Supplementary Material: Inference-based assessment of parameter identifiability in nonlinear biological models

Aidan C. Daly[1†], David Gavaghan[1], Jonathan Cooper[2], and Simon Tavener[3]

[1]Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK. † Corresponding author.
[2]Research IT Services, University College London, Gower Street, London, WC1E 6BT, UK
[3]Department of Mathematics, Colorado State University, Fort Collins, CO 80523, USA.

June 18, 2018

A Monte Carlo methods

When considering a deterministic model with a known, calculable likelihood function over its outputs, we may apply a Markov Chain Monte-Carlo (MCMC) algorithm for posterior inference over model parameters. For this study we employed the MCMC algorithm described in [1] that proposes steps in parameter space according to a multivariate normal distribution Σ that is adaptively adjusted to attain an acceptance rate of $\approx 25\%$. Given a prior distribution over parameter space $\pi(\theta)$, a likelihood function $P(D|\theta)$ of the data D given parameters θ, and a starting point θ_0, it proceeds according to Algorithm A.1.

Algorithm A.1 Adaptive MCMC sampling

Initialize $\mathcal{S} \leftarrow \emptyset$ and $\mu_0 \leftarrow \theta_0$.
for t from 1 to N **do**
  Sample θ^* from $\mathcal{N}(\theta_{t=1}, \sigma_{t=1}\Sigma_{t-1})$
  Calculate the acceptance probability $\rho = \min\left(1, \frac{P(D|\theta^*)\pi(\theta^*)}{P(D|\theta_{t-1})\pi(\theta_{t-1})}\right)$
  Sample u from $\mathcal{U}(0,1)$
  if $u < \rho$ **then**
    $\theta_t \leftarrow \theta^*$
  else
    $\theta_t \leftarrow \theta_{t-1}$
  end if
  if $t > T_B$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \theta_t$
  end if
  $\mu_t \leftarrow \mu_{t-1} + t^{0.7}(\theta_t - \mu_{t-1})$
  $\Sigma_t \leftarrow \Sigma_{t-1} + t^{0.7}\left((\theta_t - \mu_{t-1})(\theta_t - \mu_{t-1})^T - \Sigma_{t-1}\right)$
  $\log(\sigma_t) \leftarrow \log(\sigma_{t-1}) + t^{0.7}(\rho - 0.25)$
end for
return $\mathcal{S}$

T_B defines a "burn-in" time that allows the sampler to converge before samples are recorded, and is generally taken to be equal to $N/2$. We initialized Σ_0 to be a diagonal matrix with (diagonal) elements defined to be 10% of the standard deviation of each corresponding parameter, σ_0 was set to 1, and θ_0 was chosen randomly from the prior $\pi(\theta)$.

When considering a model that employs stochastic elements, or when the likelihood function on the outputs of the model is unknown or incal-

culable for another reason, we must instead employ approximate Bayesian computation (ABC) methods. The simplest form of ABC sampler is the rejection sampler outlined below. Given a prior distribution over parameter space $\pi(\theta)$, a distance function $d(f(\theta), D)$ operating between simulated data $f(\theta)$ and experimental data D, and a tolerance level ε, it proceeds according to Algorithm A.2.

Algorithm A.2 ABC rejection sampler

Initialize an empty set $\mathcal{S} = \emptyset$ of accepted parameters
while $\mathcal{S}$ contains fewer than N particles **do**
  Draw a sample θ^* from a prior distribution $\pi(\theta)$
  if $d(f(\theta^*), D) < \varepsilon$ **then**
    Add θ^* to $\mathcal{S}$
  end if
end while
return $\mathcal{S}$

Because this algorithm has a very low acceptance rate when the prior differs significantly from the posterior, we employed a sequential Monte Carlo (SMC) algorithm proposed by Del Moral et al. [6] for all inference in this study. This algorithm samples from a series of successive intermediate distributions by gradually reducing ε to attain higher acceptance rates. Given the same information as for the rejection sampler, plus a proposal distribution $K(.|\theta)$ that generates moves in parameter space, it proceeds according to Algorithm A.3.

Algorithm A.3 Del Moral ABC-SMC sampler

Initialize $t = 0$
Using Algorithm A.2, construct an initial posterior estimate $\{\theta_0^{(i)}\}$ of size N using tolerance ε_0.
Assign equal weights $w_0^{(i)} = 1/N$ to each particle in this estimate.
while $\varepsilon_t > \varepsilon_{min}$ **do**
  for $i = 1$ **to** n **do**
    Set $w_t^{(i)} = \mathbb{I}\left(d(f(\theta_t^{(i)}), D) < \varepsilon_t\right)$
  end for
  Normalize $\{w_t^{(i)}\}$
  if fewer than $N/2$ particles have nonzero weight **then**
    Resample N particles $\{\theta^{(i)*}\}$ from $\{\theta_{t-1}^{(i)}\}$ using the new weights $\{w_t^{(i)}\}$.
    for $i = 1$ **to** n **do**
      Set $\theta_{t-1}^{(i)} = \theta^{(i)*}$ and $w_t^{(i)} = 1/N$.
    end for
  end if
  for $i = 1$ **to** n **do**
    Draw θ^* from $K_t(\theta|\theta_{t-1}^{(i)})$
    if $d(f(\theta^*), D) < \varepsilon_t$ **then**
      Set $\theta_t^{(i)} = \theta^*$
    else
      Set $\theta_t^{(i)} = \theta_{t-1}^{(i)}$
    end if
  end for
  $\varepsilon_{t+1} \leftarrow 0.8\varepsilon_t$
  $t \leftarrow t + 1$
end while
return $(\{\theta_t^{(i)}, w_t^{(i)}\})$

Here, we take the proposal distribution to be independently Gaussian across all model parameters. We periodically apply an update to a scalar modifier σ of the covariance matrix of these Gaussian distributions to heuristically maintain an acceptance rate of 25% for samples. The procedure is provided in Supplementary Algorithm A.4.

Algorithm A.4 Tuning of proposal distribution for Del Moral algorithm

$K_t \sim \mathcal{N}(0, \sigma\Sigma)$
Let a be the fraction of accepted steps over the last 100 proposals
if $a < 0.001$ **then**
  $\sigma \leftarrow 0.001s$
else if $a < 0.05$ **then**
  $\sigma \leftarrow 0.5s$
else if $a < 0.2$ **then**
  $\sigma \leftarrow 0.9s$
else if $a > 0.95$ **then**
  $\sigma \leftarrow 10s$
else if $a > 0.75$ **then**
  $\sigma \leftarrow 2s$
else if $a > 0.5$ **then**
  $\sigma \leftarrow 1.1s$
end if

B Measure theoretic inverse sensitivity

An alternative approach to Monte Carlo samplers for inference on model parameters given probability distributions on the outputs, is the measure theoretic inverse sensitivity approach that was developed in [2, 3] for scalar valued functions (maps) and extended to vector valued maps in [4]. We consider a deterministic map $g : \Theta \to \mathcal{G}$ from the input (parameter) space $(\Theta, \mathcal{B}_\Theta)$ to the output space $(\mathcal{G}, \mathcal{B}_\mathcal{G})$, and a known probability distribution $P_\mathcal{G}$ on $(\mathcal{G}, \mathcal{B}_\mathcal{G})$. A simple implementation of the method appears in Algorithm B.1 to determine the corresponding probability distribution on $(\Theta, \mathcal{B}_\Theta)$.

Algorithm B.1 Measure theoretic inverse sensitivity

Partition Θ into sets $\Theta_i, i = 1, \ldots, M$
Let $\theta_i \in \Theta_i, i = 1, ..., M$ be the centroid of Θ_i
Partition $\mathcal{G}$ in to sets $Y_j, j = 1, \ldots, N$
Let A be an $M \times N$ zero matrix
for i from 1 to M **do**
  Evaluate the map g at θ_i
  $A_{ij} = A_{ij} + 1$ iff $g(\theta_i) \in Y_j$
end for
Let $q_j = P(Y_j)$ for $j = 1, \ldots, N$
Then $p_i = A_{ij} * q_j / (\sum_{i=1}^{M} A_{ij})$
return $\{\Theta_i, p_i\}$

After evaluating the forward map g at all parameters $\theta_1 \ldots \theta_M$, we may store and re-use this map for future use. Any quantities of interest that may be calculated from the model output can be calculated without the need for additional simulation, allowing modelers to easily compute the inverse map for many possible summary statistics of model output.

For all analyses in this paper, we independently divided each dimension of parameter space into 100 equally sized partitions ($M = 10^{2N_\theta}$) and divided the output space $\mathcal{G}$ into 20 equal partitions across the observable range. A detailed convergence analysis of this algorithm is provided in [5].

C Supplementary Figures

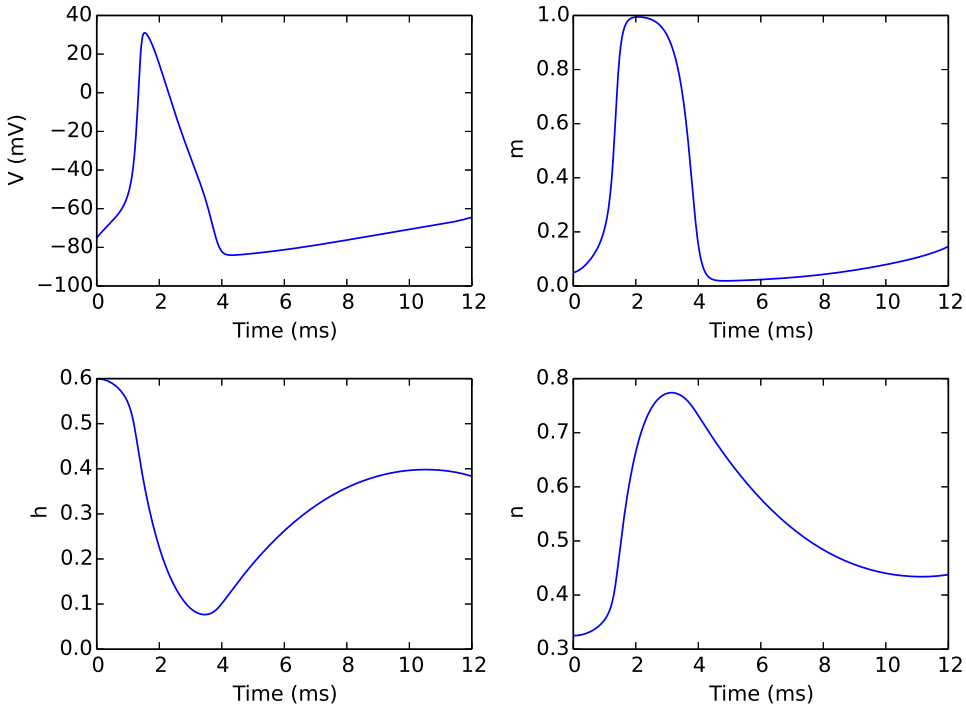

Figure C.1: Evolution of the four state variables (V, m, h, n) of the Hodgkin-Huxley model over the course of the action potential simulation with parameters $G_{Na} = 120$, $G_K = 36$, $G_l = 0.3$ (as described in Section 4). Note that while V is presented in mV, m,h, and n are dimensionless.

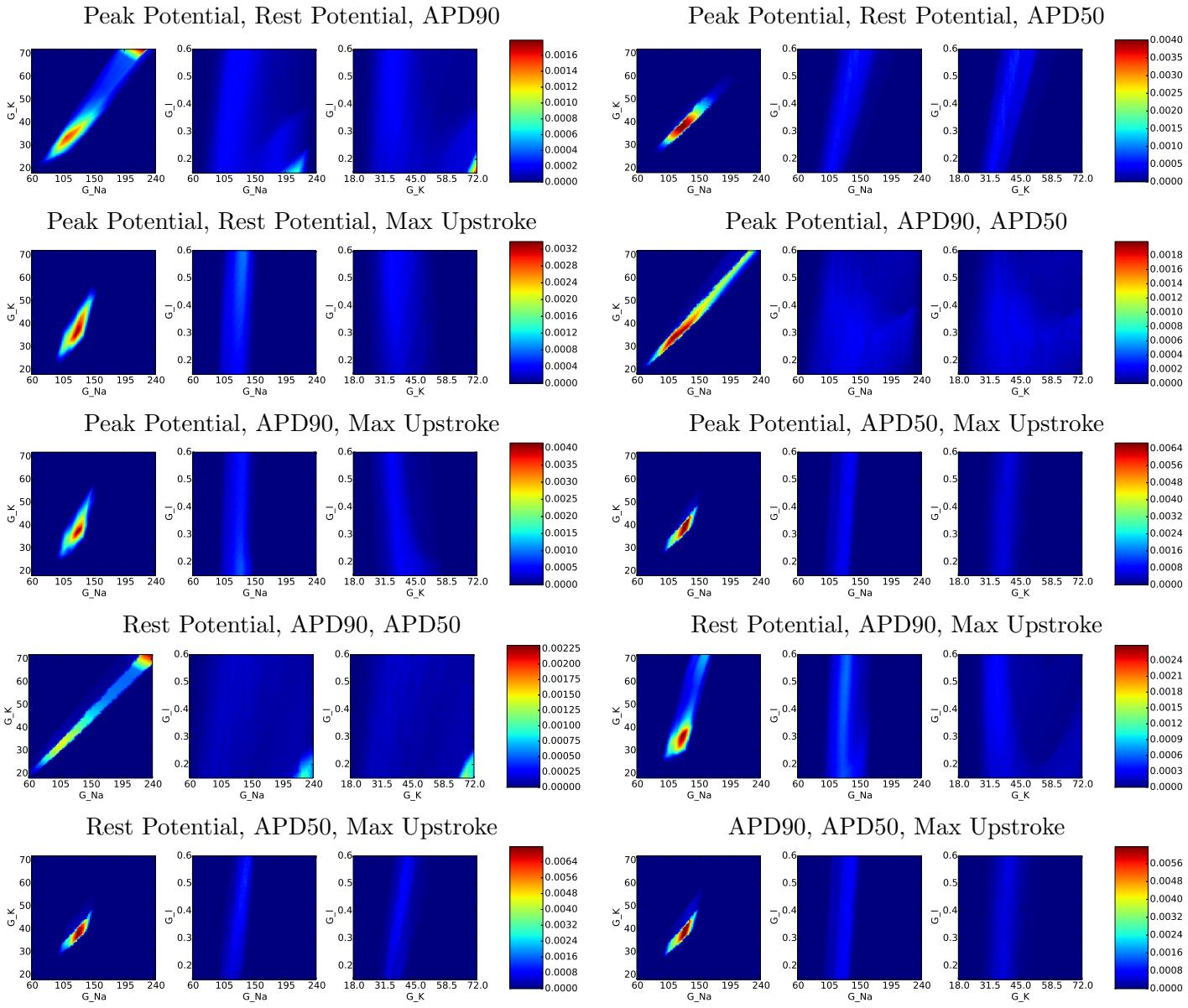

Figure C.2: Biplot visualizations of input (parameter) probability distributions for the Hodgkin-Huxley model (**??**) calculated by the measure theoretic inverse sensitivity method employing three summary statistics (indicated by labels above each triptych) observed from a "true" action potential trace. Axes (in units mS/cm^2) span the full uniform prior for each conductance parameter, and the posterior likelihood is indicated by the corresponding color legend.

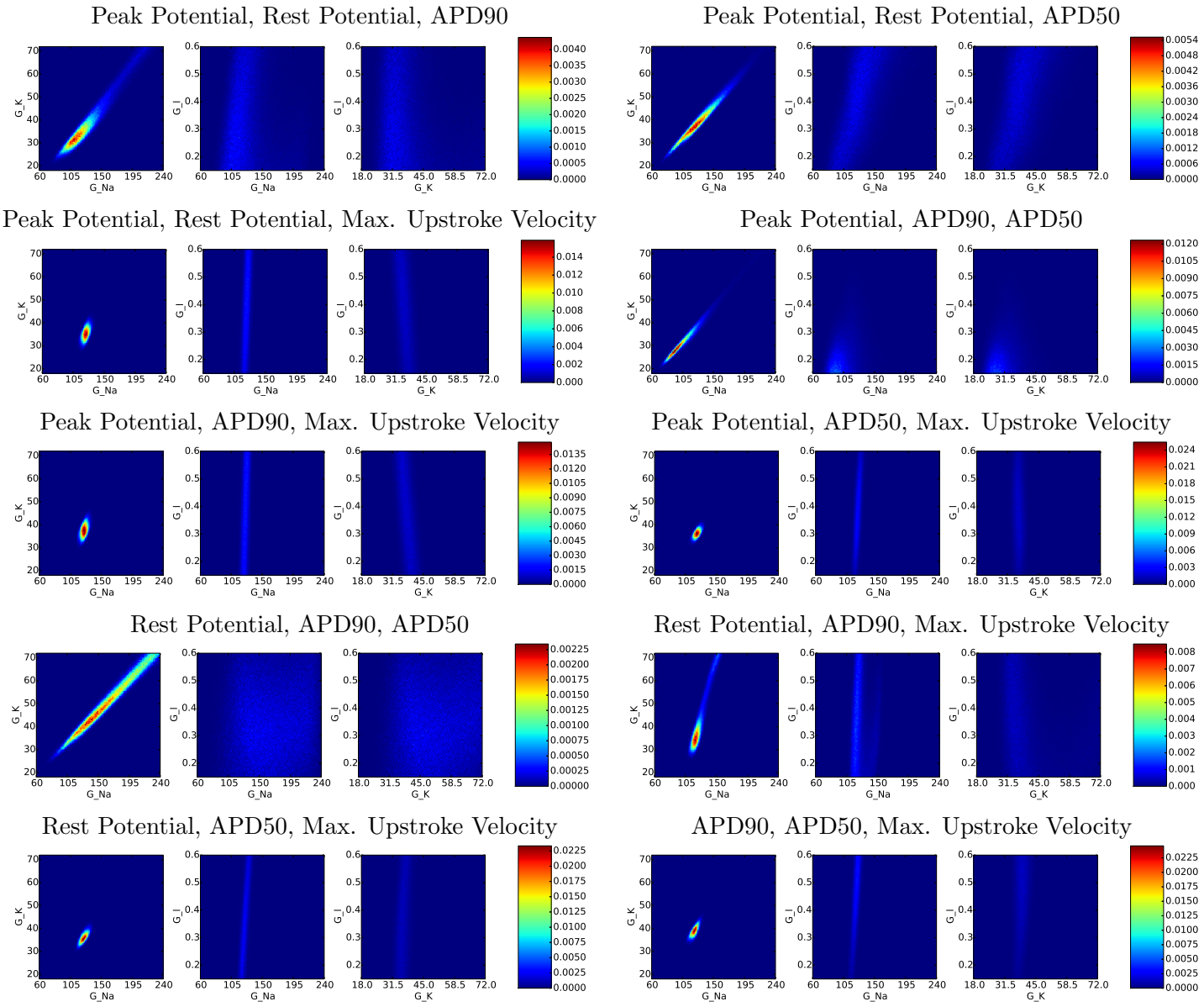

Figure C.3: Biplot visualizations of MCMC posterior distributions for parameters of the Hodgkin-Huxley model (**??**) when employing three summary statistics (indicated by labels above each triptych) observed from a "true" action potential trace. Axes (in units mS/cm^2) span the full uniform prior for each conductance parameter, and the posterior likelihood is indicated by the corresponding color legend.

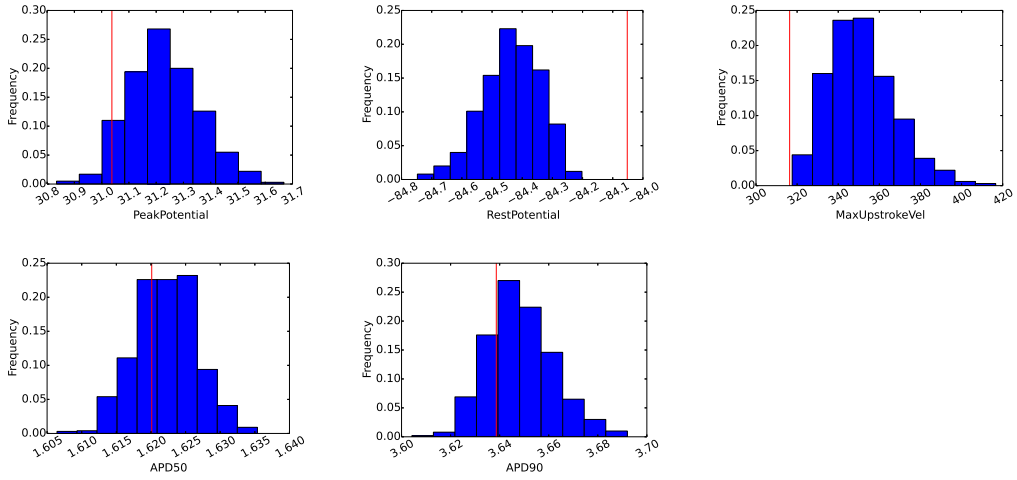

Figure C.4: Distributions of the five summary statistics for the Hodgkin-Huxley model (**??**) computed using 1000 realizations of an action potential voltage trace with additive Gaussian noise. The skew observed in the top row of figures is a result of the max or min functions used when calculating them from a noisy signal, not simply a result of random sampling. Vertical red lines indicate the "true" value of the summary statistic when calculated from the voltage trace in the absence of noise.

References

[1] R. Bardenet. Monte Carlo methods. In T. Delemontex and A. Lucotte, editors, *EDP Sciences*, volume 55. IN2P3 School of Statistics, 2013.

[2] J. Breidt, T. Butler, and D. Estep. A measure-theoretic computational method for inverse sensitivity problems I: Method and analysis. *SIAM Journal on Numerical Analysis*, 49(5):1836–1859, 2011.

[3] T. Butler, D. Estep, and J. Sandelin. A computational measure theoretic approach to inverse sensitivity problems II: a posteriori error analysis. *SIAM Journal on Numerical Analysis*, 50(1):22–45, 2012.

[4] T. Butler, D. Estep, S.J. Tavener, C. Dawson, and J. Westerink. A measure-theoretic computational method for inverse sensitivity problems III: Multiple quantities of interest. *SIAM Journal on Uncertainty Quantification*, pages 1–27, 2014.

[5] T. Butler, D. Estep, S.J. Tavener, T. Wildey, C. Dawson, and L. Graham. Solving stochastic inverse problems using sigma-algebras on contour maps. *arXiv preprint arXiv:1407.3851*, 2014.

[6] P. Del Moral, A. Doucet, and A. Jasra. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing*, 22(5):1009–1020, 2012.