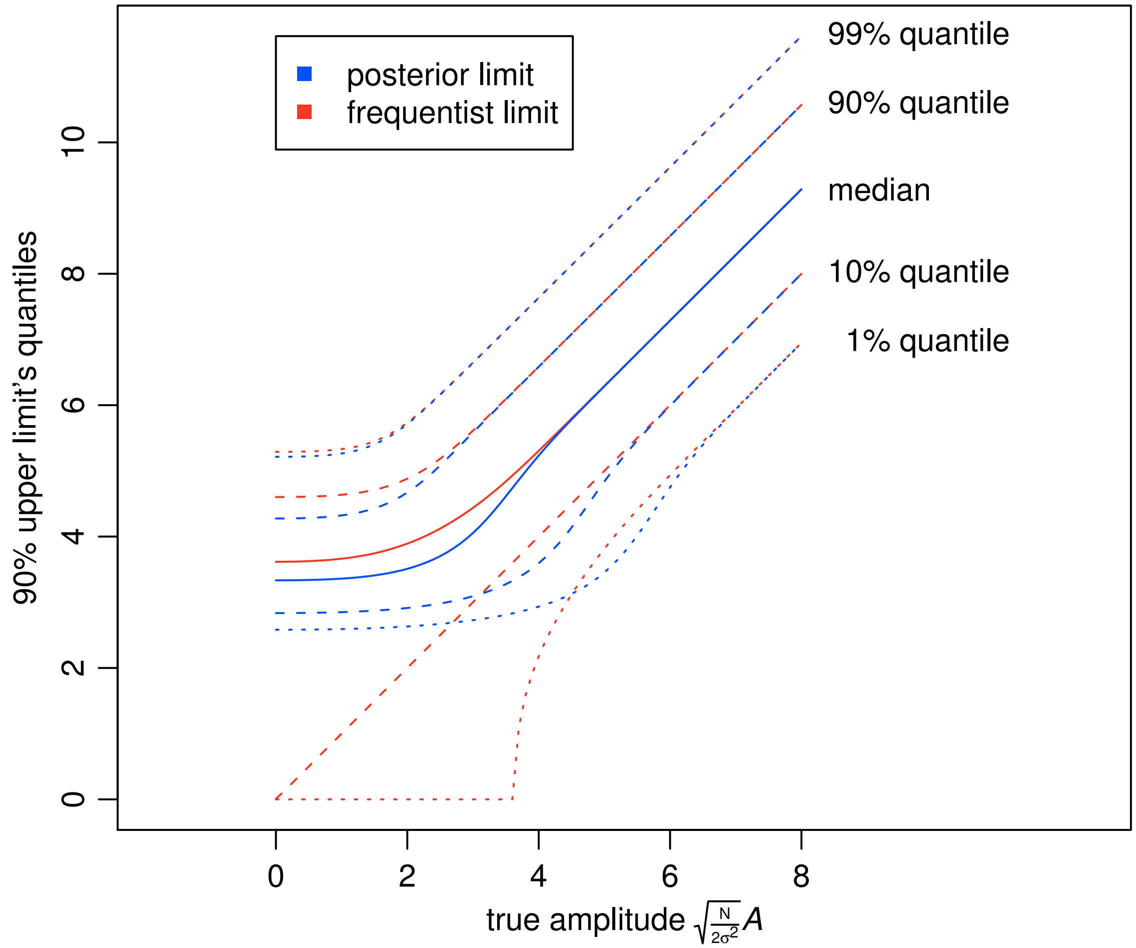
# Upper limits' distributions



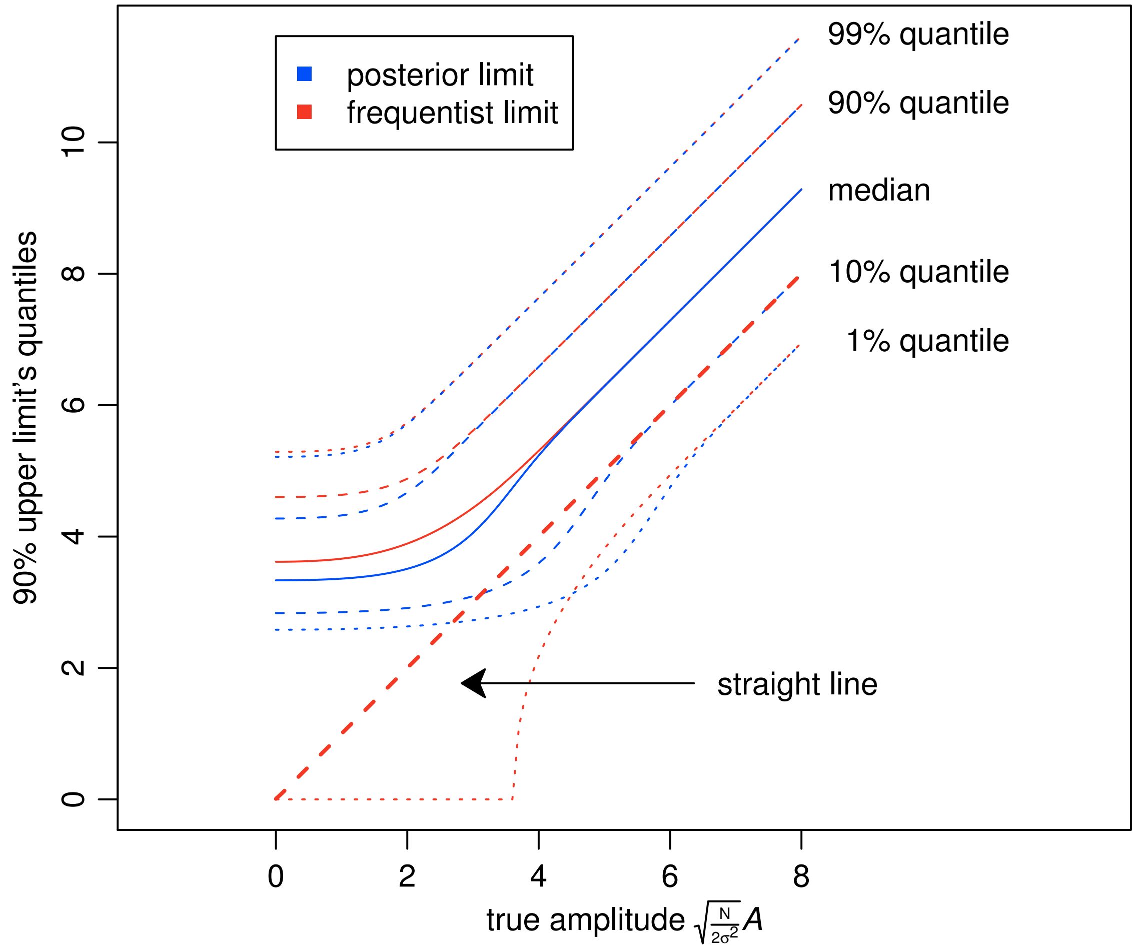
# Upper limits' distributions



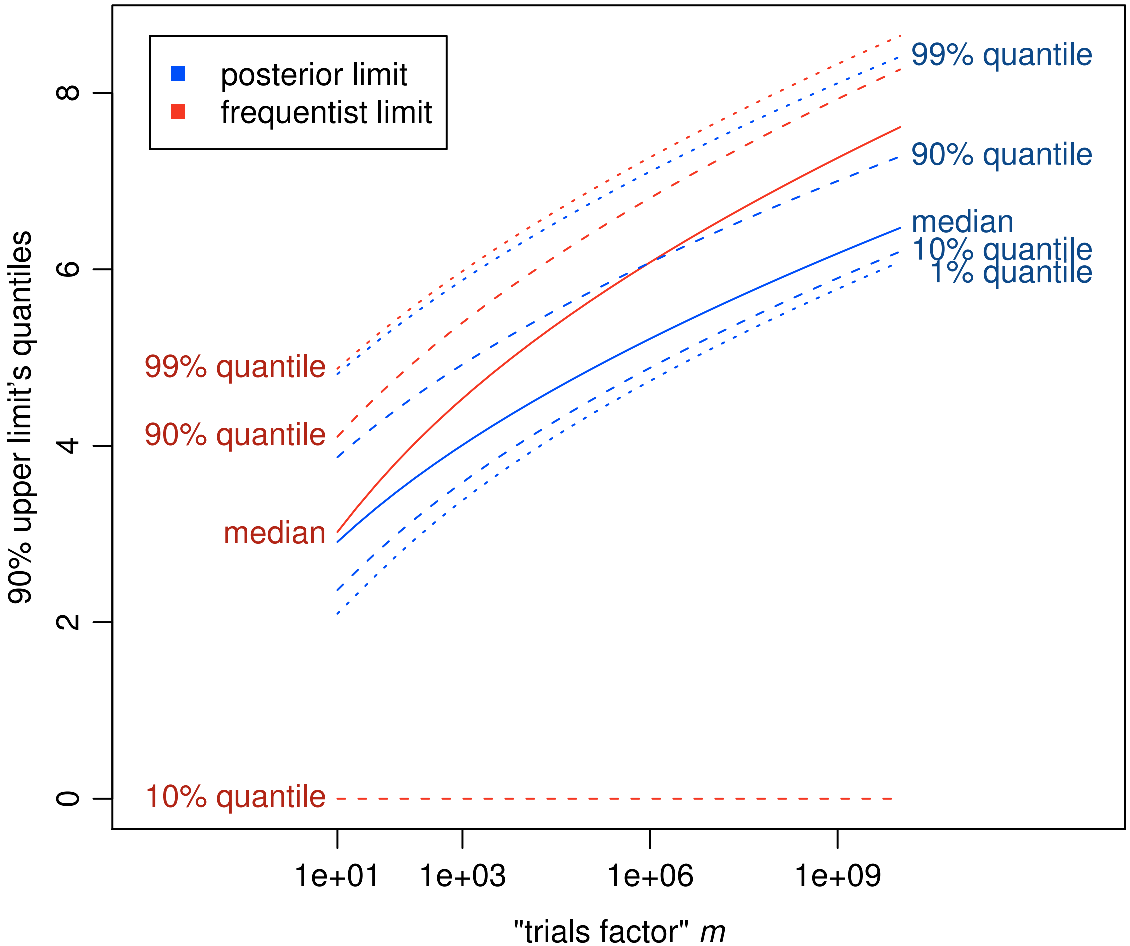
# Upper limits' distributions



# Upper limits' distributions



# Upper limits' distributions under $H_0$



# Conclusions

- maximization (instead of integration) no problem
- Bayesian and frequentist limits similar for large SNR
  (at least here, for uniform amplitude prior)
- different for (interesting case of) low SNR:
  - Bayesian upper limit states what can be ruled out with 90% certainty
  - frequentist 90% upper limit $=$ random variable with 10% quantile at true value
  - hard to interpret, occasional zero upper limits
  - posterior limit more sensitive for large $m$
- Example only, but: gravitational-wave detection problem very similar –
  nuisance parameters, partly analytically, partly numerically tractable, only subset of parameters affecting SNR, detection statistic $\chi^2$-distributed,...
- questions remaining
  - practical computation (bootstrapping?)
  - additional amplitude parameters?