Optimal experimental design for sampling voltage on dendritic trees

Jonathan Huggins

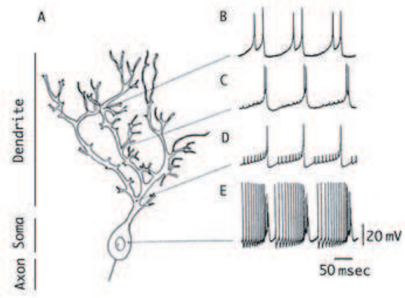
September 20, 2010

Table of contents

Table of Contents

Dendritic voltage sensing

- underlying experimental task: **voltage sensing**
- measure the potential difference across the membrane of the dendritic tree



modified from Llinas and Sugimori 1980

Voltage sensing methods

- state-of-the-art techniques:
  - multi-electrode recording
  - laser-based scanning techniques
- trade-off between spatial completeness and SNR
- our focus: the low SNR setting

What do we learn?

- biophysical quantities of interest
- e.g. passive cable parameters
- e.g. spatial membrane density distribution of voltage-gated channels

The challenges

Two challenges to address:

1. infer voltages across the full dendrite when $< 10\%$ of the dendrite is simultaneously observed
2. choose the best sampling scheme

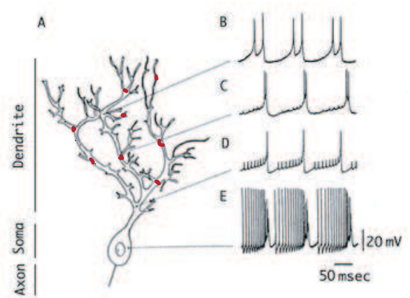


Table of Contents

The Kalman filter

- how to solve our difficulties? Statistics!
- specifically, the statistical model known as a Kalman filter-smoother
- why the Kalman filter?
    - models the dynamics to a good first approximation (in non-spiking situations)
    - incorporate noisy observations
    - provides "error bars"

Notation

- break dendrite into discrete compartments
- break time into discrete steps
- dt = time step length
- T = number of time steps
- N = number of dendritic compartments

The cable equation

$$V_{t+dt}(x) = V_t(x) + dt(-g_x V_{t+dt}(x) + \sum_{w \in N(x)} a_{xw}[V_{t+dt}(w) - V_{t+dt}(x)]). \quad (1)$$

- $V_t(x)$ is the voltage in compartment x at time t
- g_x is the membrane conductance in compartment x
- $N(x)$ is the set of compartments adjoining x
- a_{xw} is the intercompartmental conductance between compartments x and w

The Kalman equations

$$V_{t+dt} = AV_t + \epsilon_t, \; \epsilon_t \sim \mathcal{N}(0, \sigma^2 dtI) \quad (2)$$
$$y_t = B_t V_t + \eta_t, \; \eta_t \sim \mathcal{N}(\mu_t^y, W_t). \quad (3)$$

- A is a matrix that implements the backward-Euler implementation of the cable equation
- ϵ_t is Gaussian process noise
- $\{y_t\}$ are the vectors of the observed voltages
- η_t is Gaussian observation noise
- B_t is a matrix that specifies how the observations are related instantaneously to the voltage vector
- W_t is the covariance matrix that defines the noisiness of the observations

Calculating the covariance matrix

- well known how to calculate mean of all voltages given observations

$$\mu_t^s = E(V_t | Y_{1:T})$$

and the covariances

$$C_t^s = Cov(V_t | Y_{1:T})$$

- $Y_{1:t}$ denotes all of the observed data up to time t
- **problem:** standard methods to calculate mean and covariances require $\mathcal{O}(N^3)$ time and $\mathcal{O}(N^2)$ space

Calculating the covariance matrix efficiently

- previous work of [Paninski, 2010] showed how to calculate

$$\mu_t^f = E(V_t | Y_{1:t})$$

  and

$$C_t^f = Cov(V_t | Y_{1:t})$$

  in $\mathcal{O}(N)$ time and space
- [Paninski, 2010] used a low rank approximation to C_t^f:

$$C_t^f \approx C_0 + U_t D_t U_t^T$$

- C_0 is the covariance matrix if no observations are made
- similar techniques work for approximating smoothed covariance:

$$C_t^s \approx C_0 + P_t G_t P_t^T$$

Measuring sampling scheme quality

- can use C_t^s to evaluate the quality of a sampling scheme
- good news: C_t^s does not depend on the data
- two possible objective functions:
  1. (weighted) mean squared error (MSE) = (weighted) sum of the variances
     i.e. $\upsilon_w(\mathcal{O}) = \sum_{t=0}^{T} \sum_{i=1}^{N} w(i,t)[C_t^s]_{ii}$
  2. mutual information (MI) $= I(V_{1:T}, Y_{1:T}) = H(V) - H(V|Y)$
- focus on using the weighted MSE

Table of Contents

How to choose a sampling scheme?

- In general choosing the best k sample locations is an NP-hard problem
- $\binom{N}{k} = \frac{N!}{k!(N-k)!} = \mathcal{O}(N^k)$ possible combinations
- two solutions:
  1. use a heuristic
  2. take advantage of some properties of the objective function

Submodularity

- If our objective function is submodular we are in luck
- Can get to at least $(1 - 1/e) = 63\%$ of the best solution, guaranteed [Nemhauser et al., 1978]
- how? with a greedy algorithm

Greedy algorithm

Possible locations $= \{1, 2, 3, 4, 5, 6\}$

obs. set	1	2	3	4	5	6
$\emptyset$	.2	.1	.15	.2	.25	.09
$\{5\}$	.15	.08	.1	.12	X	.05
$\{5, 1\}$	X	.05	.09	.1	X	.04

Observation set $= \{5, 1, 4\}$

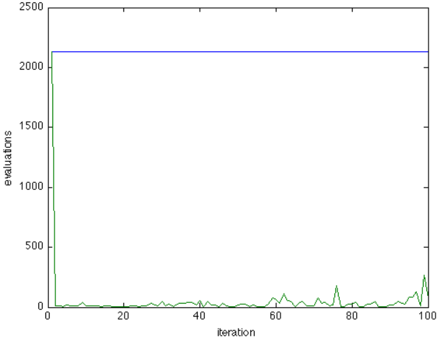
Lazy evaluation

bonus: we can use lazy evaluation

obs. set	1	2	3	4	5	6
$\emptyset$	.2	.1	.15	.2	.25	.09
$\{5\}$	.15	s	.1	.12	X	s
$\{5, 1\}$	X	.05	.09	.1	X	s

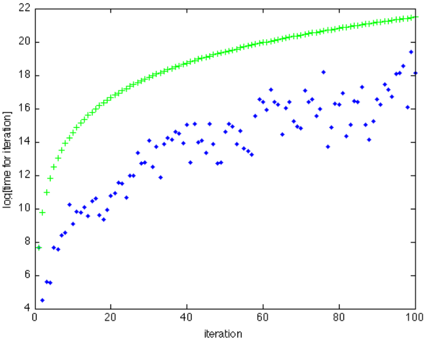
Observation set $= \{5, 1, 4\}$

Efficiency of lazy evaluation



- Number of evaluations required by the lazy vs. non-lazy greedy algorithm

Comparing run times



- Estimated relative time for each iteration of the lazy (blue) and non-lazy (green) greedy method (log scale)

What is submodularity?

- Submodularity is an intuitive diminishing returns property: the more observations added, the smaller the increase achieved by each additional observation.
- Let $\mathcal{S}$ be the set of observations from which to choose
- Formally, some real-valued function $F(\cdot)$ is submodular if for $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{S}$ and $e \in \mathcal{S} \setminus \mathcal{B}$, it holds that $F(\mathcal{A} \cup \{e\}) - F(\mathcal{A}) \geq F(\mathcal{B} \cup \{e\}) - F(\mathcal{B})$
- Said another way: including the element e in the argument set $\mathcal{O} \subseteq \mathcal{S}$ increases $F(\mathcal{O})$ less as $|\mathcal{O}|$ increases

Submodularity of objective functions

- MI submodular under very mild conditions
- MSE submodular under more onerous conditions (not fulfilled here)
- **nevertheless**, the greedy method worked well

Time dependent sampling

- greedy approach applies to this situation as well
- can start to eliminate possibilities once the per-time step quota k is reached

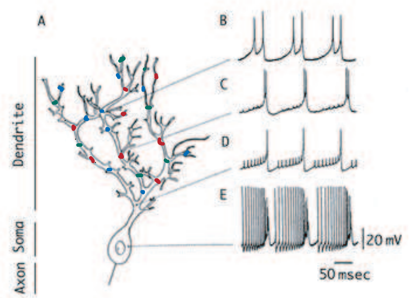
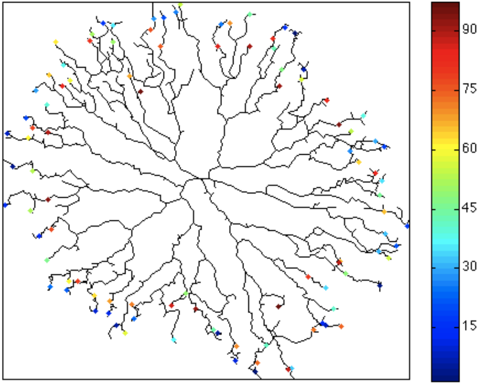


Table of Contents

Greedy sampling scheme

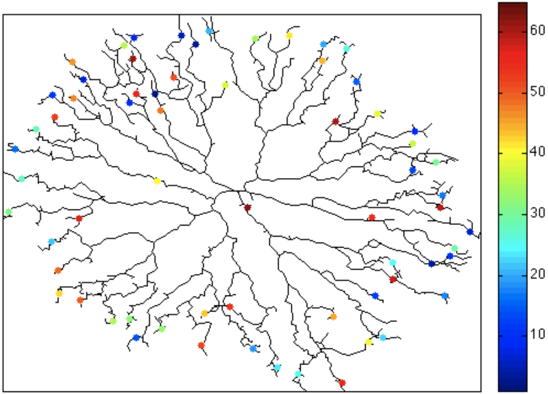


- Sampling scheme generated by greedily selecting 100 observation locations (rabbit starburst amacrine cell)

Exponentially weighted MSE

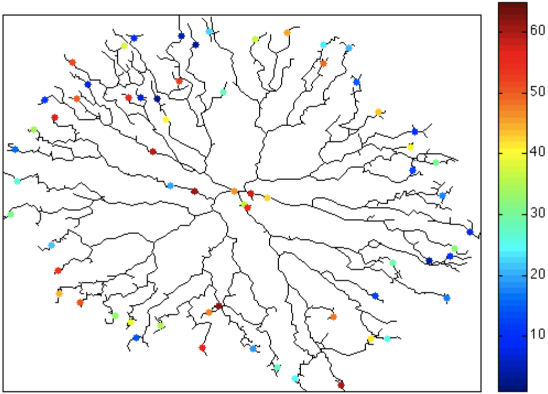
- $\upsilon_w(\mathcal{O}) = \sum_{t=0}^{T} \sum_{i=1}^{N} w(i,t)[C_t^s]_{ii}$
- in previous slide, $w(i,t) = 1 \ \forall i, t$
- what if we care about activity near the soma, not the periphery?
- instead let $w(i,t)$ decrease exponentially as the distance $d(i)$ between the soma and the compartment i increases

Sampling scheme with weighted MSE



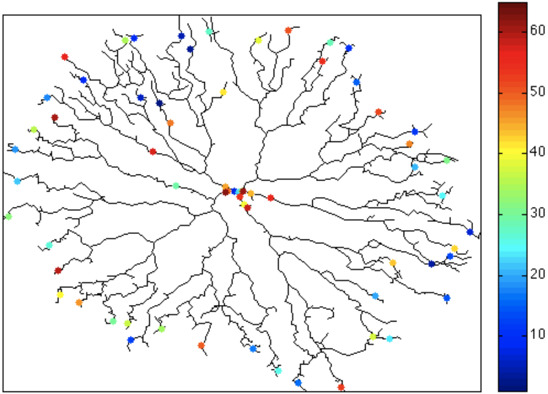
- First 65 observation locations selected when using an exponential weighting term (here $\alpha = 4$)

Sampling scheme with weighted MSE



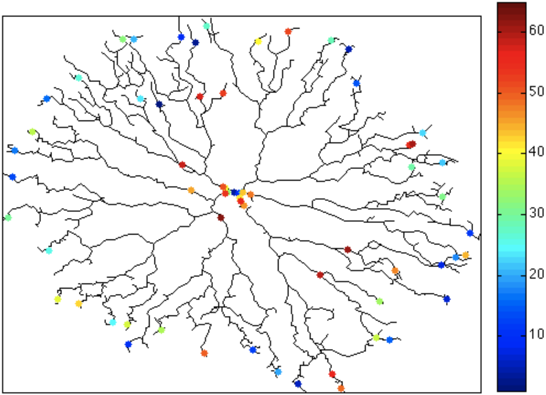
- $\alpha = 6$

Sampling scheme with weighted MSE



- $\alpha = 8$

Sampling scheme with weighted MSE

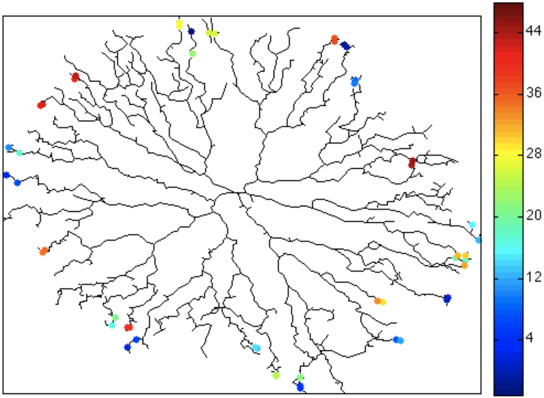


- $\alpha = 10$

Non-constant observation noise

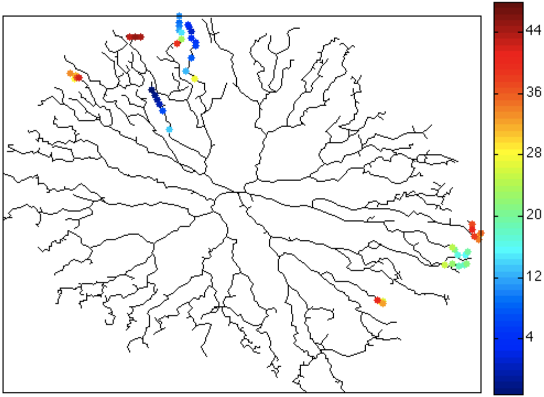
- voltage sensing generally has lower SNR along the periphery than near the soma
- model this by increasing the observation noise varied linearly with the distance of the compartment from the soma

Non-constant observation noise



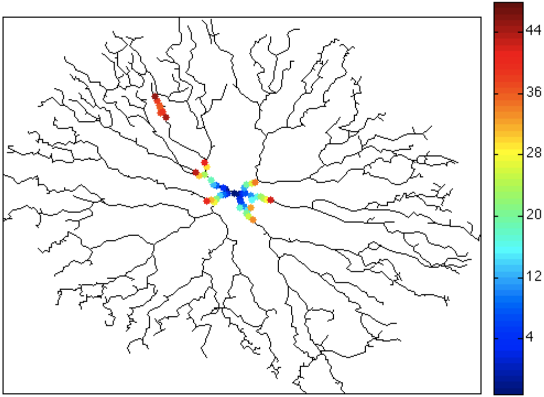
- 50 observation locations selected with higher peripheral observation noise
- min:max noise ratio = 1:2

Non-constant observation noise



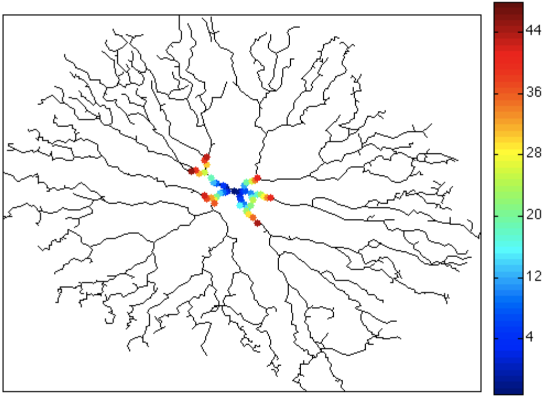
- min:max noise ratio = 1:5

Non-constant observation noise



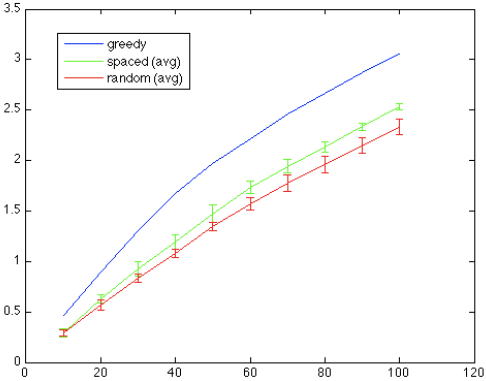
- min:max noise ratio = 1:25

Non-constant observation noise



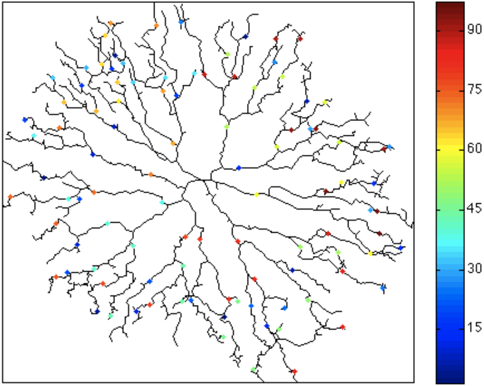
- min:max noise ratio = 1:50

Evaluating the performance of the greedy algorithm



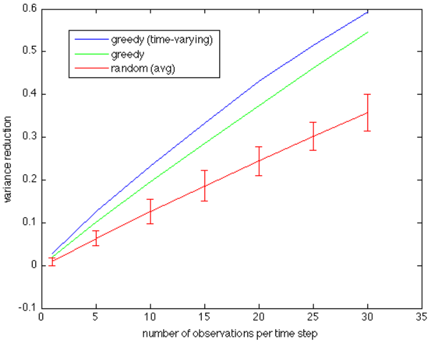
- variance reduction vs. number of observations per time step

Example of a Spaced Sampling Scheme



- 100 observation locations selected using the spaced method

Performance of greedy algorithm with time-varying observations



- variance reduction vs. number of observations per time step

Table of Contents

Conclusion

- state-space filter framework
- infer voltages from limited data
- efficient computation of metrics for optimal sampling design
- tractably design an optimal sampling scheme using a greedy algorithm with lazy evaluation

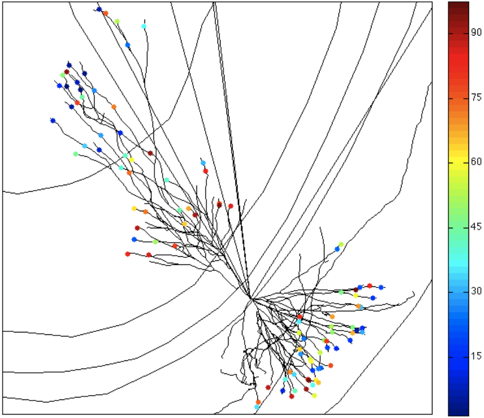
Future Work

- generalize results to non-linear and spiking cases
- generalize to allow time-variant dynamics
- investigate other metrics such as MI

Acknowledgements

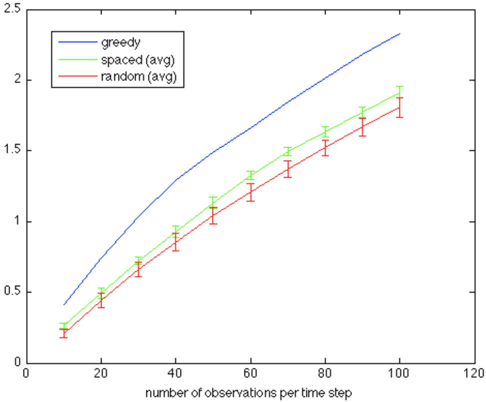
- L. Paninski
- K. Rahnama Rad
- Columbia College Rabi Scholars Program

Greedy sampling scheme



- Sampling scheme generated by greedily selecting 100 observation locations (rat hippocampal pyramidal cell)

Evaluating the performance of the greedy algorithm



- variance reduction vs. number of observations per time step

G. Nemhauser, L. Wolsey, and J. Fisher. An analysis of the approximations for maximizing submodular set functions. *Mathematical Programming*, 14:265–294, 1978.

L. Paninski. Fast Kalman filtering on dendritic trees. *Journal of Computational Neuroscience*, In press, 2010.