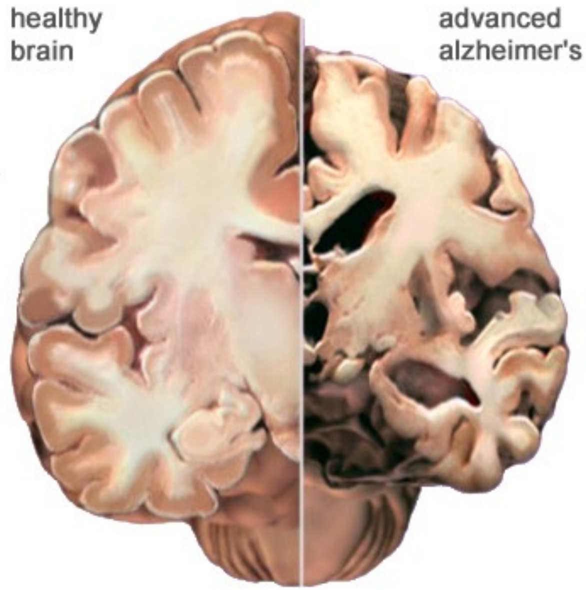
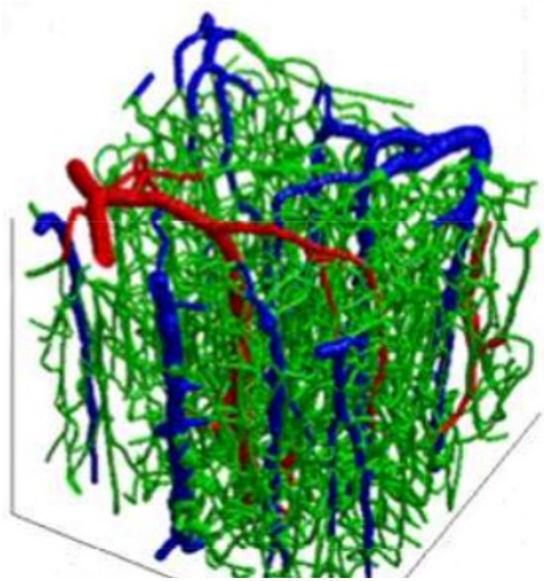
Variational formulations of the strong form for forward and inverse problems - applications to neural nets and isogeometric analysis

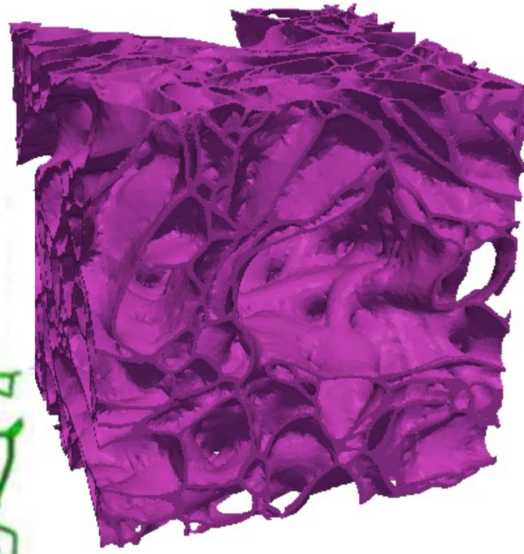
Kent-Andre Mardal

University of Oslo /

Simula Research Laboratory







CBMS Conference: Deep Learning and Numerical PDEs

Outline

- Physiology of sleep

(including some current controversies)

- Math of sleep (?!?):

Find u, f such that

$$\|-\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} \to \min$$

A basic (inverse) problem

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} \to \min$$

- u, f are the unknown state and control
- u_d is data
- Γ is a subdomain of Ω
- We notice increased regularity $u \in H^2$, $f \in L^2$

A basic question is: Can the increased regularity compensate for only partial observation?

Why do we sleep?

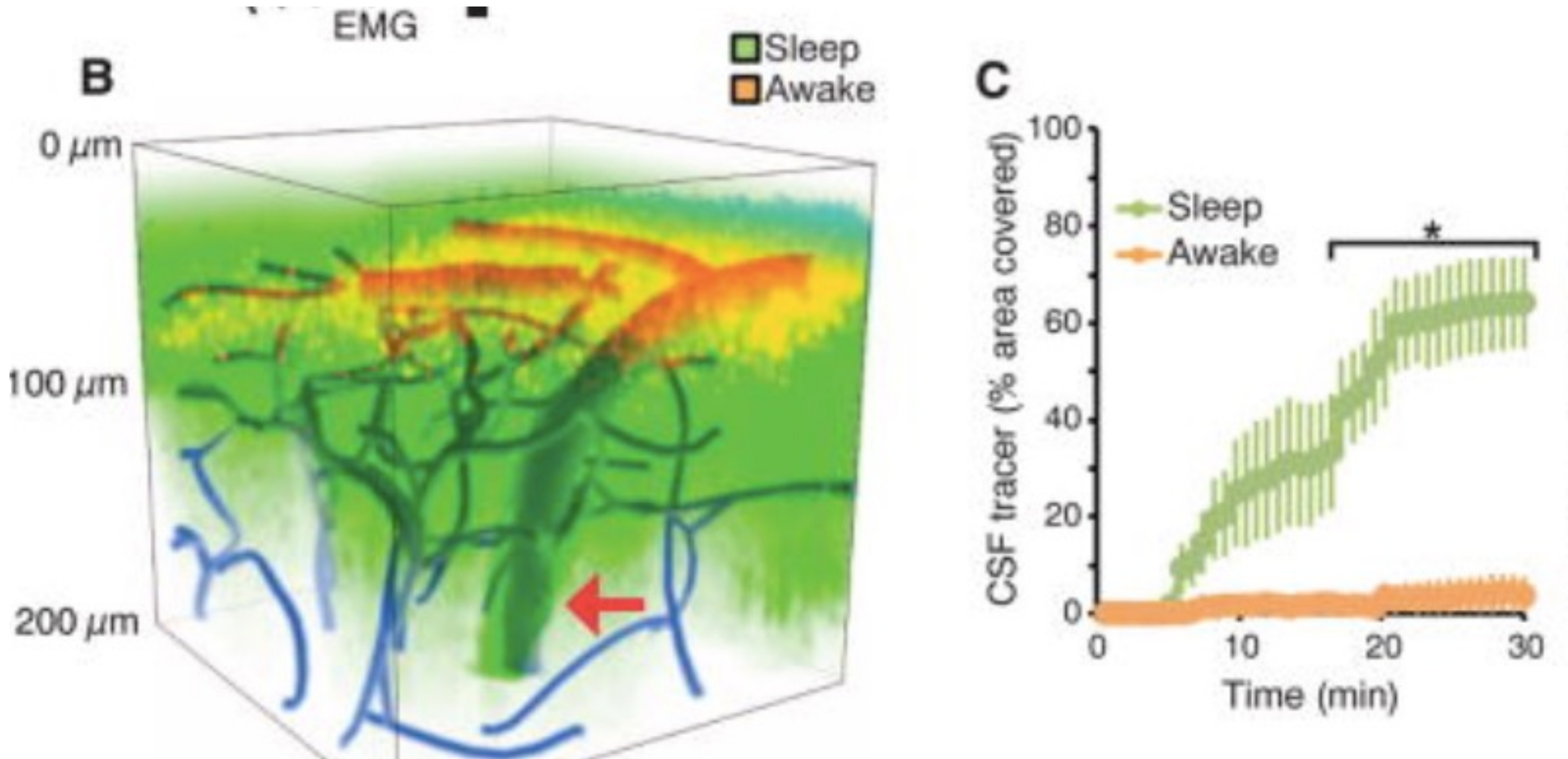
- All animals sleep; ants, jellyfish, mammals, …
- Sleep/circadian rhythm can be identified in individual cells
- From an evolutionary perspective, there is a need to save energy and a circadian rhythm makes sense
- However, *it is hard to explain why we become unconscious* during sleep – it is an obvious disadvantage from an evolutionary perspective
- Sleep is absolutely required. We die without
- Creative solutions like sleeping with half of the brain is common in nature

Why do we sleep?

- All animals sleep; ants, jellyfish, mammals, …
- Sleep/circadian rhythm can be identified in individual cells
- From an evolutionary perspective, there is a need to save energy and a circadian rhythm makes sense
- However, *it is hard to explain why we become unconscious during sleep* – it is an obvious disadvantage from an evolutionary perspective
- Sleep is absolutely required. We die without
- Creative solutions like sleeping with half of the brain is common in nature

All kinds of brain diseases relates to sleeping disorders

The glymphatic system is hyperactive during sleep because the extracellular volume increases



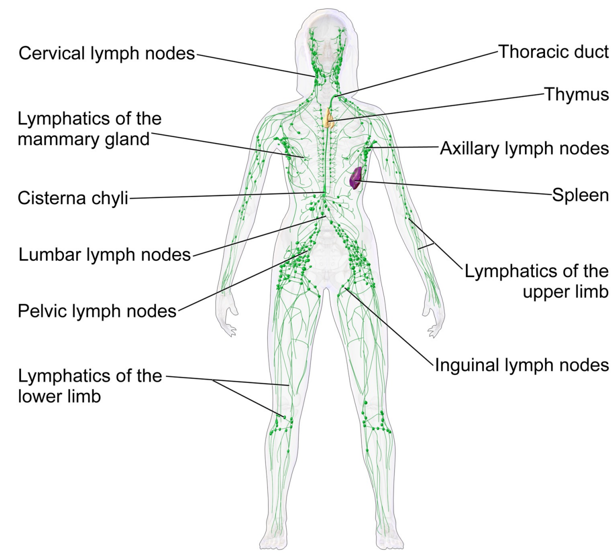
Xie, Lulu, et al. "Sleep drives metabolite clearance from the adult brain." Science 2013. (mouse study)

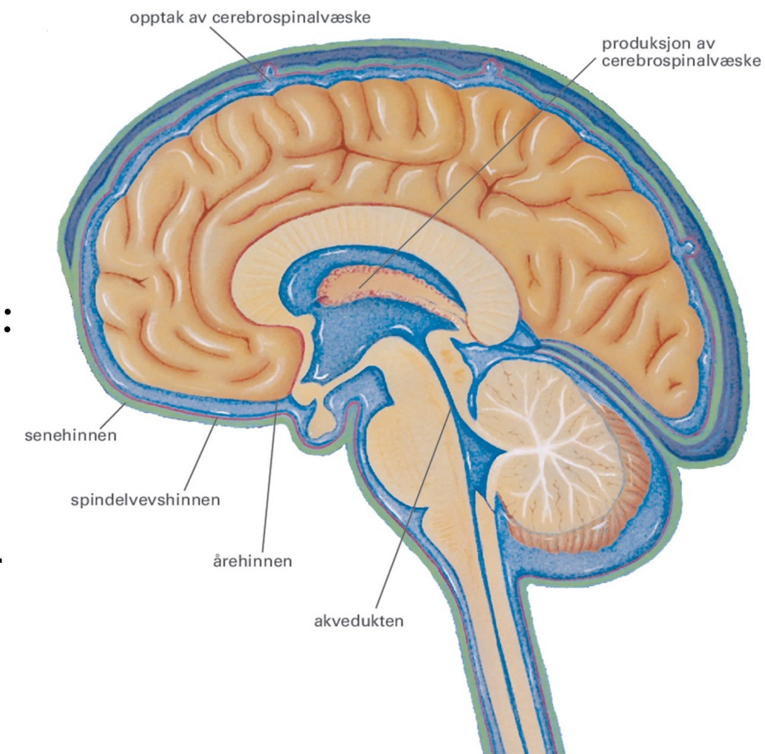
3 kDa Texas Red Dextran typically penetrated 100-200 µm in about 20 minutes

Extracellular volume fraction is 21% during sleep and 14% while awake

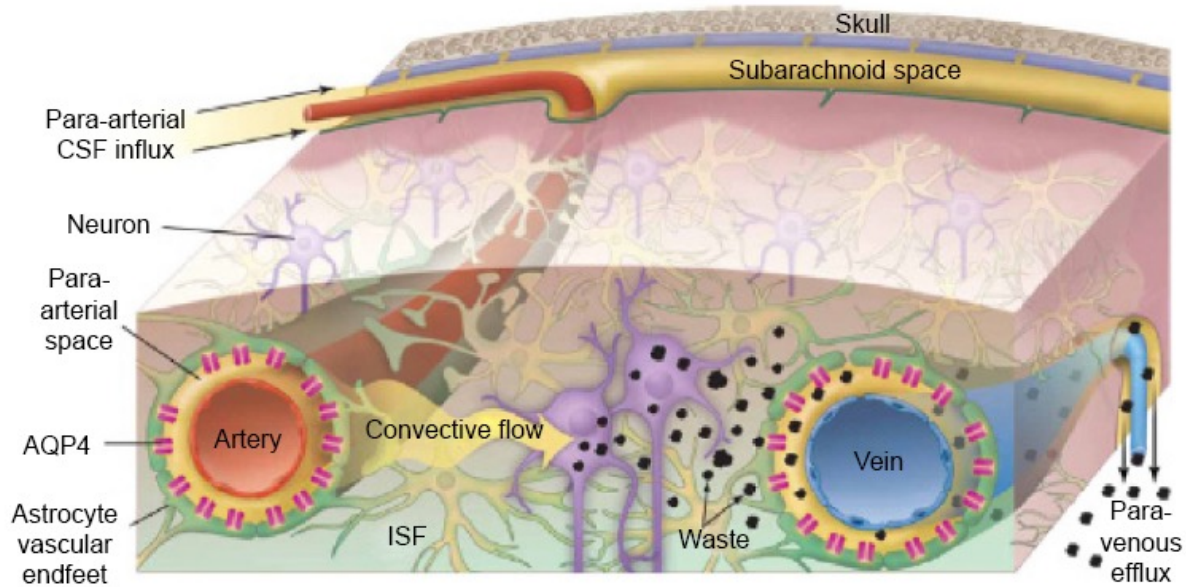
Basic facts, questions about the brain's metabolism:

- The brain occupies **1-2% of the body** in volume / weight
- The brain consumes **around 10-20%** of the body's energy, oxygen
- Elsewhere in the body, the lymphatic system plays a central role in the disposal of waste
- The lymphatic system drains from the extra-cellular matrix
- The **brain does not have a lymph** system
- This observations begs the question: how does the brain clear waste without lymphs?
- Comment: The brain is special because it is bathed in water (cerebrospinal fluid). Is this a part of an accelerated "lymph" system?





Glymphatic system: the garbage truck of the brain

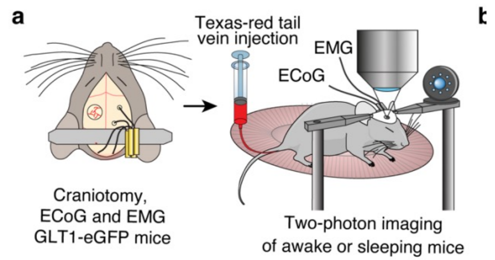


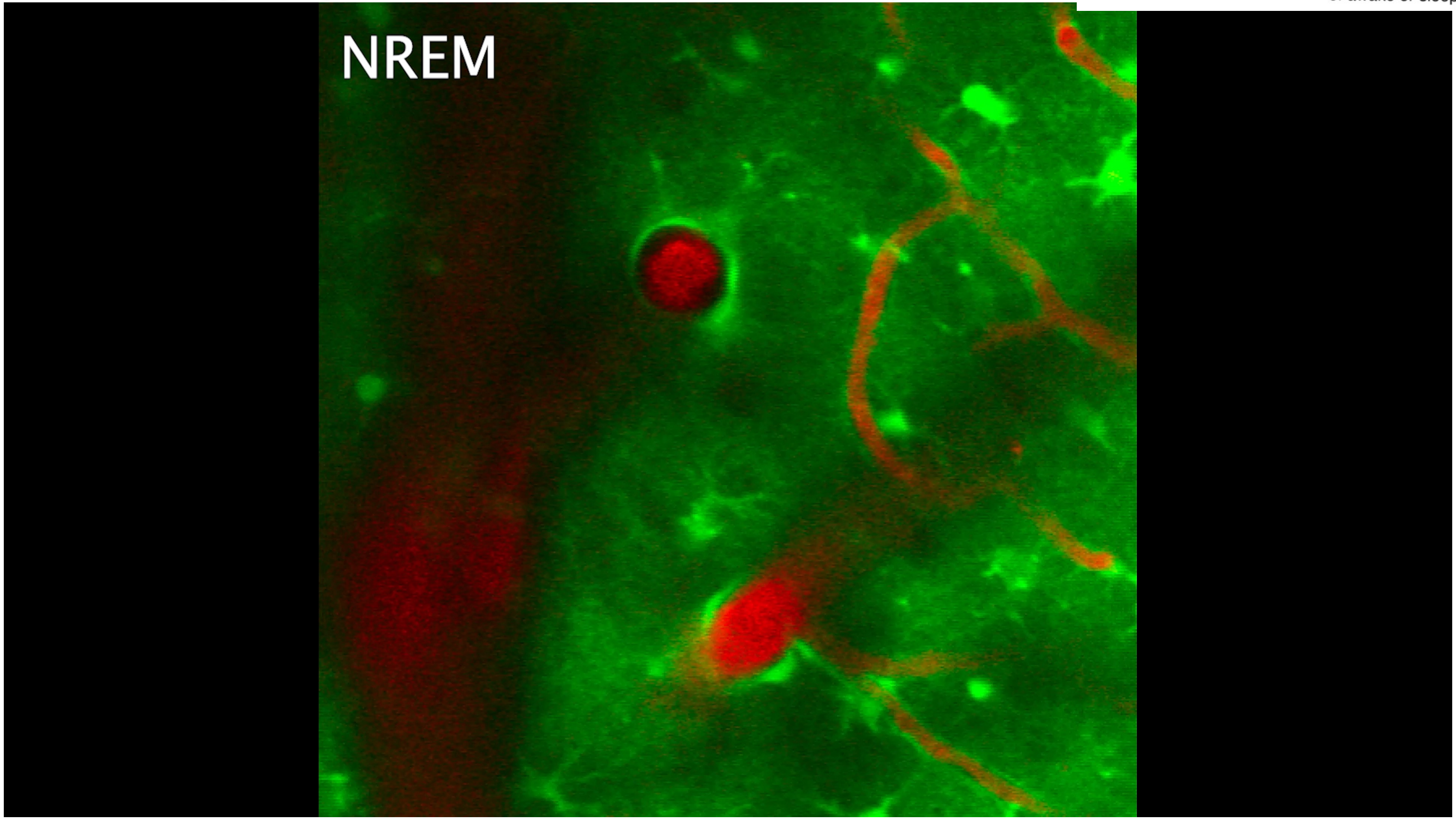
The new pathway: 3 components

- 1. the perivascular space that surrounds the arteries/arterioles are connected with the CSF that surrounds the brain. This space facilitate a bulk flow (viscous flow).
- 2. the hydrostatic pressure gradient between the arterial and venous sites facilitate a bulk flow through the interstitium (porous flow)
- 3. the waste is then removed on the venous site (viscous flow)

Nedergaard M. Garbage truck of the brain. Science. 2013

Sleep (micron-scale)

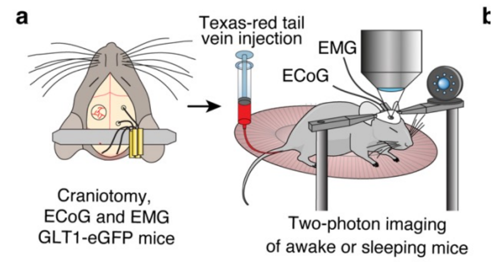


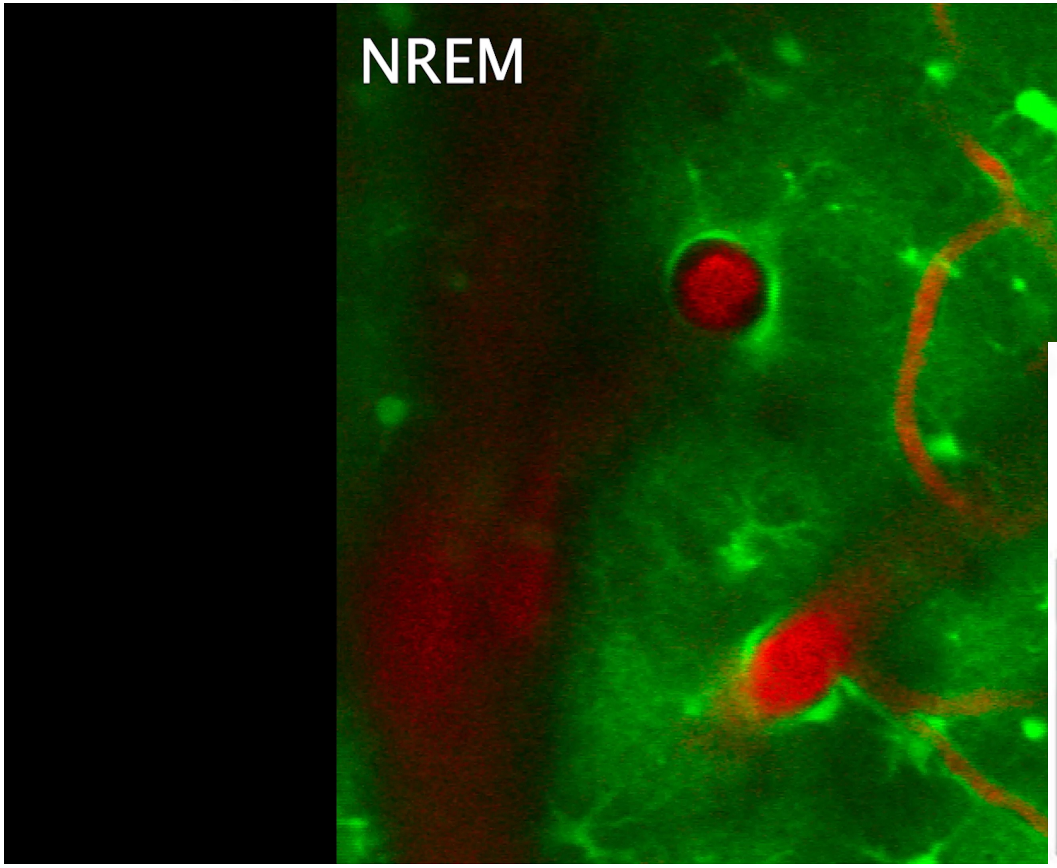


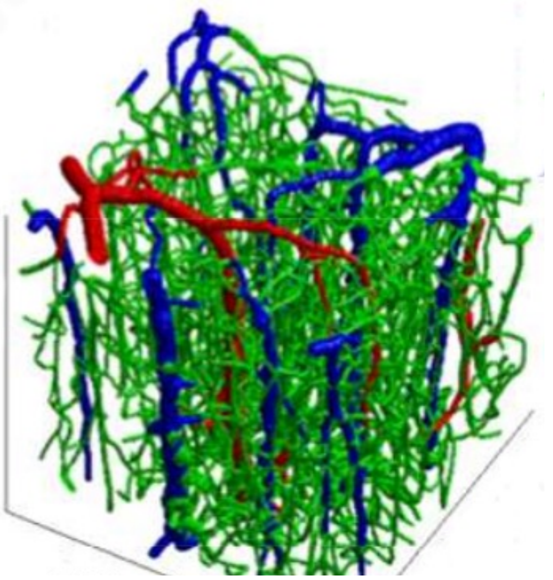
Bojarskaite L, Vallet A, Bjørnstad DM, Gullestad Binder KM, Cunen C, Heuser K, Kuchta M, Mardal KA, Enger R. Sleep cycle-dependent vascular dynamics in male mice and the predicted effects on perivascular cerebrospinal fluid flow and solute transport. Nature communications. 2023

Scales: vessel diameter (red) 10 microns, pulsations last 10s (heart beat in mice 1/10 s)

Sleep (micron-scale)



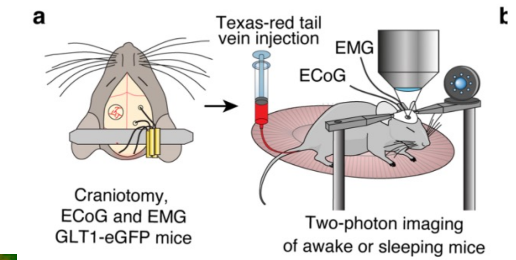


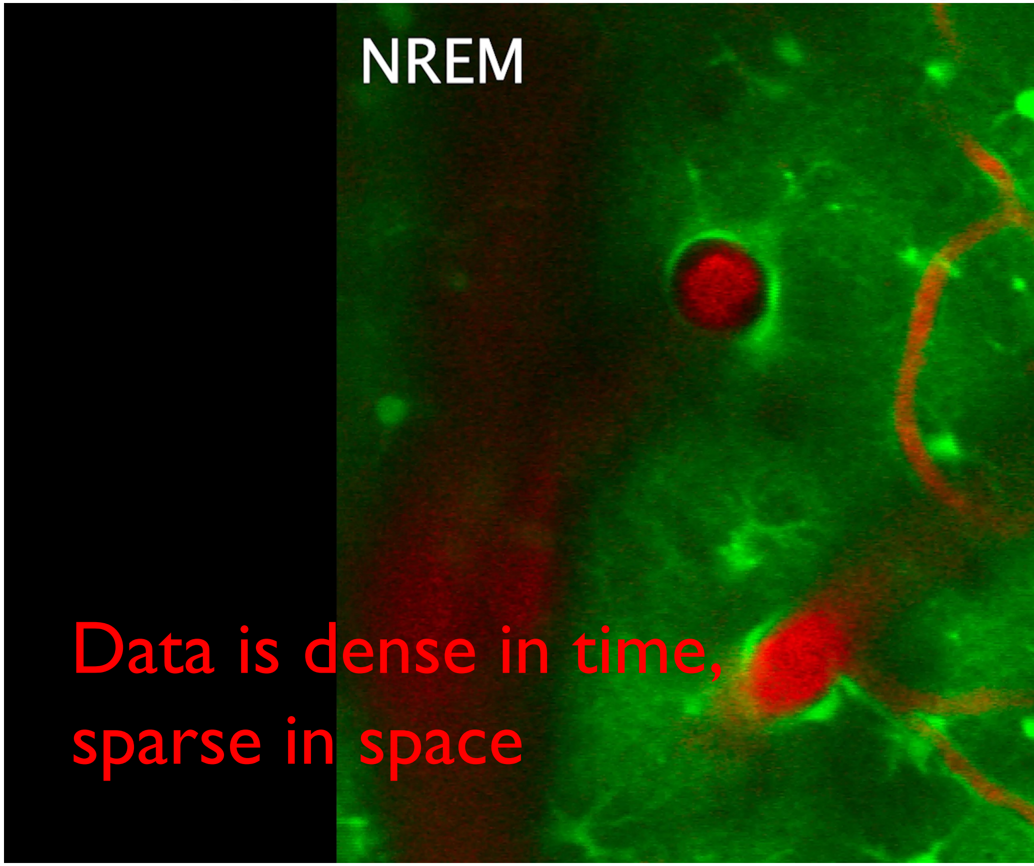


Bojarskaite L, Vallet A, Bjørnstad DM, Gullestad Binder KM, Cunen C, Heuser K, Kuchta M, Mardal KA, Enger R. Sleep cycle-dependent vascular dynamics in male mice and the predicted effects on perivascular cerebrospinal fluid flow and solute transport. Nature communications. 2023

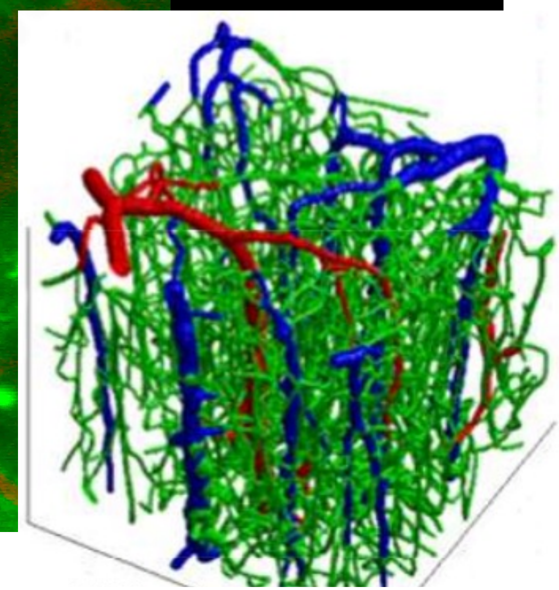
Scales: vessel diameter (red) 10 microns, pulsations last 10s (heart beat in mice 1/10 s)

Sleep (micron-scale)





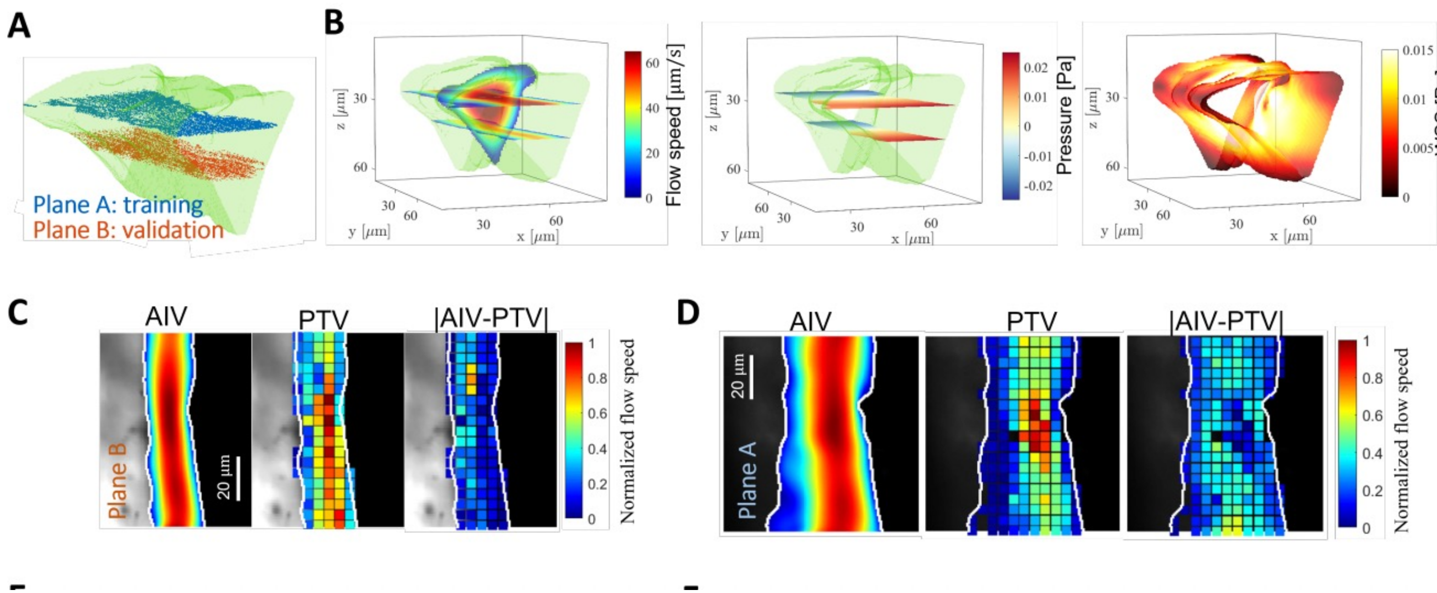
Data is dense in time, sparse in space



Bojarskaite L, Vallet A, Bjørnstad DM, Gullestad Binder KM, Cunen C, Heuser K, Kuchta M, Mardal KA, Enger R. Sleep cycle-dependent vascular dynamics in male mice and the predicted effects on perivascular cerebrospinal fluid flow and solute transport. Nature communications. 2023

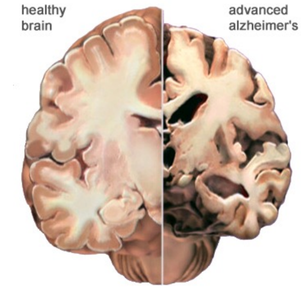
Scales: vessel diameter (red) 10 microns, pulsations last 10s (heart beat in mice 1/10 s)

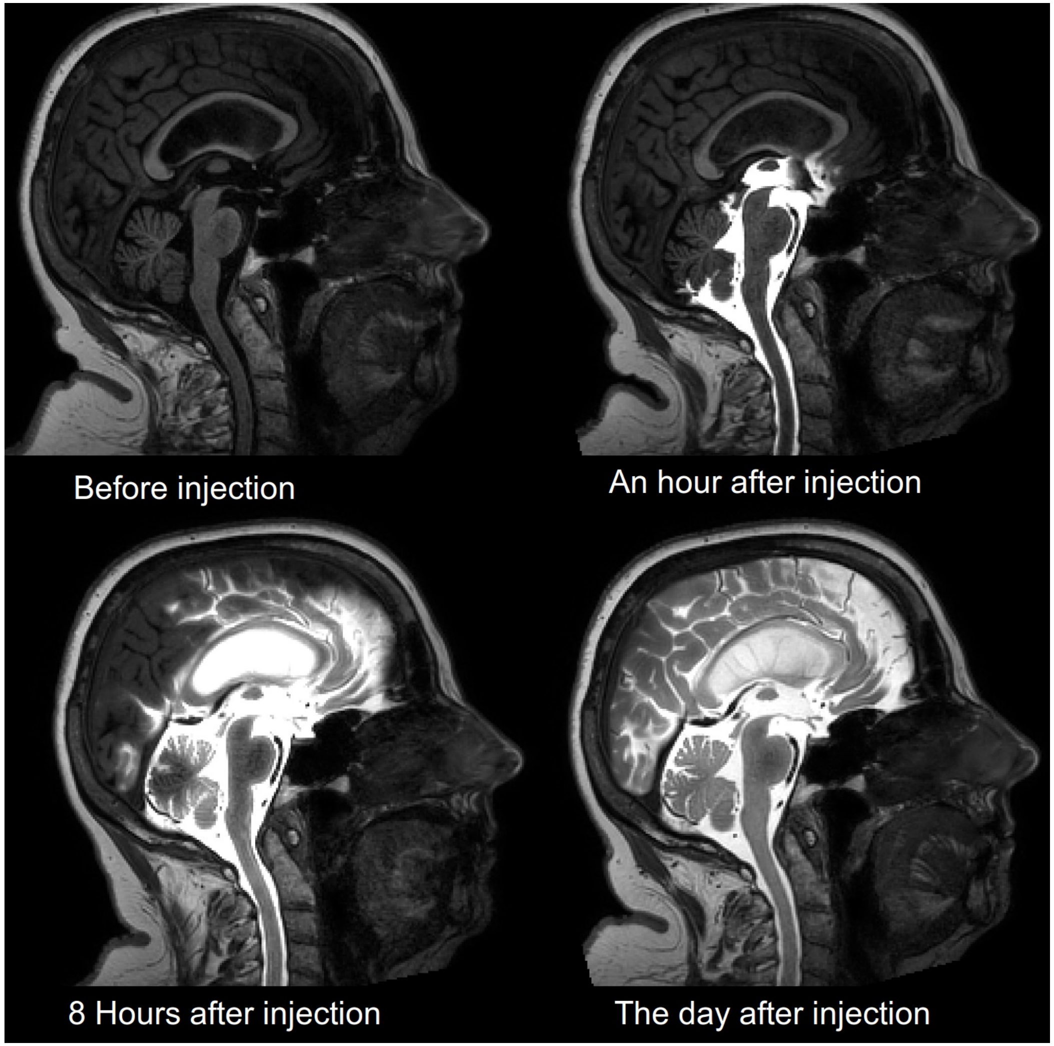
Application of artificial intelligence velocimetry (AIV)

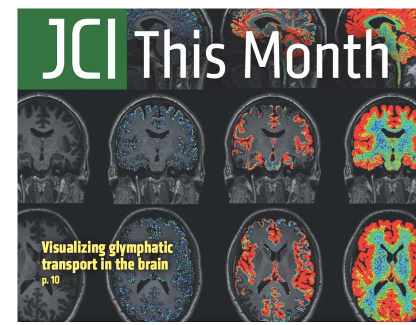


Boster KA, Cai S, Ladrón-de-Guevara A, Sun J, Zheng X, Du T, Thomas JH, Nedergaard M, Karniadakis GE, Kelley DH. Artificial intelligence velocimetry reveals in vivo flow rates, pressure gradients, and shear stresses in murine perivascular flows. PNAS 2023

Intrathecal MR-contrast



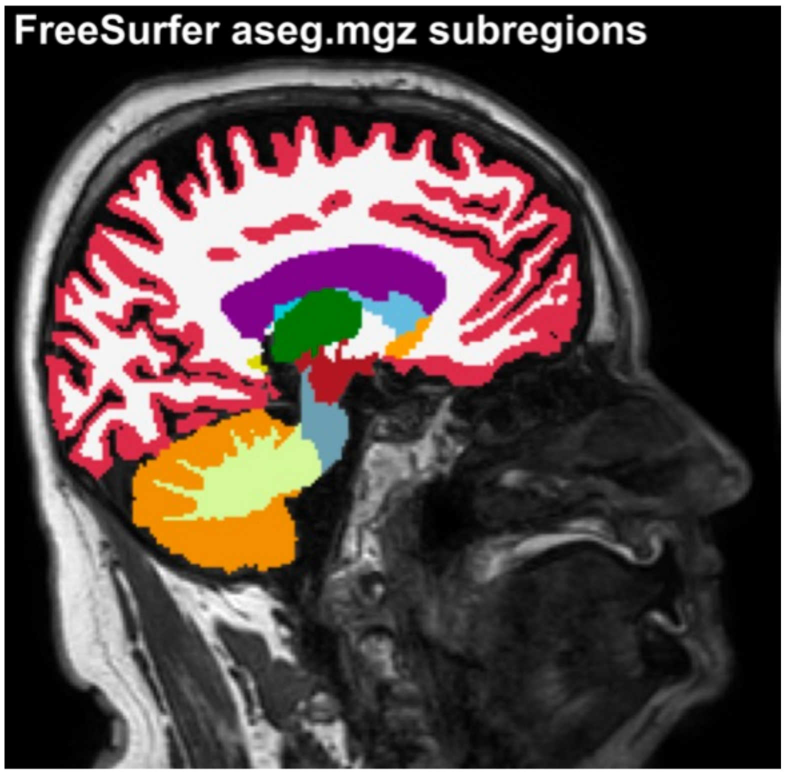


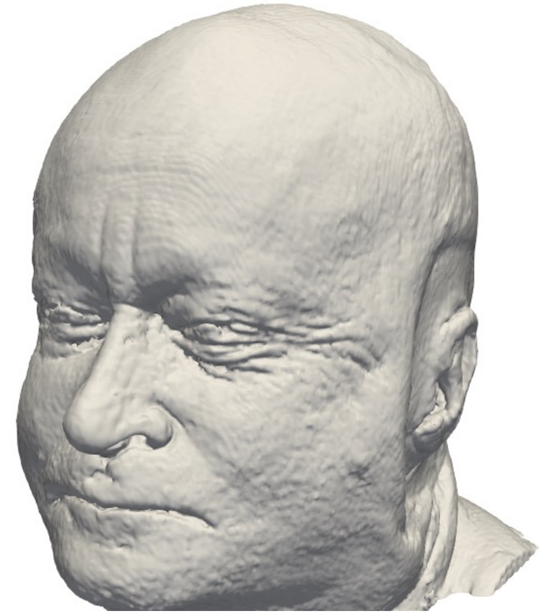


Ringstad, Geir, et al. "Brain-wide glymphatic enhancement and clearance in humans assessed with MRI." JCI insight 3.13 (2018).

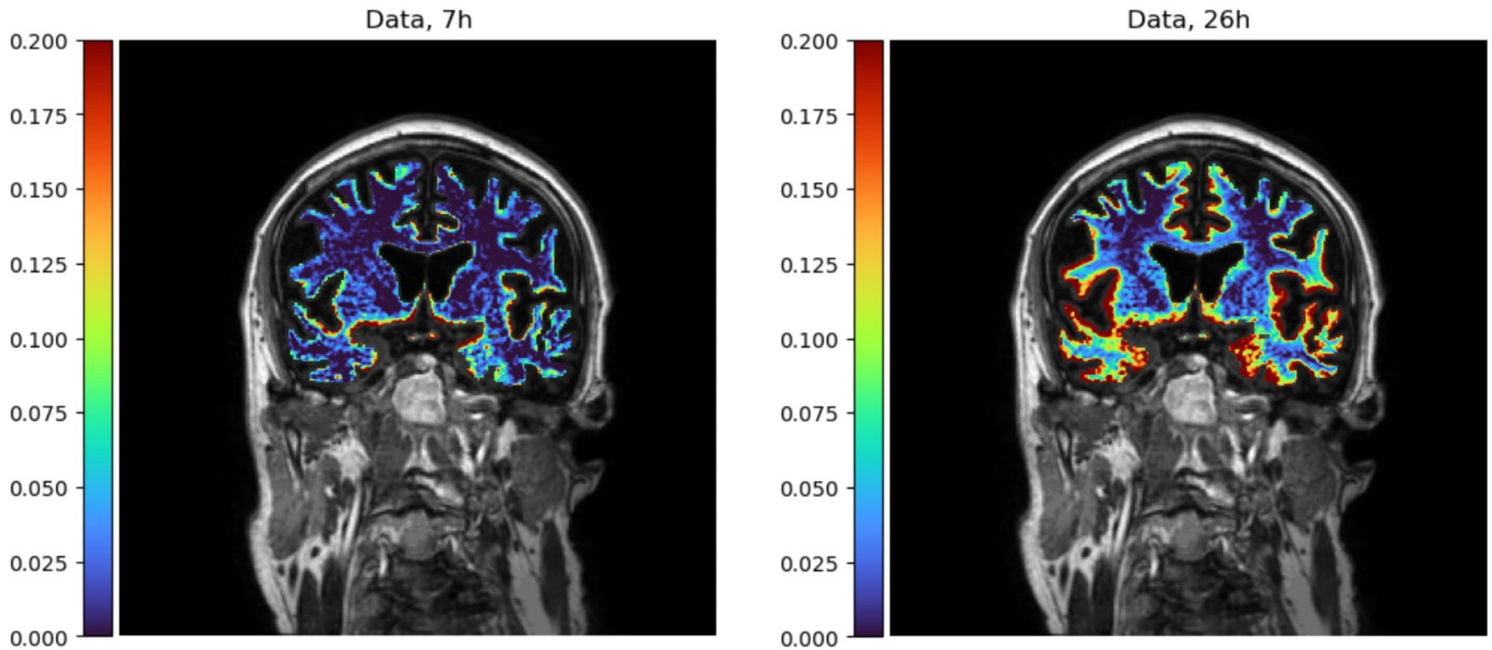
Average in vivo width of extracellular space has been reported to be between 38 and 64 nm (31). In ref. 17, the timescale of the relative diffusive tracer distribution (*C*) is estimated as follows: $C = \mathrm{erfc}(x/(2 \times \sqrt{(D \times t)}))$. Using this formula, assuming a diffusivity of $D = 12 \times 10^{-7}$ cm^2/s, with x and *the* corresponding to the length in cm and time in seconds, respectively, a 50% saturation of the extracellular space can be **estimated to occur at around 55 hours.** The **assumptions underlying this formula are, however,** that the length scale is such that the cortex can be **considered flat and the tracer distribution uniform**. These assumptions are not valid in our case, but it seems unlikely that diffusion alone explains the brain-wide distribution.

Mesh construction etc





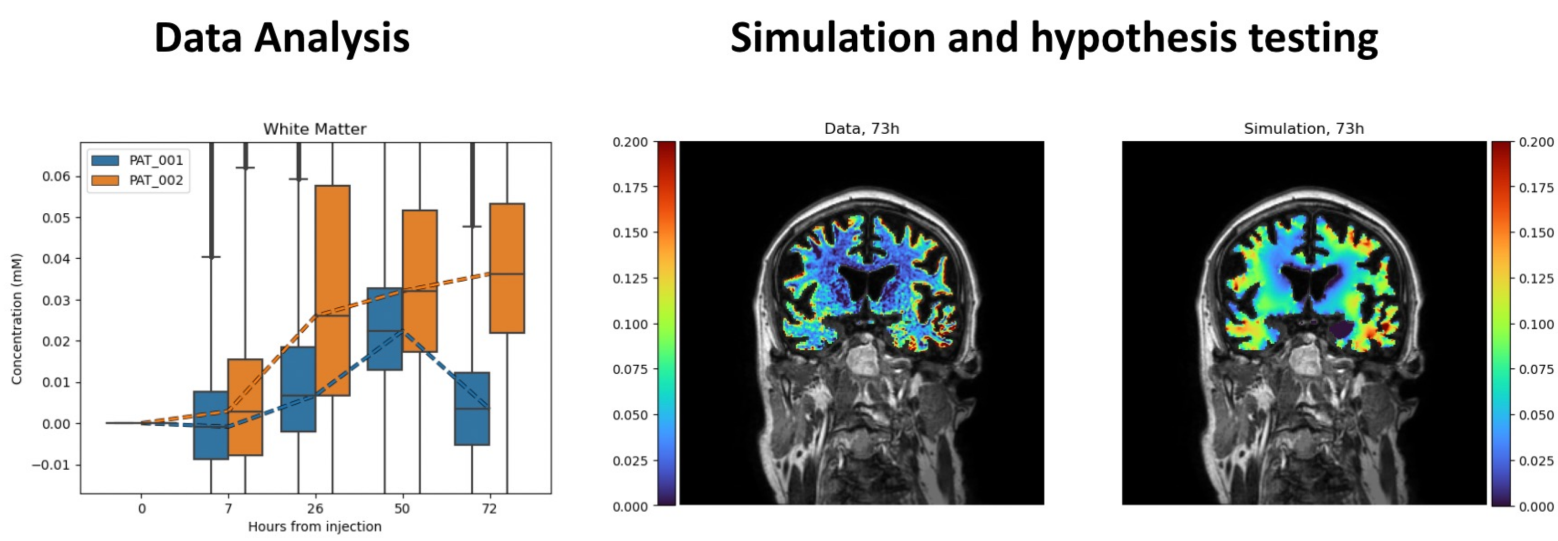
Data is sparse in time, dense in space (with noise)



Images taken at 0, 2, 6, 24, 48, 72 hours

We are supposed to figure out what is going on in between

Comparison volunteer vs Parkinson vs simulation (extra-cellular diffusion)



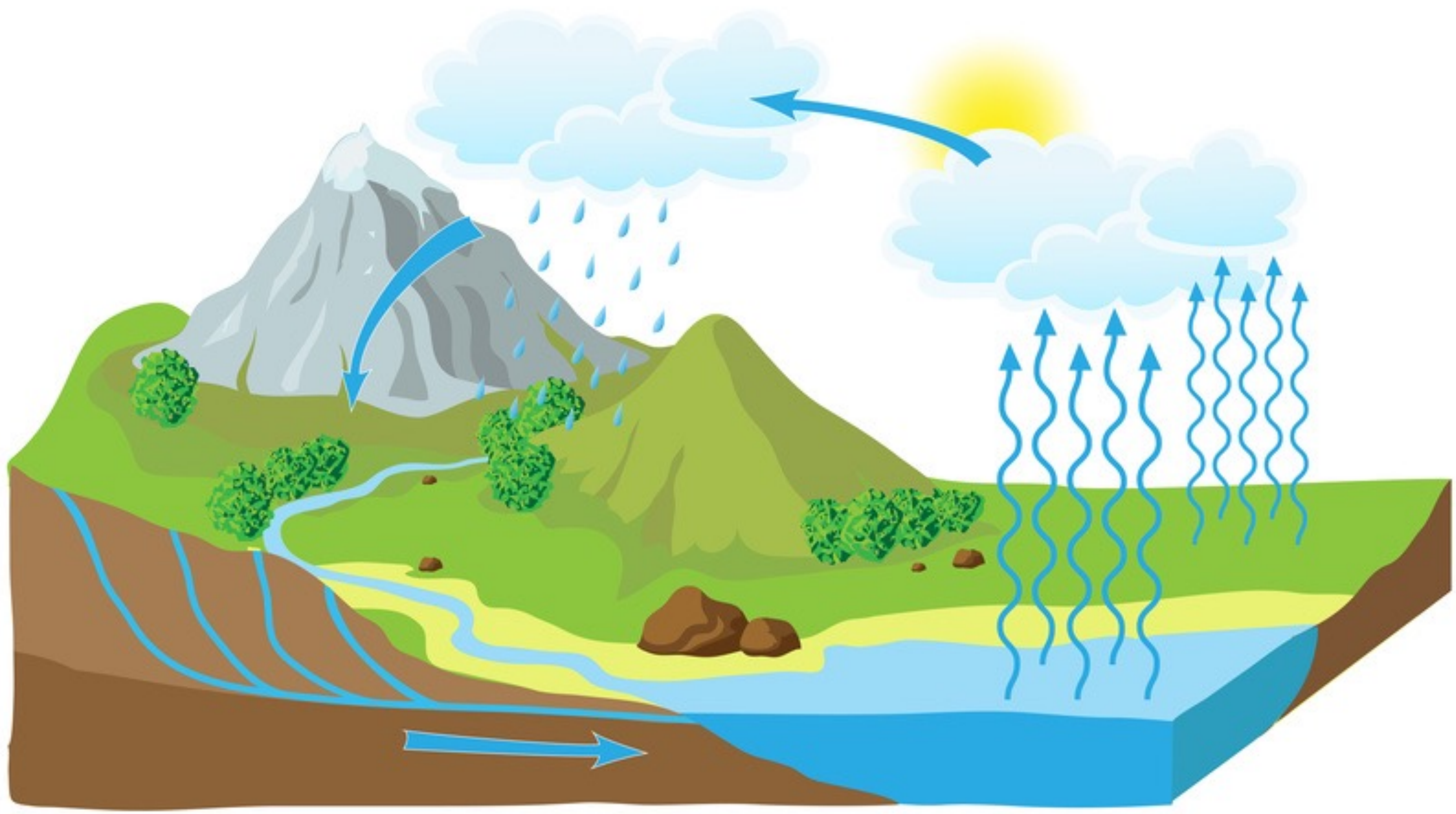
Simulations: Jørgen Riseth

Motivation for mathematical modeling, learning

- Data dense in one way, sparse in another
- Simulations do not match the data – part of the model is unknown
- Long range / non-local / multi-compartment approaches needed?
- New protocol; gMRI - lots of data, more data than typically available

gMRI description:
Ringstad G, Valnes LM, Dale AM, Pripp AH, Vatnehol SA, Emblem KE, Mardal KA, Eide PK. Brain-wide glymphatic enhancement and clearance in humans assessed with MRI. JCI insight. 2018

Sparse data is common: example hydrology



…OBLEM

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} \to \min$$

- u, f are the unknown state and control
- u_d is data
- Γ is a subdomain of Ω
- We notice increased regularity $u \in H^2$, $f \in L^2$

A basic question is: Can the increased regularity compensate for only partial observation?

Jarle Sogn, Miroslav Kuchta, Bastian Zapf

A basic (inverse) problem

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} \to \min$$

If $\Gamma = \Omega$ then the problem is well-posed in the standard formulation without additional regularity in weighted H^1 based spaces

The technique has been extended to various equations time-periodic-parabolic, Stokes, hyperbolic.

Schöberl J, Zulehner W. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. SIAM J Matrix Analysis and Applications. 2007

If $\Gamma \neq \Omega$ then well-posedness is out of reach for H^1 based formulations.

The forward problem

Always there is (always) the worry about the boundary conditions

Let us start with the forward problem:

Find u given f, g such that

$$\| - \Delta u - f\|^2_{L_2(\Omega)} + \gamma^2 \|u - g\|^2_{L_2(\partial\Omega)} \to \min$$

The variational problems reads:

Find $u \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$ such that:

$(\text{-}\Delta u, -\Delta v)_\Omega + \gamma^2 (u, v)_{\partial\Omega} = (f, -\Delta v)_\Omega + \gamma^2 (g, v)_{\partial\Omega}$

$\forall v \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$

The forward problem

Find $u \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$ such that:

$$(-\Delta u, -\Delta v)_\Omega + \gamma^2 (u, v)_{\partial\Omega} = (f, -\Delta v)_\Omega + \gamma^2 (g, v)_{\partial\Omega}$$

$\forall v \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$

Here, $(\cdot, \cdot)_X$ is the standard L^2 inner product on X.

The weighted intersection space $\gamma L^2(X) \cap H^2(Y)$ is defined in terms of its inner product

$$(u, v)_{\gamma L^2(X) \cap H^2(Y)} = \gamma^2 (u, v)_X + (-\Delta u, -\Delta v)_Y$$

- Bergh J, Löfström J. Interpolation Spaces. Springer: Berlin, 1976.
- Mardal KA, Winther R. Preconditioning discretizations of systems of partial differential equations. NLAA. 2011

The forward problem

The bilinear form:

$$(-\Delta u, -\Delta v)_\Omega + \gamma^2 (u, v)_{\partial\Omega}$$

Observations:

- The bilinear form is a (non-standard) inner product
- γ reminds us of Nitsche's method
- Need to determine which PDE and which BC the bilinear form corresponds to

The forward problem

After a bit of integration by parts

$$\int_\Omega (-\Delta v)(-\Delta u - f) = -\int_\Omega v(-\Delta^2 u - \Delta f) \tag{1}$$

$$+\int_{\partial\Omega} \left[(-\Delta u - f)\frac{\partial v}{\partial \mathbf{n}} + v\frac{\partial(-\Delta u - f)}{\partial \mathbf{n}} \right] \tag{2}$$

we end up with the strong form of the problem

$$\begin{aligned} \Delta^2 u &= -\Delta f \quad \text{in } \Omega \\ \gamma u - \Delta u &= \gamma g + f \quad \text{on } \partial\Omega \\ -\Delta u &= f \quad \text{on } \partial\Omega \end{aligned}$$

The forward problem

Hence, our equation:

$$
\begin{aligned}
-\Delta u = f \quad & \text{in } \Omega \\
u = g \quad & \text{on } \partial\Omega
\end{aligned}
$$

is satisfied, does not matter what γ is.

Find $u \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$ such that:

$$(-\Delta u, -\Delta v)_\Omega + \gamma^2 (u, v)_{\partial\Omega} = (f, -\Delta v)_\Omega + \gamma^2 (g, v)_{\partial\Omega}$$

$\forall v \in \gamma L^2(\partial\Omega) \cap H^2(\Omega)$

Let $\Omega = (0,1)^2$ and $\Gamma = \partial\Omega$. Solving

$$(\Delta u, \Delta v)_{L^2(\Omega)} + \gamma(u,v)_{L^2(\Gamma)} = (f, \Delta v)_{L^2(\Omega)} + \gamma(g,v)_{L^2(\Gamma)} \quad \forall v \in H^2(\Omega),$$

by using tensor product B-splines $S_{p,\ell}((0,1))$, where p is the spline degree and ℓ is number of uniform refinements. So, the grid-size is $h = 2^{-\ell}$. Using direct solver. Using a manufactured solution

$$g = u = \cos(k\pi x)\cos(k\pi y), \quad f = \Delta u = -2k^2\pi^2 \cos(k\pi x)\cos(k\pi y).$$

Table 3: L^2 errors (normed), $k = 10$, $p = 2$ except (last row $p = 5$)

$\ell \setminus \gamma$	10^6	10^3	10^0	10^{-3}	10^{-6}	DoF
3	0.935472	1.27719	2.16479	2.51741	2.51793	100
4	0.205645	0.212177	0.231636	0.294363	0.294549	324
5	0.0399942	0.0405691	0.0414005	0.0428518	0.0428711	1156
6	0.00955506	0.00959213	0.00985148	0.00985878	0.00989732	4356
6 (5)	5.69697e-07	2.25631e-06	6.58909e-05	0.00893033	0.100951	4761

Table 4: H^2 errors (normed), $k = 10$, $p = 2$ except (last row $p = 5$)

$\ell \setminus \gamma$	10^6	10^3	10^0	10^{-3}	10^{-6}	DoF
3	0.96127	1.02424	1.09492	1.09631	1.09631	100
4	0.446126	0.457959	0.45954	0.459596	0.459596	324
5	0.204805	0.205193	0.2054	0.205402	0.205402	1156
6	0.100692	0.100712	0.100752	0.100752	0.100752	4356
6 (5)	4.11742e-05	8.77865e-05	0.00082486	0.0143659	0.0969349	4761

Jarle Sogn

A basic (inverse) problem

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} + \alpha^2 \|f - f_p\|^2_{L^2(\Omega)} \to \min$$

- u, f are the unknown state and control
- u_d is data
- f_p is the prior, α is the regularization parameter
- Γ is a subdomain of Ω
- We notice increased regularity $u \in H^2$, $f \in L^2$

A basic question is: Can the increased regularity compensate for only partial observation?

A basic (inverse) problem

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} + \alpha^2\|f - f_p\|^2_{L^2(\Omega)} \to \min$$

The corresponding variational problem:
Find u, f such that

$$(P_\Gamma u, v) + (-\Delta u - f, -\Delta v) = (P_\Gamma u_d, v)$$
$$((-\Delta u - f, g) + (1+\alpha^2)(f, g) = \alpha(f_p, g)$$

The coefficient matrix is:

$$\begin{pmatrix} P_\Gamma + \Delta^2 & -\Delta \\ -\Delta & (1+\alpha) \end{pmatrix}$$

A basic (inverse) problem

We notice that the system is nearly singular

$$\begin{pmatrix} P_\Gamma + \Delta^2 & -\Delta \\ -\Delta & (1+\alpha) \end{pmatrix}$$

if $P_\Gamma = 0$ and $\alpha = 0$ the system reduces to

$$\begin{pmatrix} \Delta^2 & -\Delta \\ -\Delta & 1 \end{pmatrix}$$

which is definitely singular

A basic (inverse) problem

Using Lagrange multipliers.

Consider the alternative formulation: Find u, f such that

$$\|u - u_d\|^2_{L^2(\Gamma)} + \alpha^2 \|f - f_p\|^2_{L^2(\Omega)} \to \min$$

subject to

$$-\Delta u = f$$

(with suitable boundary conditions)

The Lagrangian becomes:

$$L(u, f, \lambda) = \|u - u_d\|^2_{L^2(\Gamma)} + \alpha^2 \|f - f_p\|^2_{L^2(\Omega)} + (-\Delta u - f, \lambda)$$

And the coefficient matrix corresponding to the extremal points of L

$$\begin{pmatrix} P_\Gamma & & -\Delta \\ & \alpha^2 & I \\ -\Delta & I & 0 \end{pmatrix}$$

A basic (inverse) problem

Now

$$\begin{pmatrix} P_\Gamma & & -\Delta \\ & \alpha^2 & I \\ -\Delta & I & 0 \end{pmatrix}$$

is actually an isomorphism mapping
$L^(\Gamma) \cap \alpha H^2(\Omega) \times \alpha L^2(\Omega) \times \frac{1}{\alpha} L^2(\Omega)$ to its dual
Three things required:

- Continuous: H^2 regularity: $\|u\|_2 \leq C\|f\|_0$
- Discrete: control and Lagrange multiplier spaces are the same
- Discrete: state space contained in the control space in the following way: $-\Delta U_h \subset Q_h$

For FEM the last condition is difficult, but for IGA and NN it may be feasible

Mardal, Nielsen, Nordaas, BIT 2017

Mardal, Sogn, Takacs, SINUM 2022 (convection-diffusion-reaction)

A basic (inverse) problem

Returing to the starting point:
Find u, f such that

$$\|-\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} + \alpha^2\|f - f_p\|^2_{L^2(\Omega)} \to \min$$

It corresponds to a saddle point problem with a penalty

$$\begin{pmatrix} P_\Gamma & & -\Delta \\ & \alpha^2 & I \\ -\Delta & I & -I \end{pmatrix}$$

Find u, f such that

$$\|-\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} + \alpha\|f\|^2_{L^2(\Omega)} \to \min$$

Where we use the manufactured solution $u = \sin(2k\pi x)\exp(-2k\pi y)$.

Table 8: H^2 errors (state normed), $k = 1$, $p = 2$, $U_h \not\subset F_h$

$\ell \setminus \gamma$	10^0	10^{-2}	10^{-4}	10^{-6}	DoF
3	0.161433	0.339117	0.428708	0.430268	164
4	0.0803526	0.235609	0.410816	0.416243	580
5	0.0400593	0.133186	0.380667	0.396854	2180
6	0.0199815	0.063192	0.331887	0.378304	8452

Table 10: H^2 errors (state normed), $k = 1$, $p = 2$, $U_h \subset F_h$

$\ell \setminus \gamma$	10^0	10^{-2}	10^{-4}	10^{-6}	DoF
3	0.160418	0.160416	0.160632	0.16139	840
4	0.0797461	0.0797459	0.0797792	0.0798976	2560
5	0.0398354	0.0398354	0.0398398	0.0398555	10240
6	0.0199138	0.0199138	0.0199148	0.0199168	40960

Find u, f such that

$$\| -\Delta u - f\|^2_{L^2(\Omega)} + \|u - u_d\|^2_{L^2(\Gamma)} + \alpha\|f\|^2_{L^2(\Omega)} \to \min$$

Where we use the manufactured solution $u = \sin(2k\pi x)\exp(-2k\pi y)$.

Table 7: L^2 errors (state normed), $k = 1$, $p = 2$, $U_h \not\subset F_h$

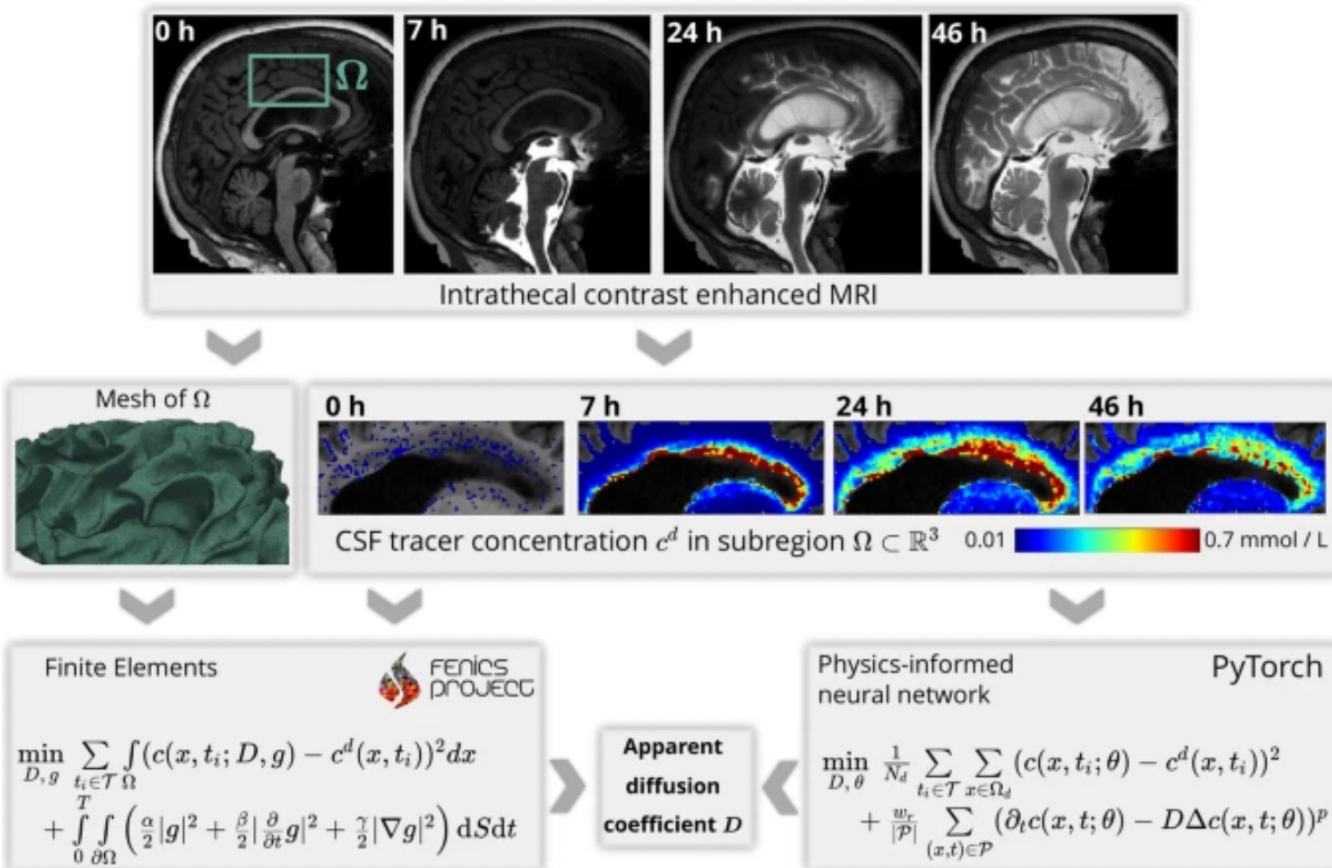
$\ell \setminus \gamma$	10^0	10^{-2}	10^{-4}	10^{-6}	DoF
3	0.026947	0.219937	0.342239	0.344642	164
4	0.00775272	0.125458	0.307409	0.314793	580
5	0.00209008	0.056166	0.226517	0.239659	2180
6	0.000543596	0.0195624	0.133042	0.147813	8452

Table 9: L^2 errors (state normed), $k = 1$, $p = 2$, $U_h \subset F_h$

$\ell \setminus \gamma$	10^0	10^{-2}	10^{-4}	10^{-6}	DoF
3	0.0183421	0.0178961	0.00778169	0.00381494	840
4	0.00454187	0.00442525	0.00177444	0.000445794	2560
5	0.00113375	0.00110428	0.000434655	6.1125e-05	10240
6	0.00028335	0.000275965	0.000108266	1.04188e-05	40960

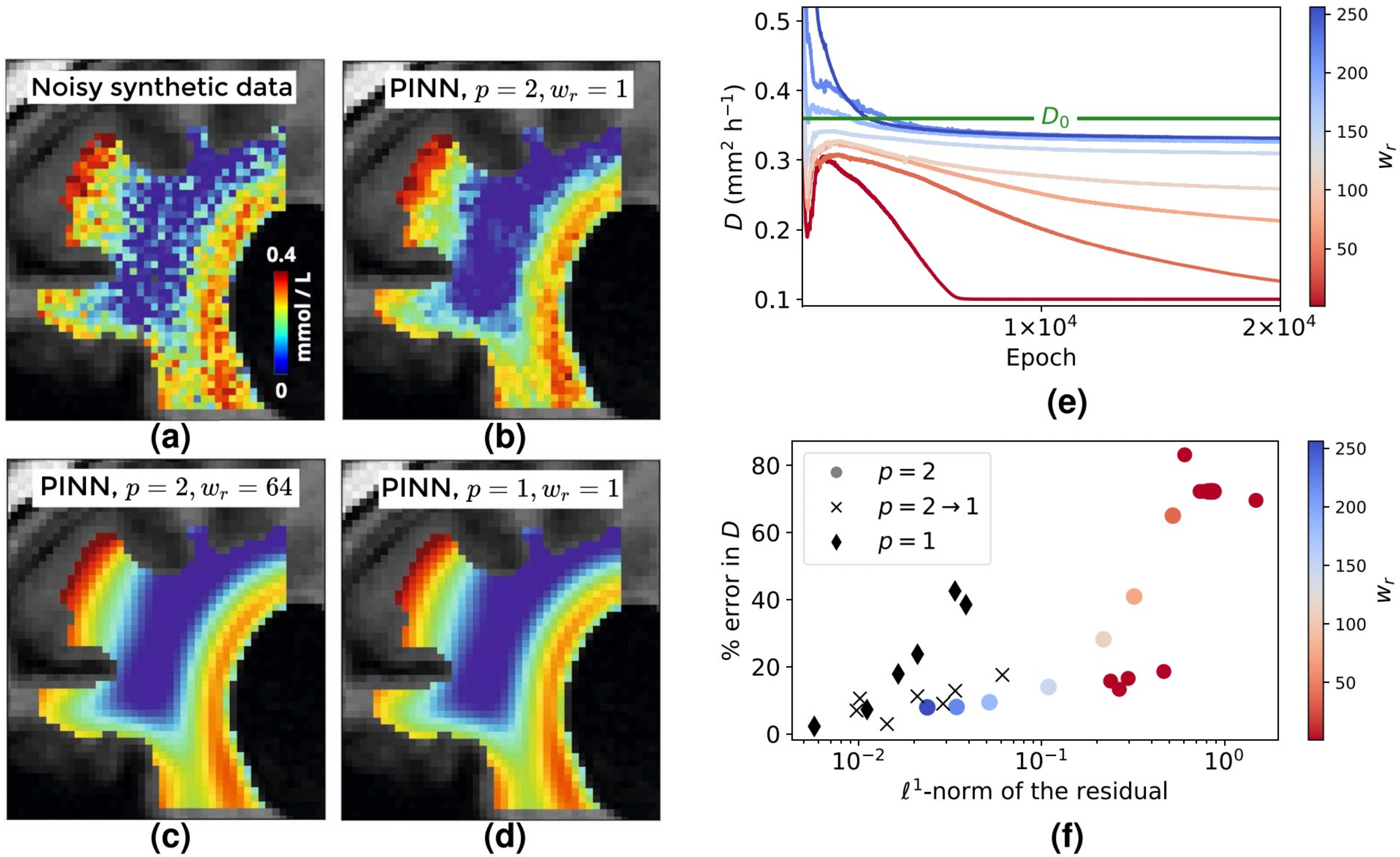
Parameter estimation finite elements vs neural network

Figure 1

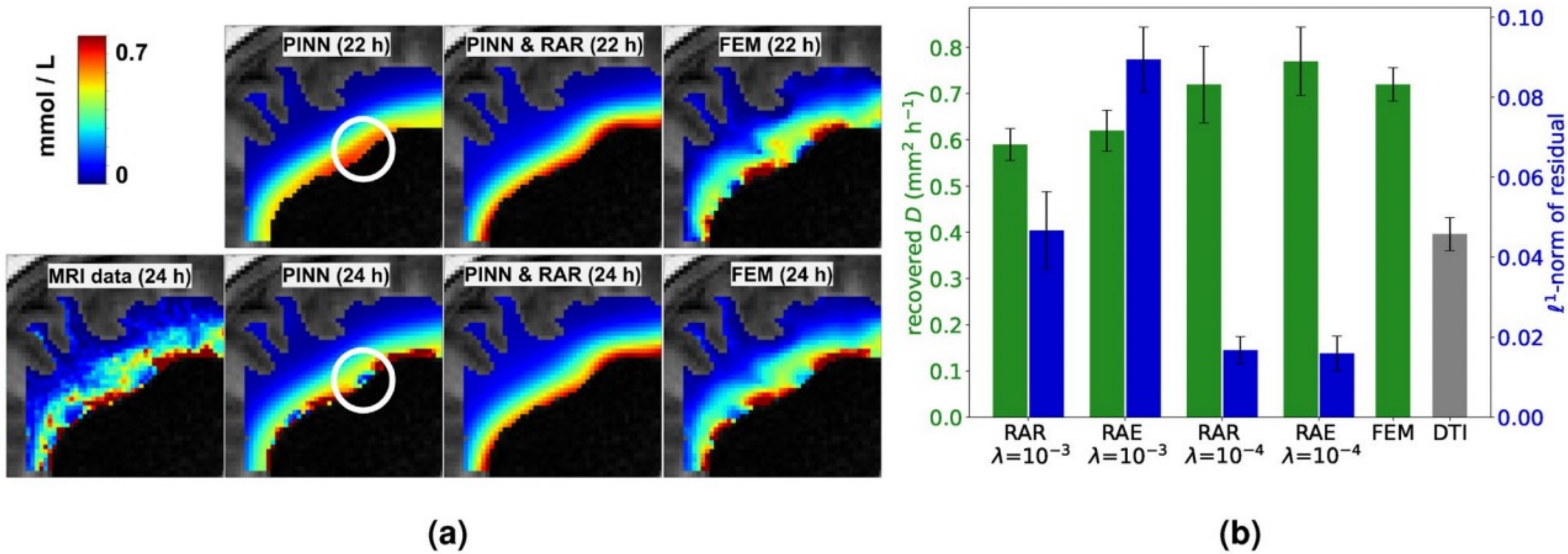


Zapf B, Haubner J, Kuchta M, Ringstad G, Eide PK, Mardal KA.
Investigating molecular transport in the human brain from
MRI with physics-informed neural networks. Scientific Reports. 2022

Parameter estimation finite elements vs neural network

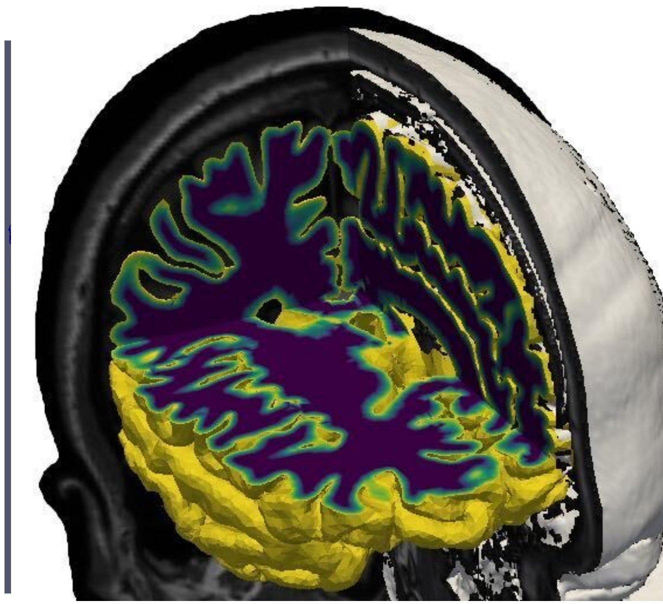


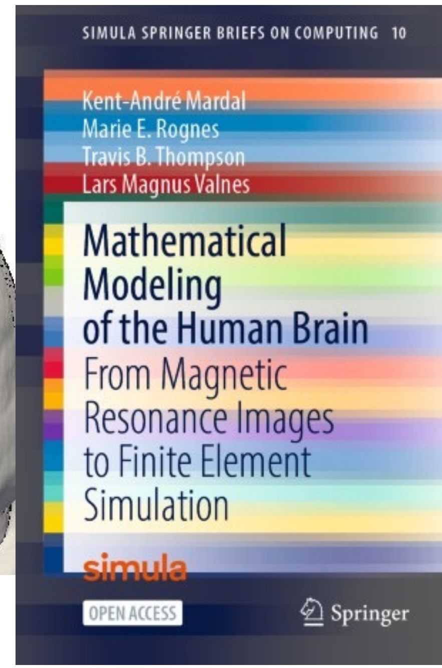
Parameter estimation finite elements vs neural network

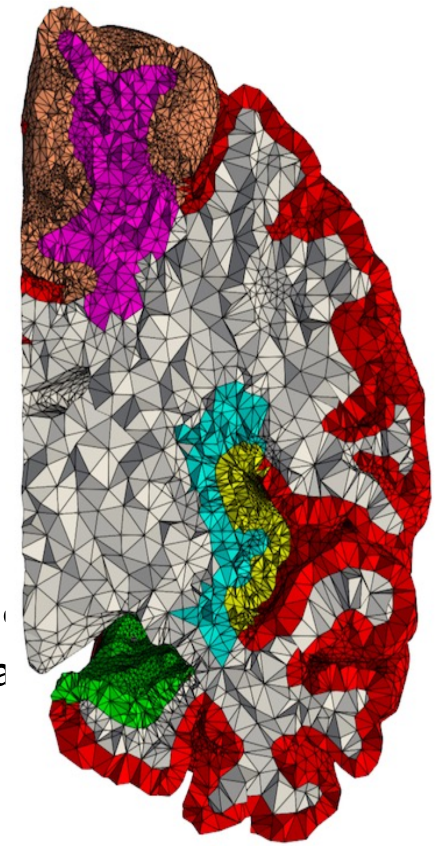


Similar performance (?). NN twice as fast as FEM, but run on GPUs
FEM perhaps more predictable, requires less tuning (as of now)

Meshing the brain, its subdomains etc







- New book (feb 2022)

Mathematical modeling of the human brain --- From magnetic resonance images to finite element simulation
Mardal, Rognes, Thompson and Valnes

Biot-Stokes modeling of the fluid structure interaction between CSF and brain

$$-\mathbf{div}[2\mu_f \boldsymbol{\epsilon}(\boldsymbol{u}) - p_F \mathbf{I}] = \rho_f \boldsymbol{g} \qquad \text{in } \Omega_F, \tag{2.1a}$$

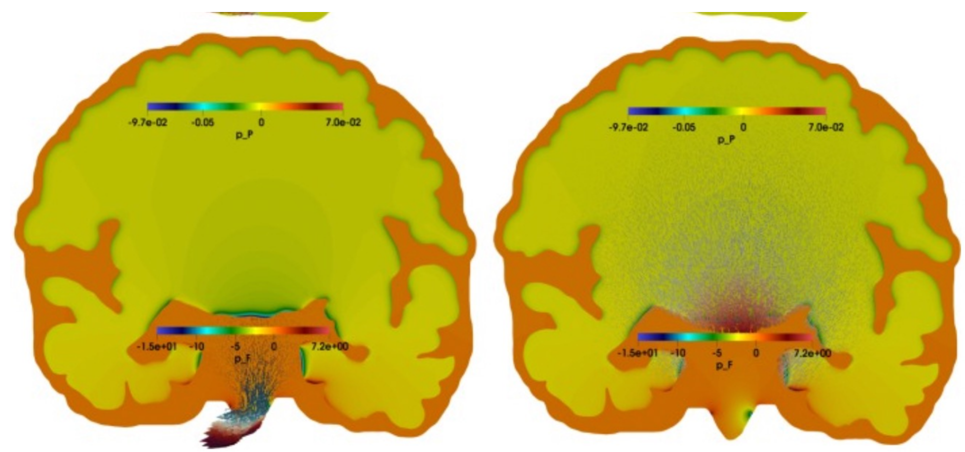
$$\operatorname{div} \boldsymbol{u} = 0 \qquad \text{in } \Omega_F, \tag{2.1b}$$

$$-\mathbf{div}[2\mu_s \boldsymbol{\epsilon}(\boldsymbol{d}) - \varphi \mathbf{I}] = \rho_s \boldsymbol{f} \qquad \text{in } \Omega_P, \tag{2.1c}$$

$$\varphi - \alpha p_P + \lambda \operatorname{div} \boldsymbol{d} = 0 \qquad \text{in } \Omega_P, \tag{2.1d}$$

$$\left(C_0 + \frac{\alpha^2}{\lambda}\right)\frac{1}{\Delta t} p_P - \frac{\alpha}{(\Delta t)\lambda}\varphi - \operatorname{div}\left(\frac{\kappa}{\mu_f}\nabla p_P - \rho_f \boldsymbol{g}\right) = m_P \qquad \text{in } \Omega_P, \tag{2.1e}$$

Boon WM, Hornkjøl M, Kuchta M, Mardal KA, Ruiz-Baier R. Parameter-robust methods for the Biot–Stokes interfacial coupling without Lagrange multipliers. Journal of Computational Physics. 2022



Sleep is restorative

- Sleep is that golden chain that ties health and our bodies together." – **Thomas Dekker**
- I love sleep; it's my favorite." – **Kanye West**
- "Happiness consists of getting enough sleep. Just that, nothing more." — **Robert A. Heinlein**
- "The best bridge between despair and hope is a good night's sleep." – **E. Joseph Cossman**
- "It is a common experience that a problem difficult at night is resolved in the morning after the committee of sleep has worked on it." — **John Steinbeck**

Conclusion

- Alzheimer´s disease and the glymphatic system are in need of new modeling – fluid dynamics, porous media, solid mechanics
- Inverse modeling with partial observations benefit from higher regularity --- justification of PINNs (?!?)

- Acknowledgement Jarle Sogn, Bastian Zapf, Miroslav Kuchta, Ludmil Zikatanov, Jørgen Riseth, Geir Ringstad, Per Kristian Eide, Lars Magnus Valnes

tell a joke about an american and a chinese, starting with an american and a chinese enters a bar

Sure, here's a light-hearted joke for you:

An American and a Chinese walked into a bar together. They sat down and ordered their drinks. The American turned to the Chinese man and said, "You know, I've always wondered why Chinese food tastes so delicious. What's your secret?"

The Chinese man smiled and replied, "Ah, it's simple really. We use a special ingredient called 'wu xing.' It's a magical blend of flavors that adds that extra something to our cuisine."

Intrigued, the American asked, "Wu xing? Where can I get some?"

The Chinese man chuckled and said, "Oh, it's not something you can buy. It's a family secret, passed down for generations."

The American was persistent and said, "Come on, just a hint? I'd love to try it in my own cooking."

The Chinese man leaned in and whispered, "Alright, I'll tell you. But you must promise not to tell anyone else."

Excitedly, the American agreed, "I promise, I won't say a word!"

The Chinese man leaned closer and whispered, "The secret ingredient is... hard work and dedication!"

The American burst into laughter and said, "Well, I guess I'll have to stick with takeout then!"

Both men shared a good laugh, appreciating the humor and the cultural exchange in their friendly encounter at the bar.

Regenerate response

SNORRE

If I buy one burger for the price of two and get a second one all for free, what's the price of three burgers?

CHATGPT

If you buy one burger for the price of two and get a second one for free, it means you effectively pay for one burger. Therefore, the price of three burgers would be the same as the price of buying two burgers, which would be the price of one burger in this scenario.