Bayesian estimation of optical properties of the human head via 3D structural MRI

June 23, 2003 at ECBO 2003

Alex Barnett

Courant Institute, New York University

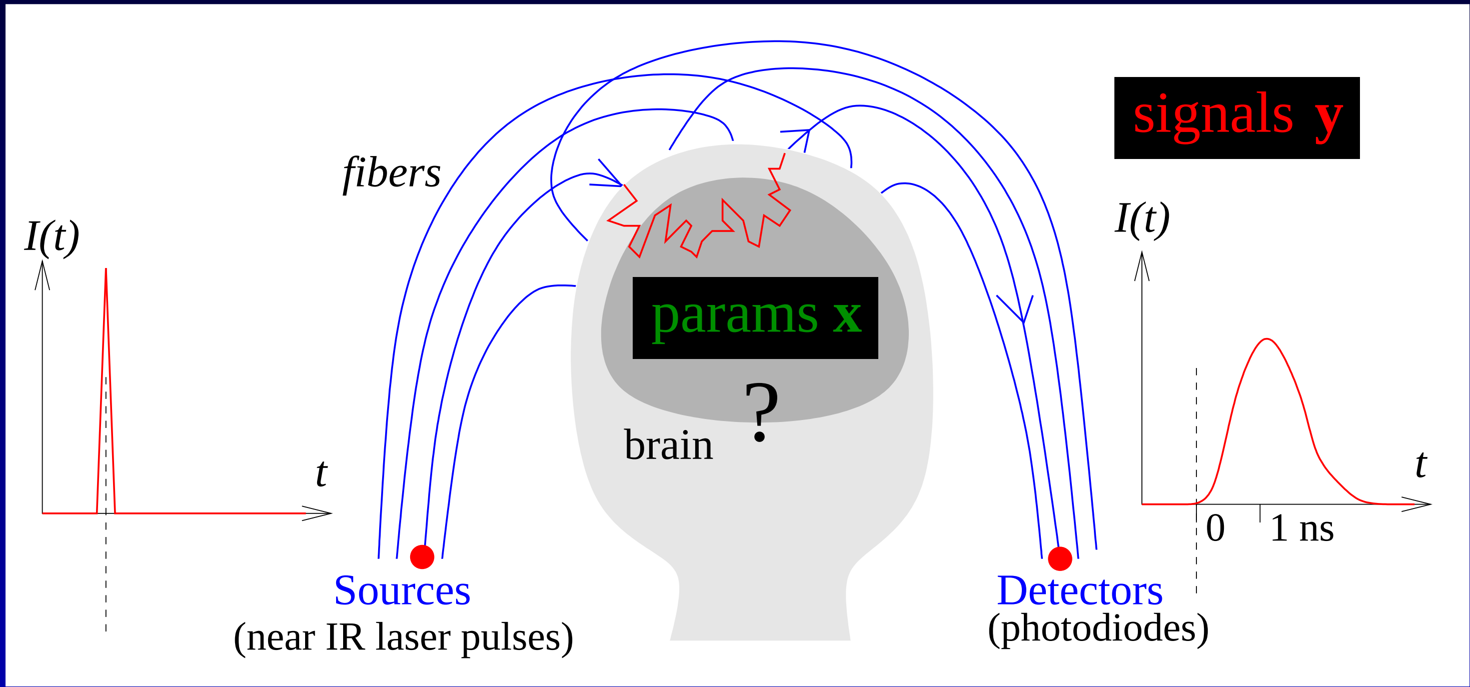
Collaborators (*NMR Center, Mass. Gen. Hosp., Boston*)

David Boas	Joe Culver	Anna Custo (MIT)
Gregory Sorensen	Anders Dale	

Funding: CIMS, NIH, CIMIT

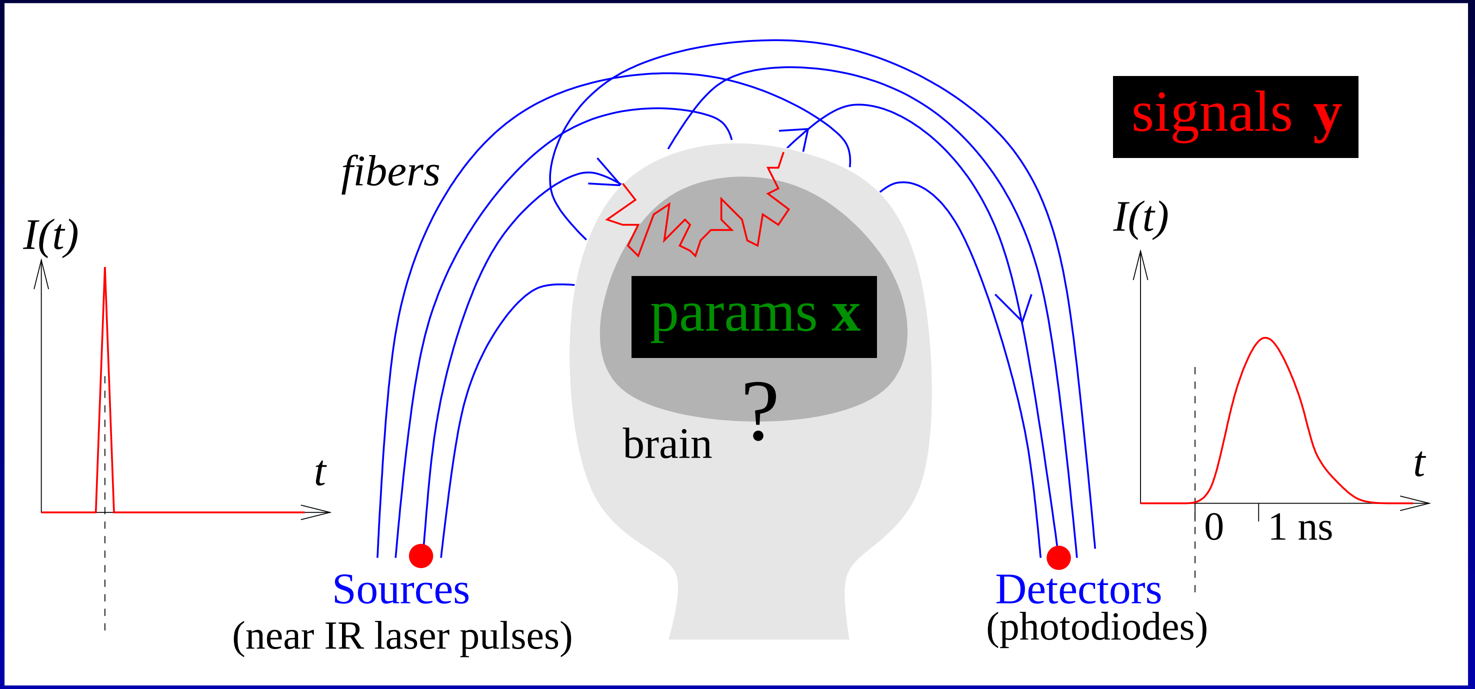
The big picture

Diffuse Optical Tomography (DOT):



The big picture

Diffuse Optical Tomography (DOT):



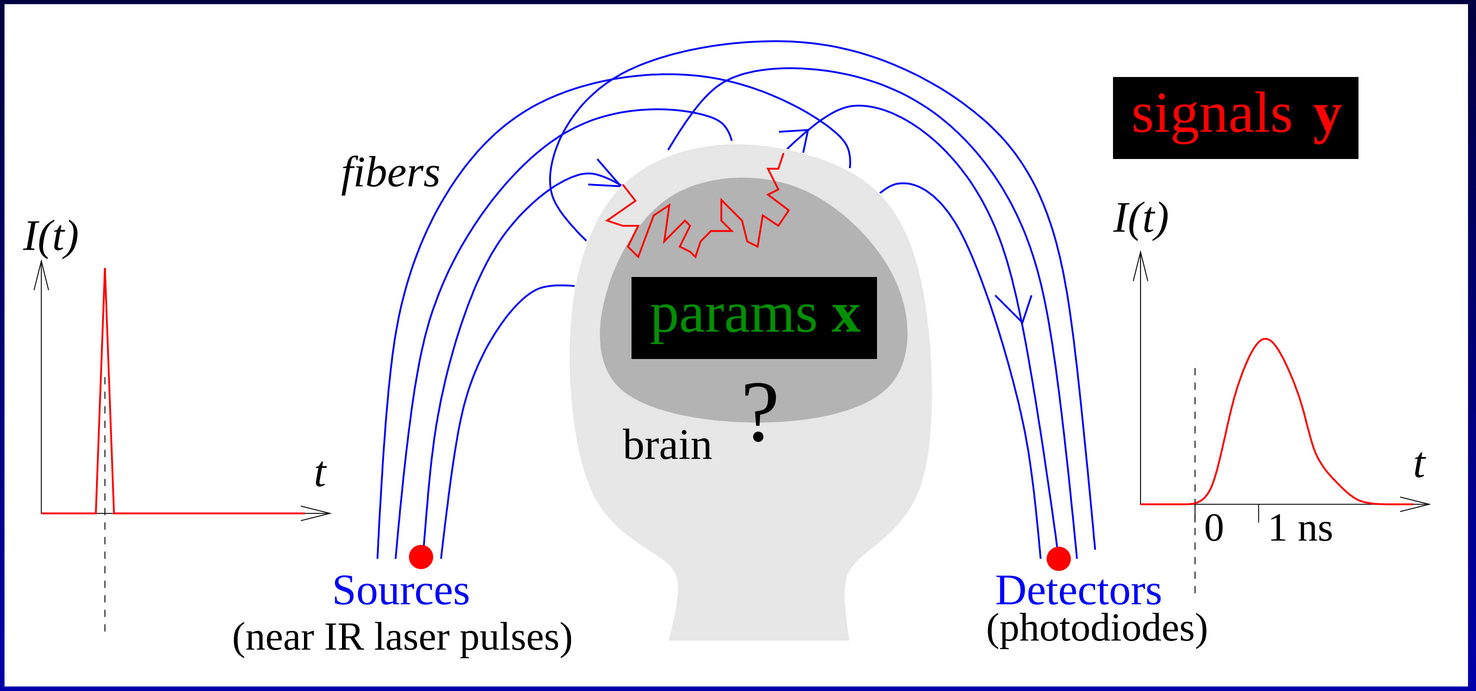
vector **x** = spatial absorption & scattering info

vector **y** = components of measured signals

Inverse problem: given **y** find **x**.

The big picture

Diffuse Optical Tomography (DOT):



vector **x** = spatial absorption & scattering info

vector **y** = components of measured signals

Inverse problem: given **y** find **x**.

Many wavelengths $\longrightarrow$ [HbO_2], [HbR], activation...

Baseline optical measurement

Assuming head tissues optically homogeneous:
How well could we measure their baseline properties?

- *absolute* cortical absorption μ_a
  - cerebral oximetry, neonatal, stroke, trauma...
- required for *quantitative* brain activation studies

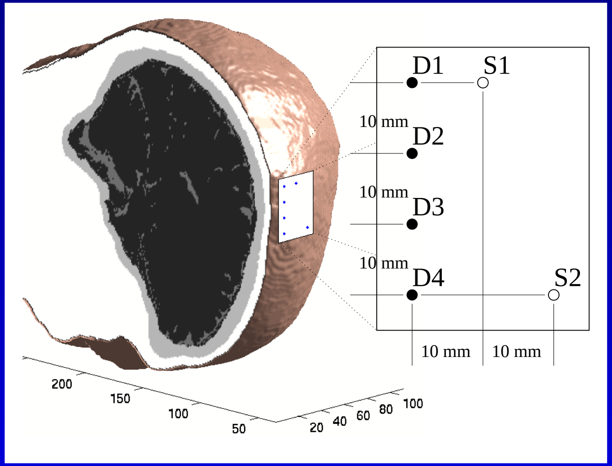
Baseline optical measurement

Assuming head tissues optically homogeneous:
How well could we measure their baseline properties?

- *absolute* cortical absorption μ_a
  - cerebral oximetry, neonatal, stroke, trauma...
- required for *quantitative* brain activation studies

small # unknowns ($N = 6$):
$\mathbf{x} \equiv \{\mu_a, \mu_s'\}$ scalp, skull, brain

- time-resolved DOT
- small system eg 2 Src, 4 Det



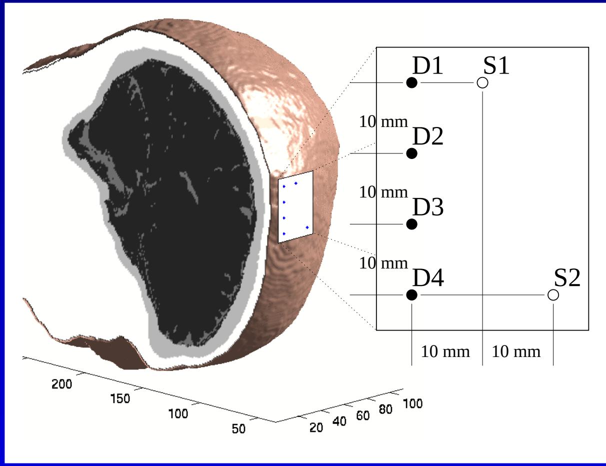
Baseline optical measurement

Assuming head tissues optically homogeneous:
How well could we measure their baseline properties?

- *absolute* cortical absorption μ_a
  - · cerebral oximetry, neonatal, stroke, trauma. . .
- required for *quantitative* brain activation studies

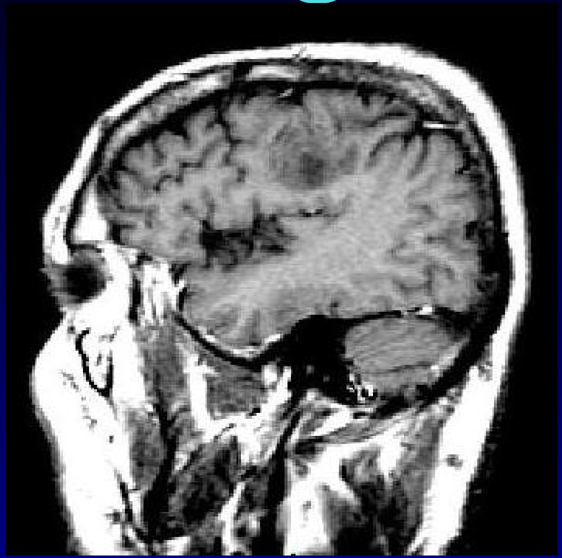
small # unknowns ($N = 6$):
$\mathbf{x} \equiv \{\mu_a, \mu'_s\}$ scalp, skull, brain

- time-resolved DOT
- small system eg 2 Src, 4 Det

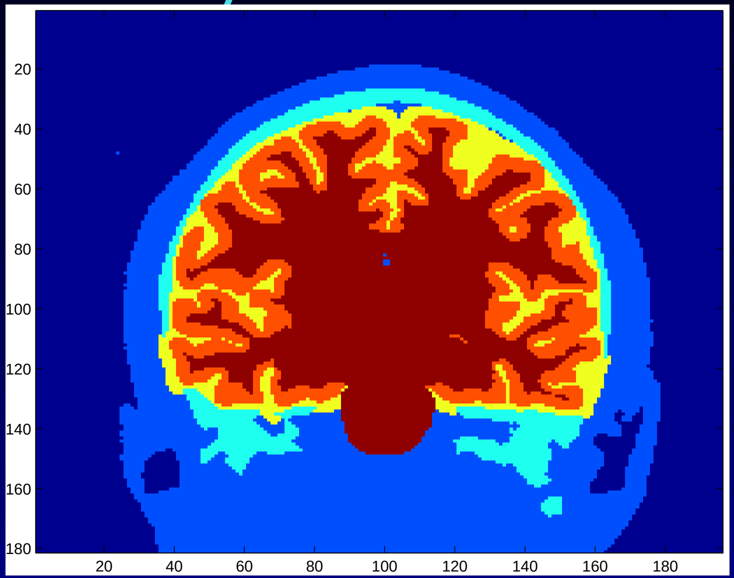


Numerical study: simulated noisy signals $\mathbf{y}$

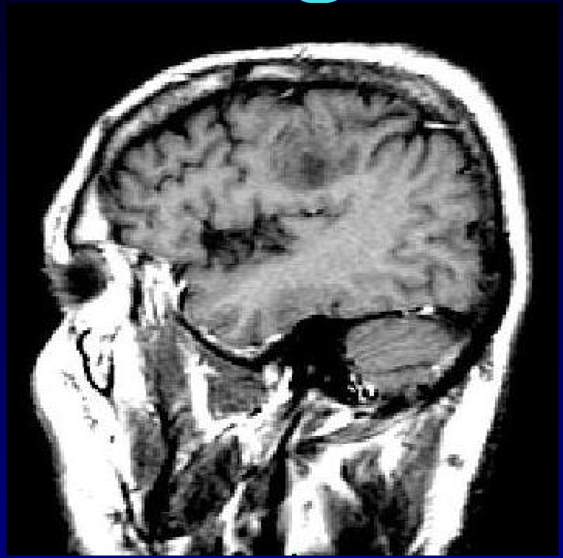
MRI segmented geometry



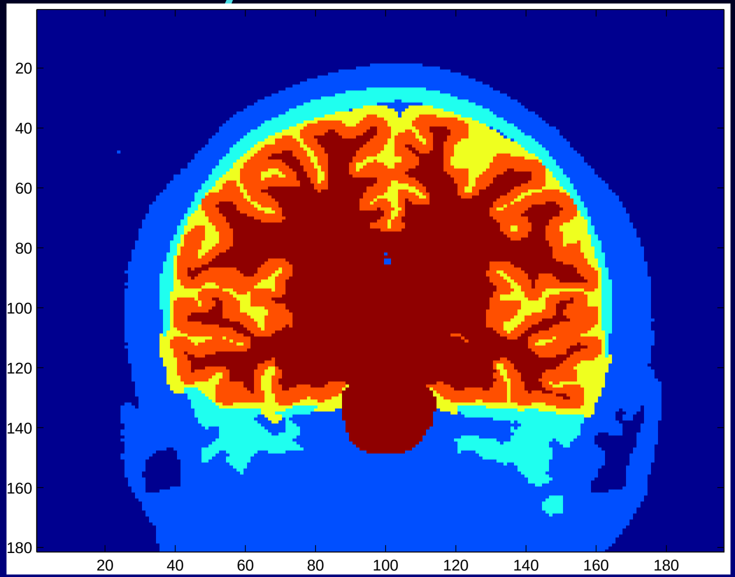
segment
$\longrightarrow$



MRI segmented geometry



segment $\longrightarrow$



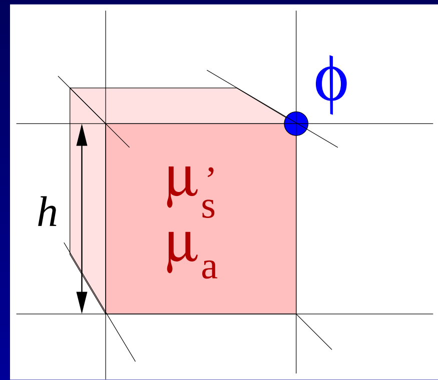
tissue	μ_a (mm^{-1})	μ'_s (mm^{-1})	shape
scalp	0.015	0.8	$\sim$ 7 mm layer
skull	0.01	1.0	$\sim$ 7 mm layer
CSF	0.0004	0.01*	folded 1–3 mm sheet
brain	0.018	1.3	$\sim$ 1 cm folds (sulci)

Much uncertainty. * Diffusion, we use $\mu'_{s,\mathrm{eff}} \sim 0.4$ mm^{-1}.

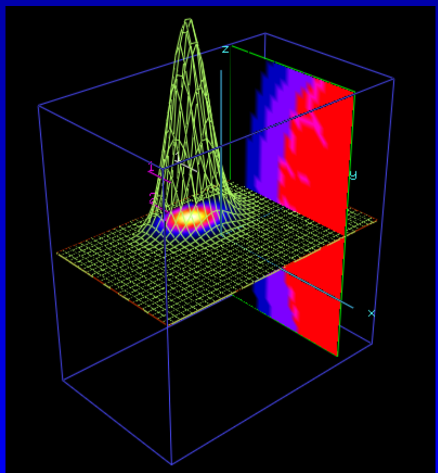
Diffusion forward model

$$\frac{1}{c}\frac{\partial \phi}{\partial t} = \nabla \cdot \left(\frac{1}{3\mu_s'(\mathbf{r})} \nabla \phi \right) - \mu_a(\mathbf{r})\phi + q(\mathbf{r}, t)$$

- Finite difference, lattice size h
- Forward Euler, timestep $\Delta t \sim$ ps
- accuracy $O(h^2)$, typ few % error
- $h = 2$ mm: 4×10^4 cells, 10s CPU



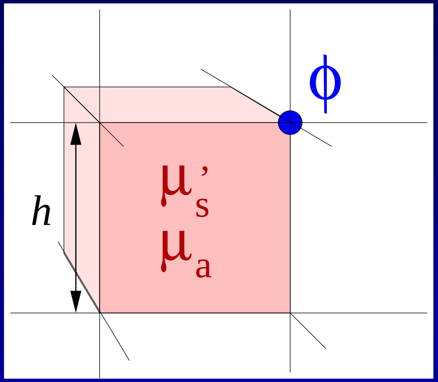
Signals: time-dep fluence at detectors = vector $\mathbf{f}(\mathbf{x})$



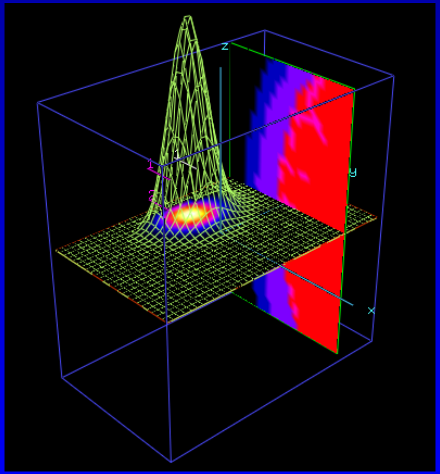
Diffusion forward model

$$\frac{1}{c}\frac{\partial \phi}{\partial t} = \nabla \cdot \left(\frac{1}{3\mu_s'(\mathbf{r})} \nabla \phi \right) - \mu_a(\mathbf{r})\phi + q(\mathbf{r}, t)$$

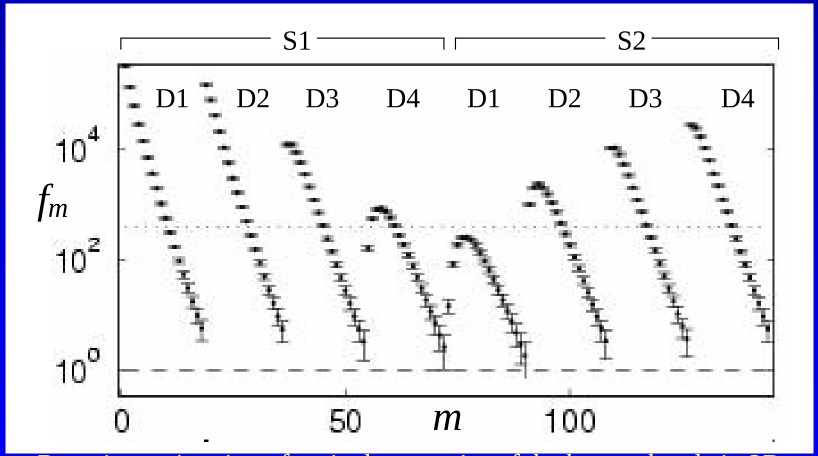
- Finite difference, lattice size h
- Forward Euler, timestep $\Delta t \sim$ ps
- accuracy $O(h^2)$, typ few % error
- $h = 2$ mm: 4×10^4 cells, 10s CPU



Signals: time-dep fluence at detectors = vector $\mathbf{f}(\mathbf{x})$



detect $\longrightarrow$



Bayesian inverse problem

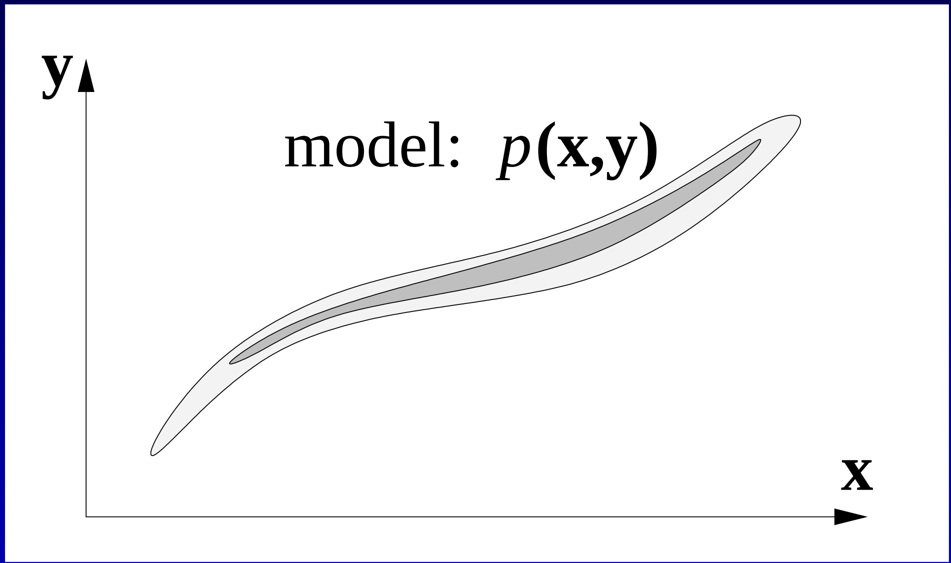
$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\longrightarrow 0$: ‘ill-posed’ (many $\mathbf{x}$ equally valid)

Incomplete info on $\mathbf{x} \longrightarrow$ *probability density function*

Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\to 0$: 'ill-posed' (many $\mathbf{x}$ equally valid)

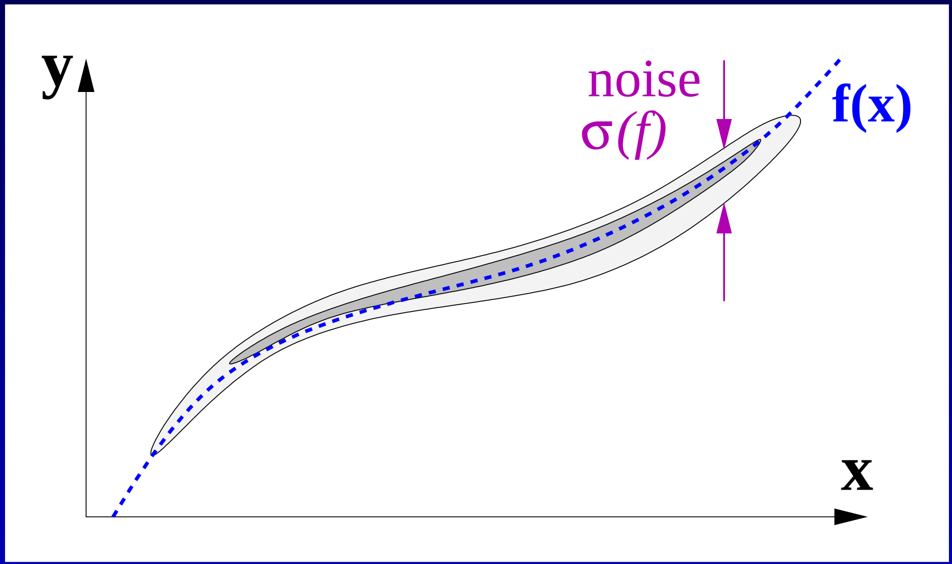
Incomplete info on $\mathbf{x} \to$ *probability density function*



Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\longrightarrow 0$: ‘ill-posed’ (many $\mathbf{x}$ equally valid)

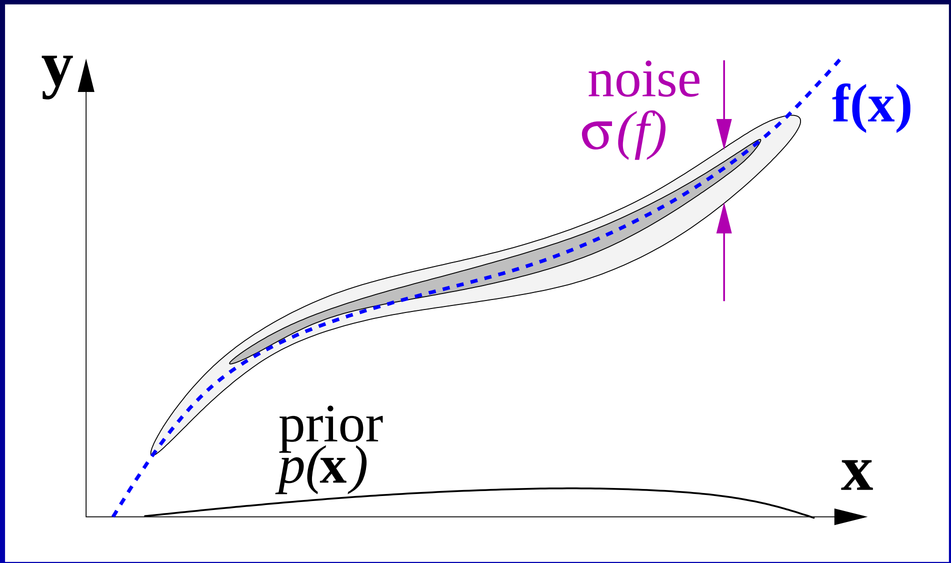
Incomplete info on $\mathbf{x} \longrightarrow$ *probability density function*



Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\longrightarrow 0$: 'ill-posed' (many $\mathbf{x}$ equally valid)

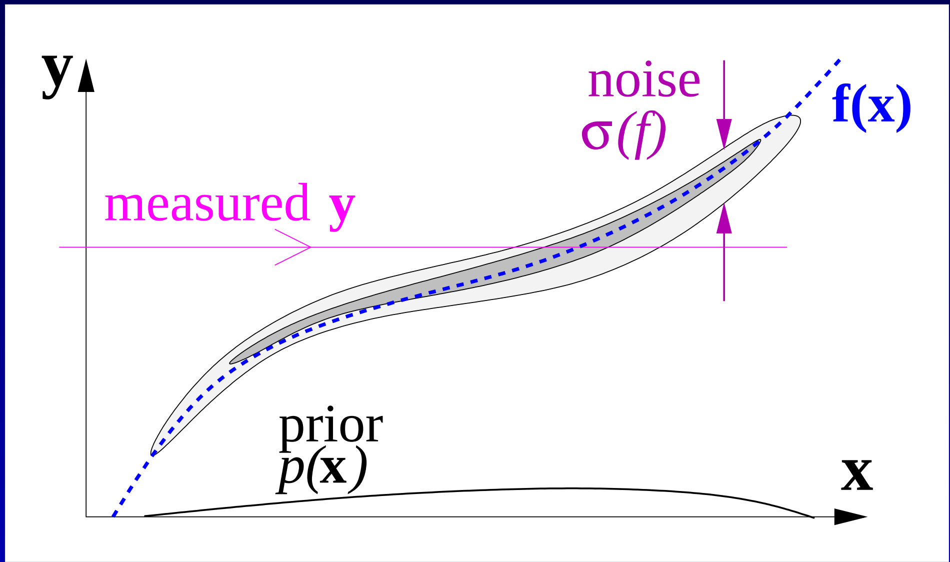
Incomplete info on $\mathbf{x} \longrightarrow$ *probability density function*



Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\longrightarrow 0$: 'ill-posed' (many $\mathbf{x}$ equally valid)

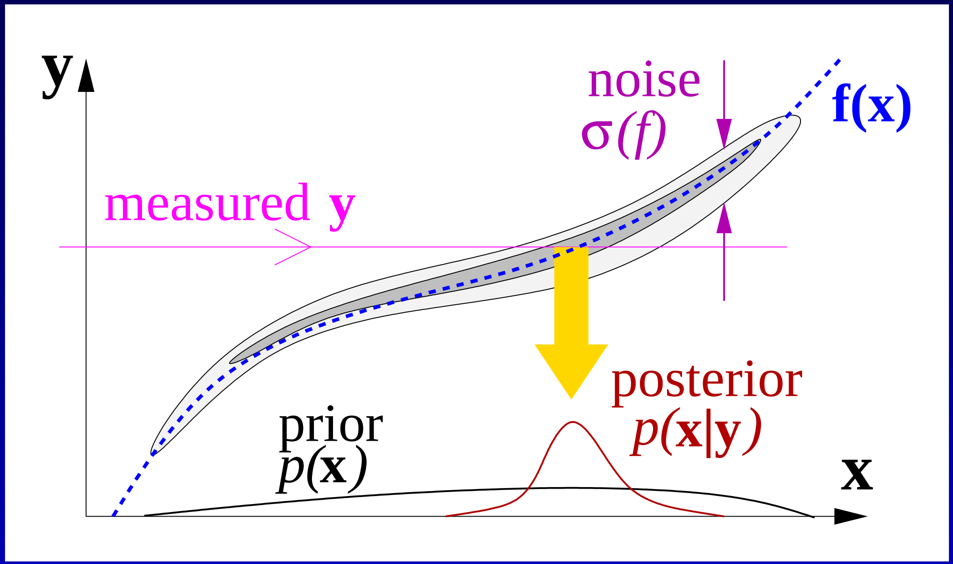
Incomplete info on $\mathbf{x} \longrightarrow$ *probability density function*



Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\rightarrow 0$: ‘ill-posed’ (many $\mathbf{x}$ equally valid)

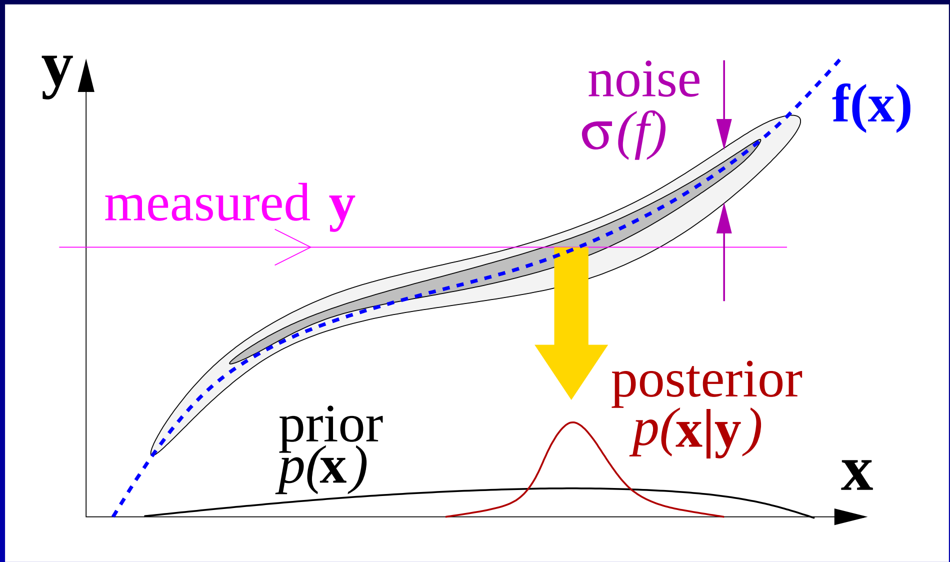
Incomplete info on $\mathbf{x}$ $\rightarrow$ *probability density function*



Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\longrightarrow 0$: 'ill-posed' (many $\mathbf{x}$ equally valid)

Incomplete info on $\mathbf{x} \longrightarrow$ *probability density function*



Bayesian inference

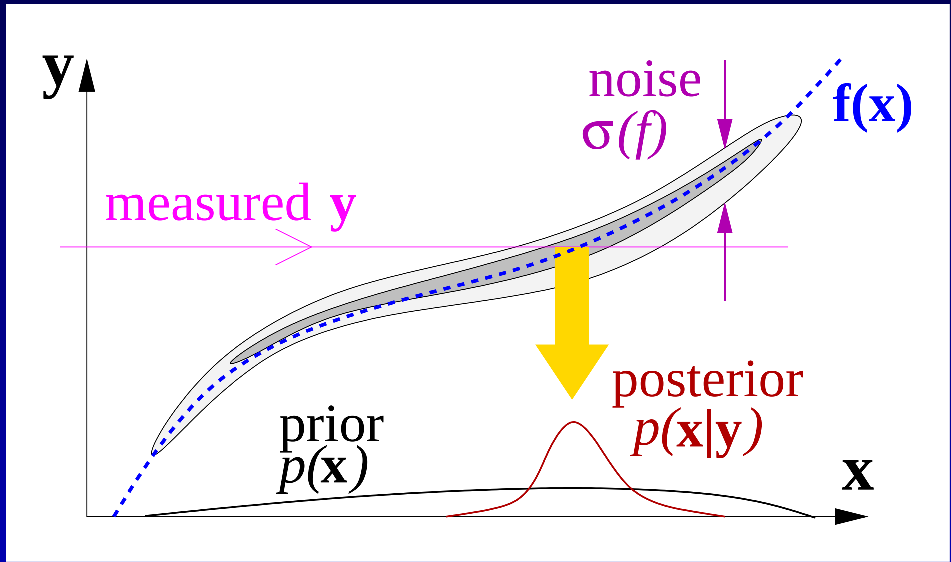
$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &\propto p(\mathbf{x}, \mathbf{y}) \\ &= p(\mathbf{y}|\mathbf{x}) \cdot p(\mathbf{x}) \end{aligned}$$

posterior likelihood prior

Bayesian inverse problem

$\frac{\partial f_m}{\partial x_n}$ sing. vals. $\rightarrow 0$: 'ill-posed' (many $\mathbf{x}$ equally valid)

Incomplete info on $\mathbf{x} \rightarrow$ *probability density function*



Bayesian inference

$$
\begin{aligned}
p(\mathbf{x}|\mathbf{y}) &\propto p(\mathbf{x},\mathbf{y}) \\
&= p(\mathbf{y}|\mathbf{x}) \cdot p(\mathbf{x})
\end{aligned}
$$

posterior likelihood prior

- assumptions about noise $\rightarrow$ width of likelihood
- Embraces ill-posedness, statistically rigorous.
- Need to explore N-dim posterior: many $\mathbf{f}(\mathbf{x})$ evals required ($> 10^2$).

Why full PDF?

Common “Bayesian” method for DOT inversion: find single best-fit $\mathbf{x} = \mathbf{x}_{\mathrm{MAP}}$. (MAP = *maximum a-posteriori*)

Why full PDF?

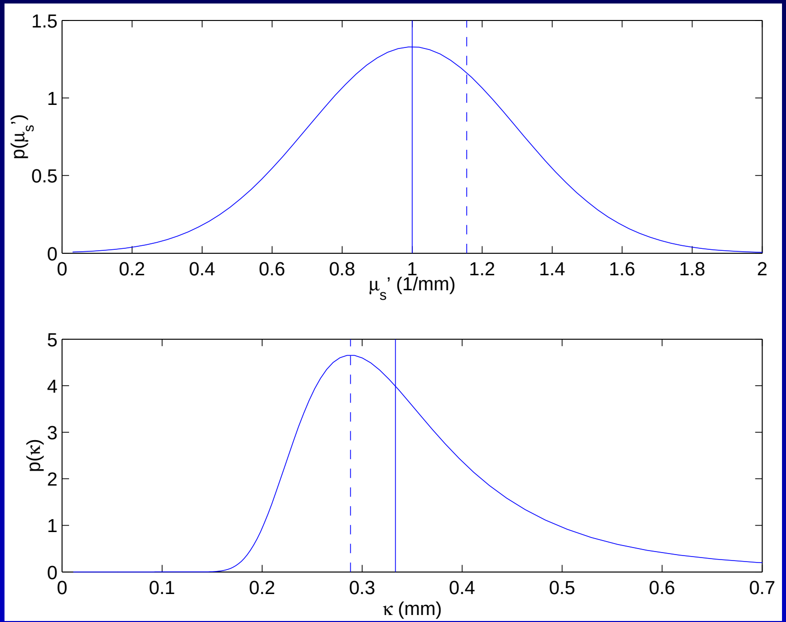
Common “Bayesian” method for DOT inversion: find single best-fit $\mathbf{x} = \mathbf{x}_{\mathrm{MAP}}$. (MAP = *maximum a-posteriori*)

However $\mathbf{x}_{\mathrm{MAP}}$ can be moved by reparametrization!

Why full PDF?

Common "Bayesian" method for DOT inversion: find single best-fit $\mathbf{x} = \mathbf{x}_{\mathrm{MAP}}$. (MAP = *maximum a-posteriori*)

However $\mathbf{x}_{\mathrm{MAP}}$ can be moved by reparametrization!



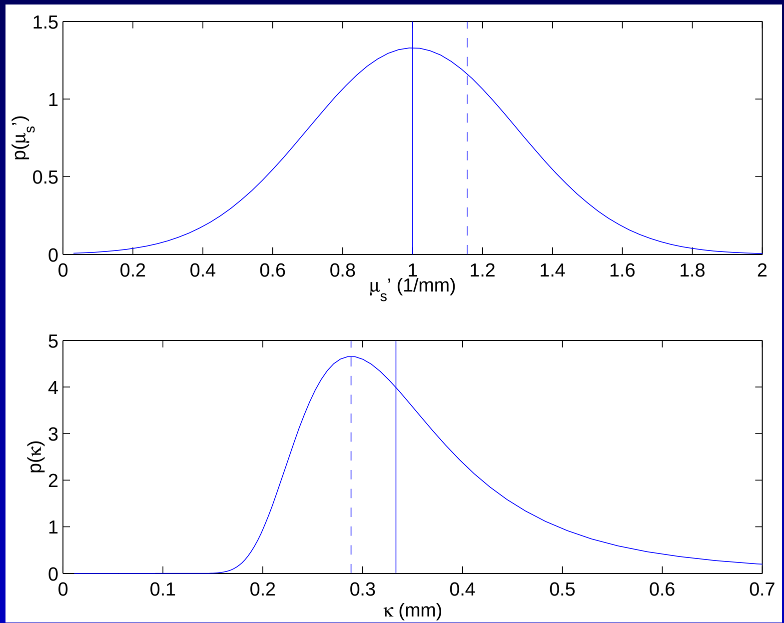
Example: broad PDF on μ_s' Normal, $\sigma = 30\%$ of mean.

16% shift in MAP between $p(\mu_s')$ and $p(\mu_s'^{-1})$

Why full PDF?

Common “Bayesian” method for DOT inversion: find single best-fit $\mathbf{x} = \mathbf{x}_{\text{MAP}}$. (MAP = *maximum a-posteriori*)

However $\mathbf{x}_{\text{MAP}}$ can be moved by reparametrization!



Example: broad PDF on μ_s' Normal, $\sigma = 30\%$ of mean.

16% shift in MAP between $p(\mu_s')$ and $p(\mu_s'^{-1})$

- CPUs advance faster than DOT instrumentation $\longrightarrow$ best to make maximal use of data
- Statistical answers $\longrightarrow$ multimodal imaging.

Realistic new noise model

Each signal component $f_m(\mathbf{x})$ independent noise.

Photons Poissonian: gaussian approx $\sigma(f) = f^{1/2}$

E.g. 10^6 photons = 0.1% relative error

But: we do not trust forward model to 0.1% !

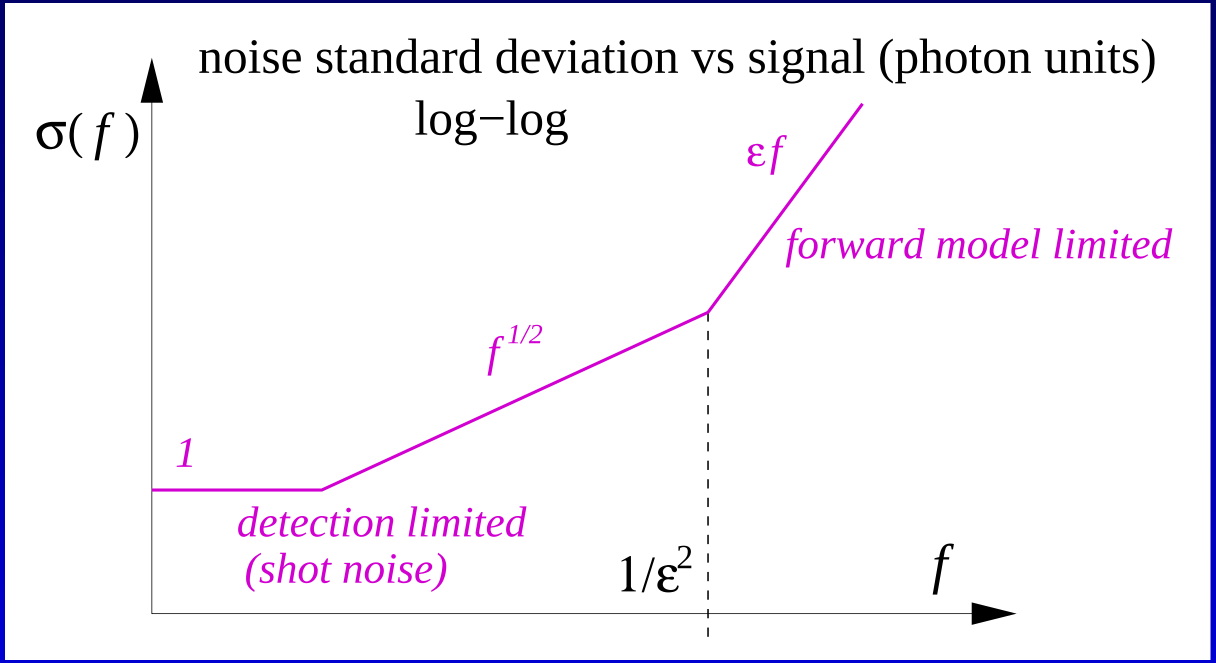
Realistic new noise model

Each signal component $f_m(\mathbf{x})$ independent noise.

Photons Poissonian: gaussian approx $\sigma(f) = f^{1/2}$

E.g. 10^6 photons = 0.1% relative error

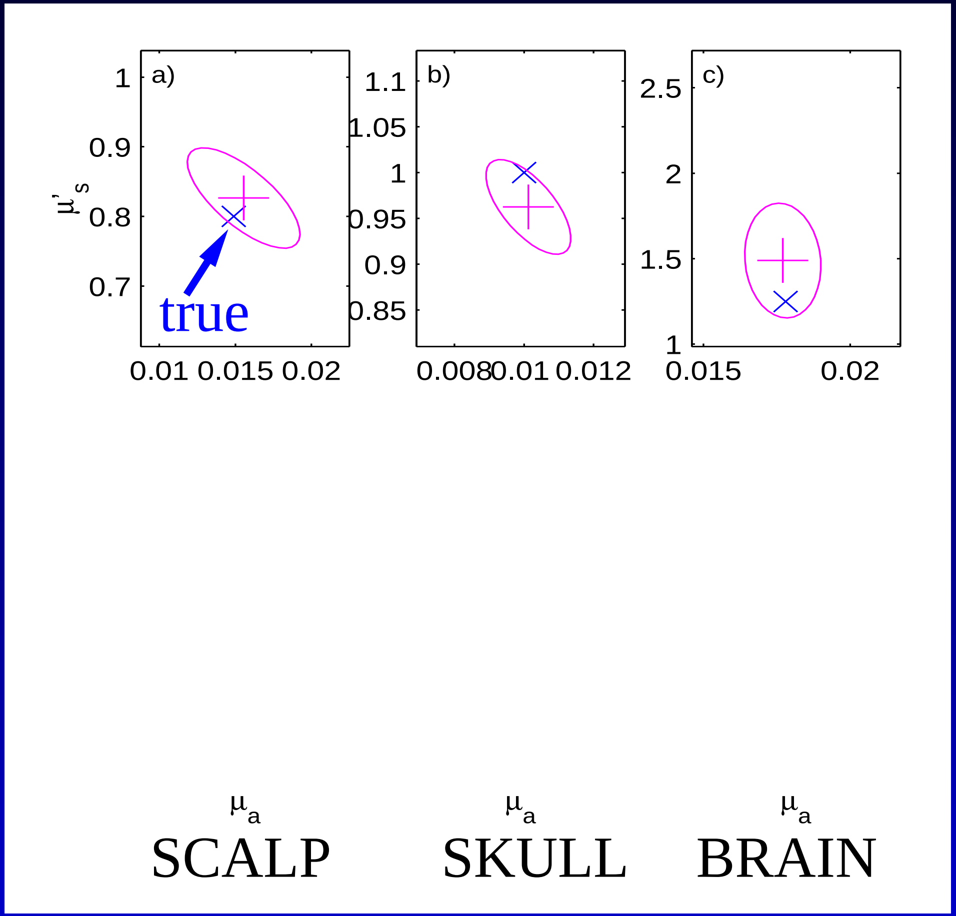
But: we do not trust forward model to 0.1% !



- $\varepsilon =$ relative forward model error *e.g.* 1–5% (errors: physics, numerical. . .)

Result: marginal posterior PDF

Applied Optics, special issue on biomedical optics, June 2003.

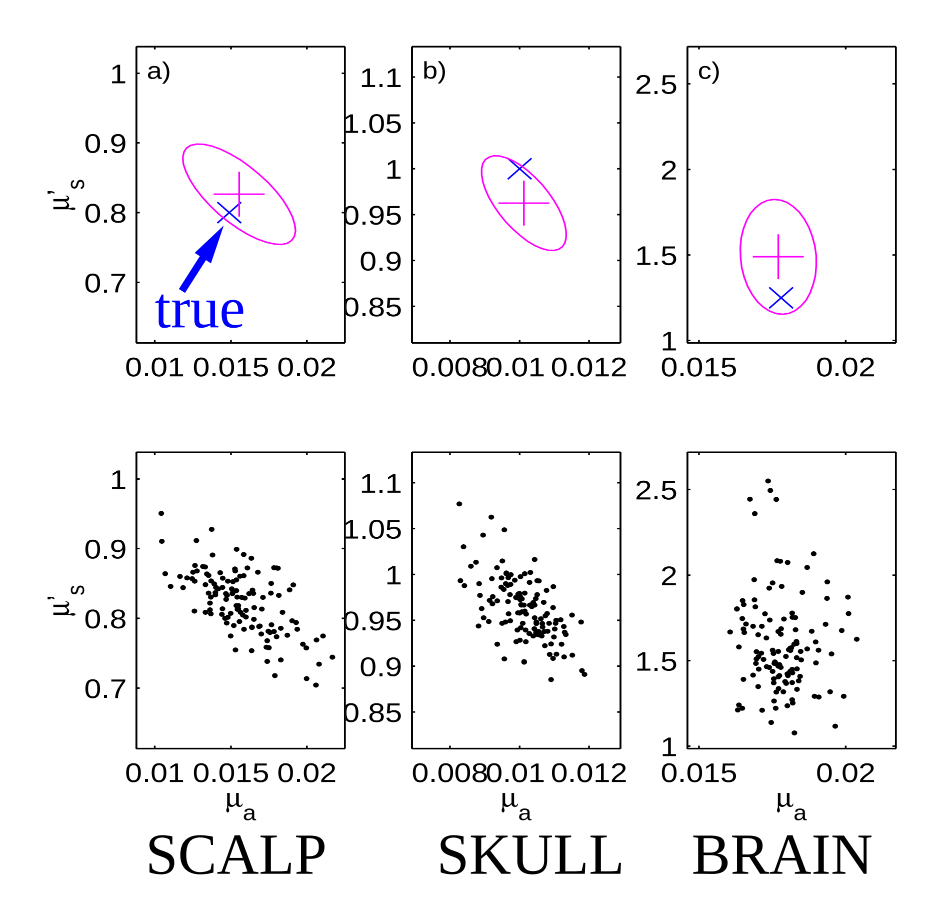


Gaussian PDF approx

via Levenburg-Marquardt, ellipse encloses 65% prob

Result: marginal posterior PDF

Applied Optics, special issue on biomedical optics, June 2003.



Gaussian PDF approx

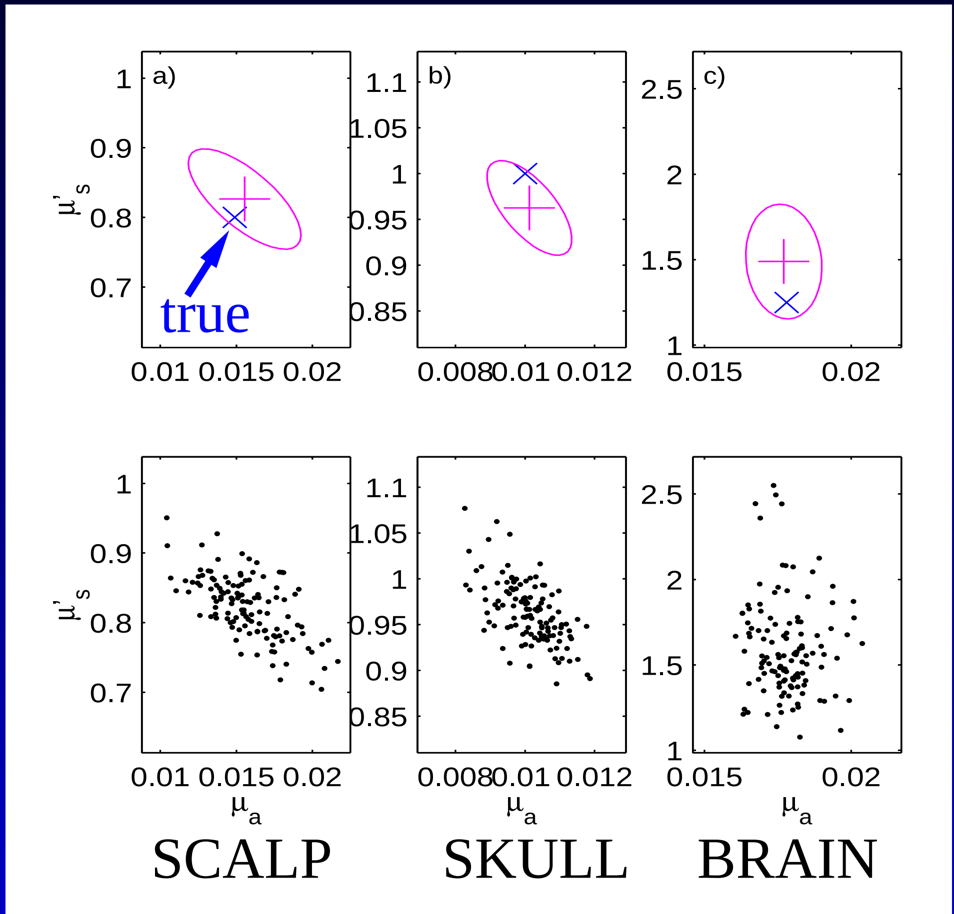
via Levenburg-Marquardt, ellipse encloses 65% prob

Sampling exact PDF

Markov chain Monte Carlo $\sim 40\ \mathbf{f}(\mathbf{x})$ evals per sample (CPU intensive)

Result: marginal posterior PDF

Applied Optics, special issue on biomedical optics, June 2003.



Gaussian PDF approx

via Levenburg-Marquardt, ellipse encloses 65% prob

Sampling exact PDF

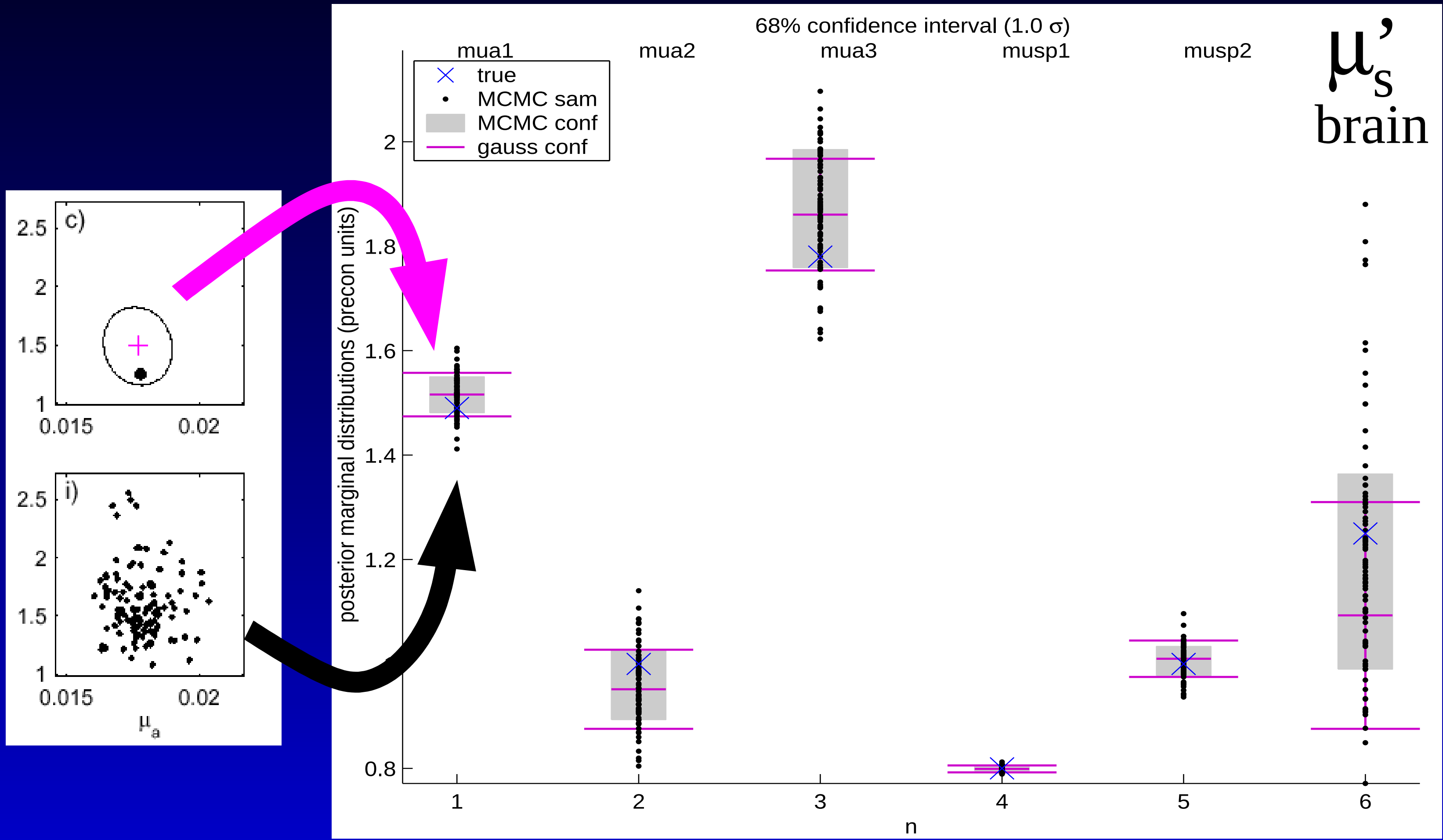
Markov chain Monte Carlo

~ 40 $\mathbf{f}(\mathbf{x})$ evals per sample

(CPU intensive)

Pancake-like PDF: major-to-minor axis ratio $\sim 10^2$

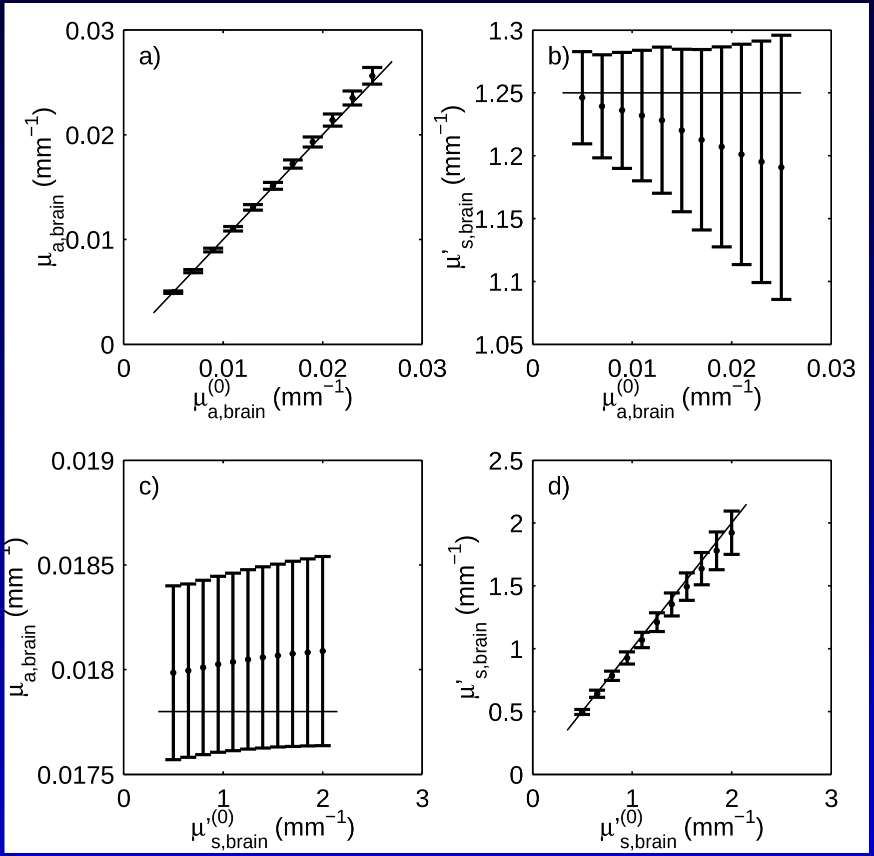
Results: confidence intervals



$\mu'_{s,\mathrm{brain}}$: Gaussian approx bad, need MCMC sampling

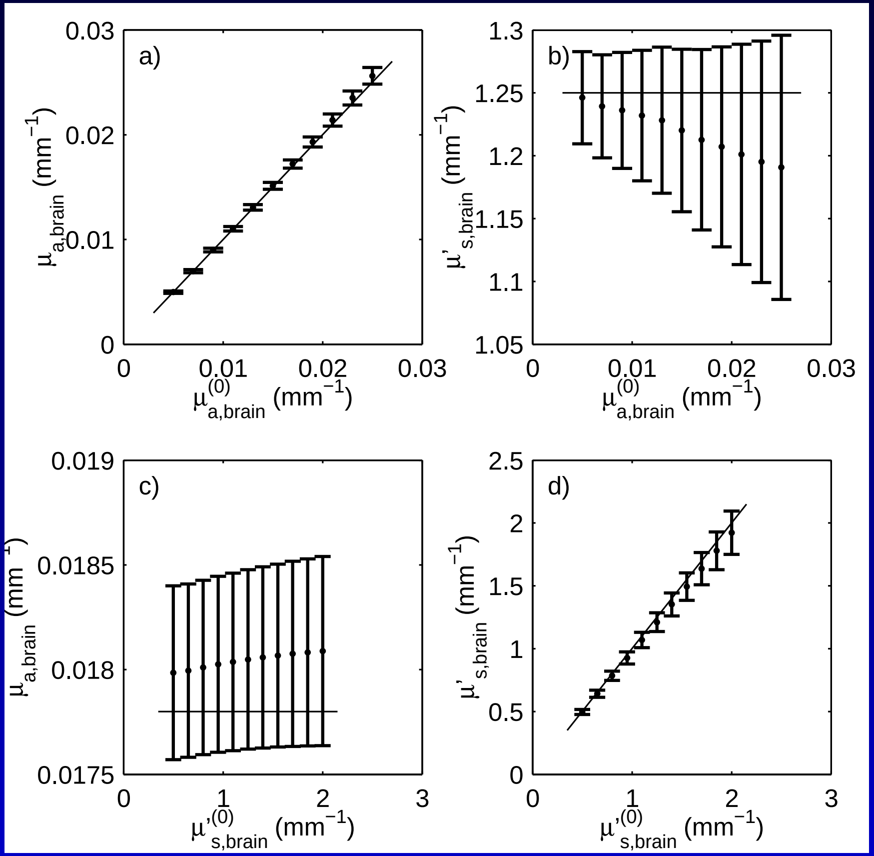
Tracking brain parameters

measured tracks true optical params, little crosstalk



Tracking brain parameters

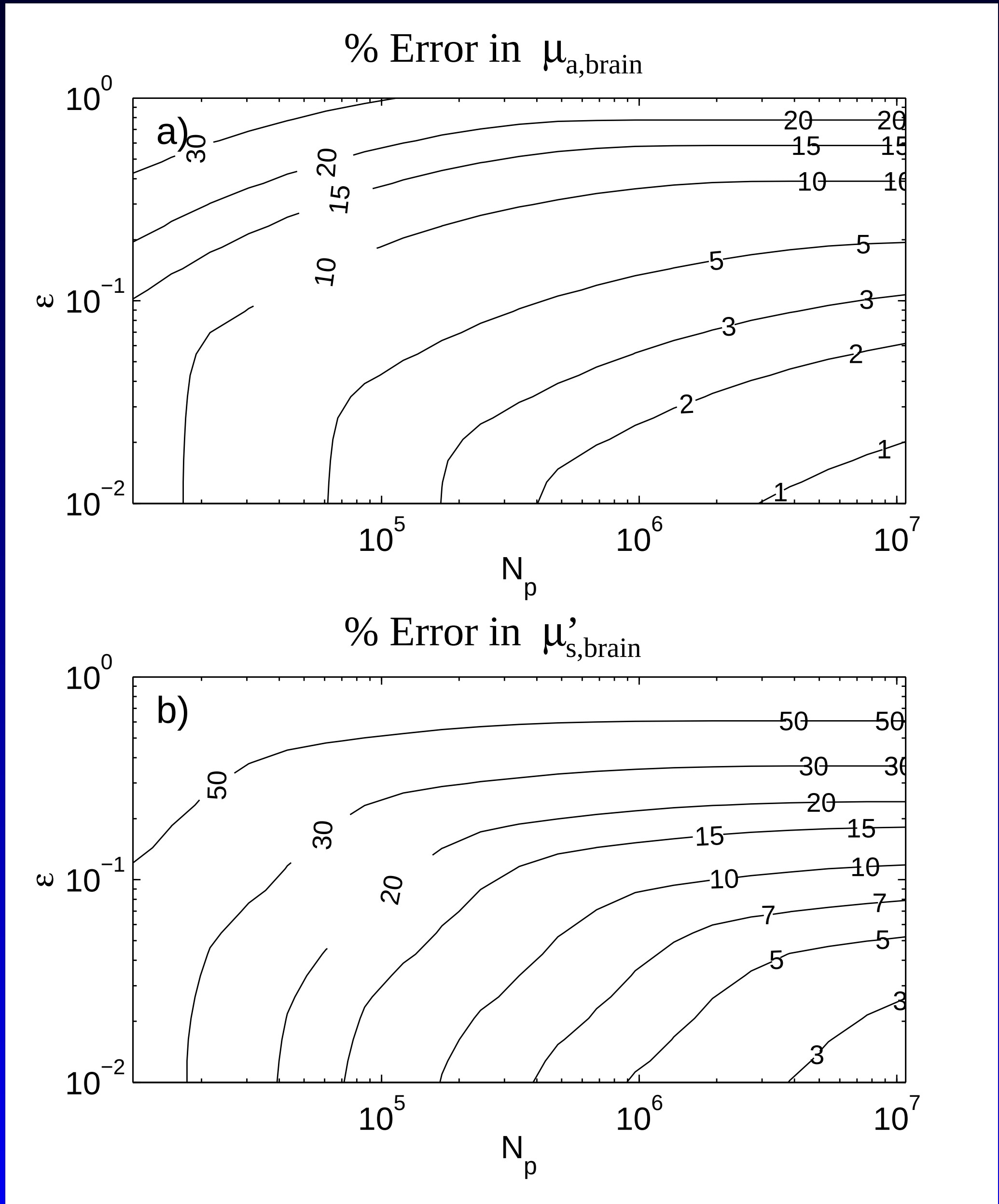
measured tracks true optical params, little crosstalk



2×10^6 det photons: errorbars 3% $\mu_{a,\text{brain}}$, 10% $\mu'_{s,\text{brain}}$

for $\varepsilon = 3\%$, flat prior

How many photons needed?



Baseline brain errorbars vs…

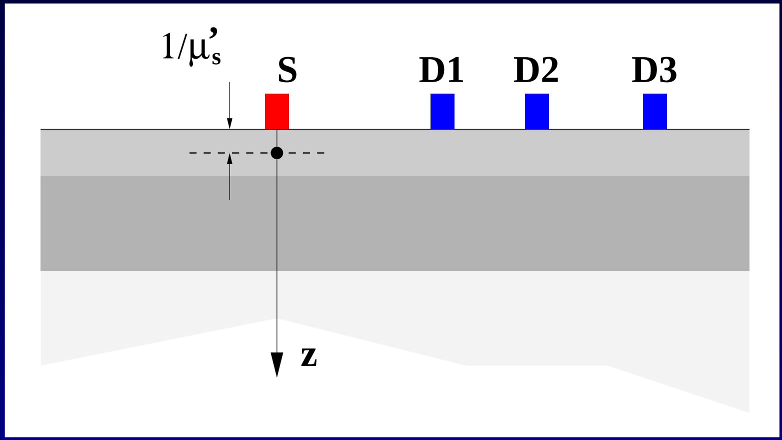
N_p = total det photons

ϵ = fwd model accuracy

Can optimize design of DOT apparatus

Partial geometric info

multilayer slab diffusion forward model:



Crank-Nicholson

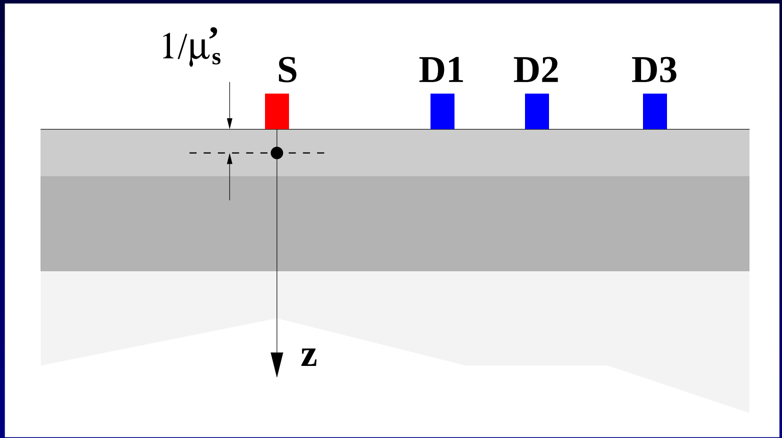
novel logarithmic timesteps

<1% error

Fast! typ. 0.2 sec CPU

Partial geometric info

multilayer slab diffusion forward model:



Crank-Nicholson

novel logarithmic timesteps

<1% error

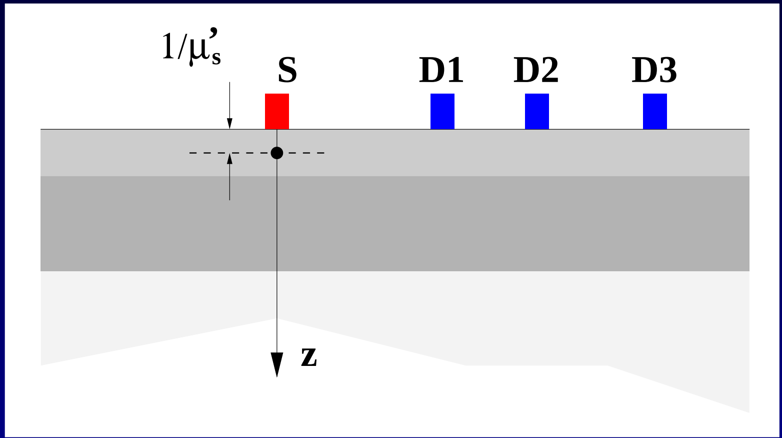
Fast! typ. 0.2 sec CPU

Preliminary results. . .

- slab fits MRI head signals well (if no CSF)

Partial geometric info

multilayer slab diffusion forward model:



Crank-Nicholson

novel logarithmic timesteps

<1% error

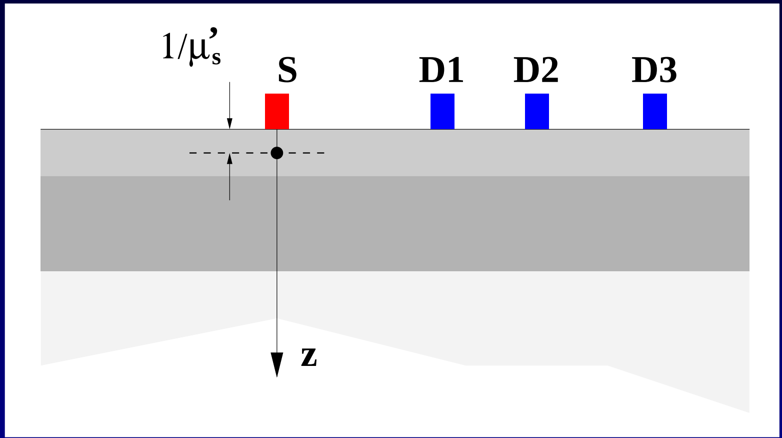
Fast! typ. 0.2 sec CPU

Preliminary results. . .

- slab fits MRI head signals well (if no CSF)
- can also fit for unknown scalp+skull thickness
  but $\mu_{a,\mathrm{brain}}$ is $3\times$ more accurate if MRI geom used

Partial geometric info

multilayer slab diffusion forward model:



Crank-Nicholson

novel logarithmic timesteps

<1% error

Fast! typ. 0.2 sec CPU

Preliminary results. . .

- slab fits MRI head signals well (if no CSF)
- can also fit for unknown scalp+skull thickness
  but $\mu_{a,\mathrm{brain}}$ is $3\times$ more accurate if MRI geom used

Optode amplitude & time-offset calibrations work
(Bayes $\longrightarrow$ marginalize over the ‘nuisance’ params)

Conclusions

- Optical tissue model from structural MRI
  - can infer baseline tissue parameters
  - errors $< 10\%$ $\mu_{a,\text{brain}}$, $\mu'_{s,\text{brain}}$, no crosstalk
  - noise model: shot noise + fwd model error

Conclusions

- Optical tissue model from structural MRI
  - can infer baseline tissue parameters
  - errors $< 10\%$ $\mu_{a,\text{brain}}$, $\mu'_{s,\text{brain}}$, no crosstalk
  - noise model: shot noise + fwd model error
- Forward models in complex geometry
  - need to be fast $\longrightarrow$ will call many times
  - known errors $\longrightarrow$ include in noise model

Conclusions

- Optical tissue model from structural MRI
  - · can infer baseline tissue parameters
  - · errors $< 10\%$ $\mu_{a,\text{brain}}$, $\mu'_{s,\text{brain}}$, no crosstalk
  - · noise model: shot noise + fwd model error
- Forward models in complex geometry
  - · need to be fast $\longrightarrow$ will call many times
  - · known errors $\longrightarrow$ include in noise model
- Bayes: understand full PDF on unknowns
  - · predict all errorbars, correlations
  - · CPU intensive but optimal use of data
  - · handle calibration (nuisance) params