



Perfect MCMC Sampling in Bayesian MRFs for Uncertainty Estimation in Segmentation

Saurabh Garg and Suyash P. Awate^(✉)

Computer Science and Engineering Department,
Indian Institute of Technology (IIT) Bombay, Mumbai, India
suyash@cse.iitb.ac.in

Abstract. Typical segmentation methods produce a single optimal solution and fail to inform about (i) the confidence/*uncertainty* in the object boundaries or (ii) alternate close-to-optimal solutions. To estimate uncertainty, some methods intend to sample segmentations from an associated posterior model using Markov chain Monte Carlo (MCMC) sampling or perturbation models. However, they cannot guarantee sampling from the true posterior, deviating significantly in practice. We propose a novel method that guarantees *exact MCMC sampling*, in finite time, of multi-label segmentations from *generic Bayesian Markov random field* (MRF) models. For exact sampling, we propose Fill's strategy and *extend it to generic MRF models* via a novel *bounding chain* algorithm. Results on simulated data and clinical brain images from 4 classic problems show that our uncertainty estimates gain accuracy over the state of the art.

Keywords: Segmentation · Uncertainty · Hidden MRF · Bayesian EM · MCMC · Perfect/exact sampling · Brain · Multiatlas · Tumor Tissue · Lesion

1 Introduction and Related Work

Accounting for *uncertainty* in automated segmentation results can improve risk analysis in clinical procedures (e.g., neurosurgery [1], radiotherapy [8]) and reliability in clinical diagnosis and studies. Segmentation methods, e.g., using expectation maximization (EM) and hidden Markov random fields (MRFs) or graph cuts, typically produce a single optimal solution, failing to inform about object-boundary uncertainty or alternate close-to-optimal solutions.

For a small class of MRF models that allow segmentation inference via graph cuts, efficient methods exist to exactly estimate label uncertainty [6]. For general MRFs, typical uncertainty estimation methods approximate modeling or sampling. While [3] uses (non-exact) Markov chain Monte Carlo (MCMC) to sample

The authors are grateful for funding from IIT Bombay Seed Grant 14IRCCSG010.

nonparametric curves, [8] uses a Gaussian-process approximation for label distributions. In tumor segmentation, [1] approximates the Gumbel perturbation models in [9] to sample from the underlying Bayesian MRF. For multiatlas segmentation, [2] uses bootstrap resampling to learn nonparametric regression models and error-convergence rates indicating voxelwise uncertainty for a *population* of images (not a specific image). In contrast, we propose the *perfect/exact MCMC* paradigm and a novel perfect-MCMC sampler for *generic Bayesian MRFs*, to estimate uncertainty in multilabel and multiatlas segmentation.

For uncertainty estimation in image registration, while some methods [7] use bootstrap data resampling to approximate the data distribution (unlike the posterior), others use MCMC sampling. Unlike typical MCMC [3] that is only asymptotically exact and can suffer from insufficient burn-in (fixing one very large burn-in for *all* tasks makes computational costs exorbitant), we guarantee exact MCMC in *finite time* and eliminate adhoc heuristics to determine burn-in.

We introduce a new framework for uncertainty estimation in segmentation by relying on *perfect MCMC sampling*, in finite time, from *generic Bayesian MRF* models. We propose to perfect-sample label images: (i) by combining coupling-from-the-past (CFTP) [10] with the bounding-chain (BC) [5] scheme, called CFTP-BC, and, more importantly, (ii) by extending Fill’s algorithm (FA) [4] using the BC scheme, called FA-BC. Results on clinical brain images from 4 applications (segmenting tissues, subcortical structures, tumor, lobes) show that our uncertainty estimates gain accuracy over the state of the art.

2 Methods

We describe our frameworks for perfect MCMC sampling to estimate uncertainty.

MCMC Sampling. Let observed image y , with V voxels, be generated from (i) a hidden label image x that is modeled by MRF X with prior probability mass function (PMF) $P(X)$ and (ii) a likelihood model $P(Y|X)$. MRF X has a neighborhood system $\mathcal{N} := \{\mathcal{N}_v\}_{v=1}^V$, where \mathcal{N}_v is the set of voxels neighboring voxel v . To sample from the posterior $Q(X) := P(X|y)$, MCMC methods construct a Markov chain \mathcal{M} as the MRF sequence $X^1, X^2, \dots, X^t, \dots$, an associated *transition kernel* $K(\cdot, \cdot)$ with $P(X^{t+1}|X^t) := K(X^t, \cdot)$, and *stationary PMF* $Q(X)$. Typically, \mathcal{M} is positive, recurrent, and aperiodic (such a chain is called *ergodic*), thereby having a unique stationary PMF and that PMF being $Q(X)$. \mathcal{M} also typically satisfies *detailed balance*, or *reversibility*, which implies that kernel $K(\cdot, \cdot)$ also applies to the time-reversed chain. The Gibbs sampler is a Metropolis-Hastings MCMC sampler (with an ergodic reversible Markov chain); it iteratively selects a random voxel and draws from its local conditional PMF.

Coupling from the Past (CFTP) for Perfect MCMC Sampling. Gibbs sampler, and typical Metropolis samplers, need to run infinitely long to guarantee draws from the associated stationary PMF $Q(X)$. CFTP [10] theoretically guarantees the sampled state to be from the desired PMF $Q(X)$ by ensuring that any long-running Markov chain, irrespective of its initial state, would have reached

the chosen sampled state, using a specific sequence of interstate-transition maps. CFTP tracks *coupled* parallel chains, one chain started in each possible state of the state space, until all of them *coalesce* to a single state.

Theorem 1. *Propp-Wilson [10]: The CFTP algorithm terminates in finite time and returns a draw from the stationary distribution of the Markov chain.*

Interpretation: Markov chain ergodicity implies, \forall states x , a non-zero probability $> \epsilon > 0$ of reaching x , from any state x' in a finite number of transitions N_x . For a given instance of a sequence of interstate-transition maps (or, equivalently, random numbers) in the Markov chain, coalescence to some state must occur for some finite number of transitions $M \geq \max_x N_x$. Indeed, the probability of coalescence failing to occur $\rightarrow 0$ as $M \rightarrow \infty$. M is almost-surely *finite* because the probability of coalescence in any finite number of transitions is positive. Assume that coalescence occurred when the chain ran from time $t = -M$ to $t = 0$, using a specific sequence of transition maps. A chain running from $-\infty$ to 0 that uses this sequence of transition maps within $[-M, 0]$ reaches the same state at $t = 0$. Because the state reached by a chain running infinitely long is a draw from the stationary PMF $Q(X)$, the coalesced state at $t = 0$ is a draw from $Q(X)$.

For some PMFs $Q(X)$, the Gibbs sampler is *monotonic* [10] where transitions of coupled chains preserve a partial order on the states, and this allows CFTP to simplify parallel-chain tracking to tracking only two chains, each started from one of the extremal states (minimum and maximum) under the partial order. While monotonicity holds for the special case of the ferromagnetic Ising model, it fails to apply to many popular binary-MRF/ Potts models. For general cases, perfect sampling can use the *bounding chain* principle [5] as we propose next.

CFTP with Bounding Chain (CFTP-BC). For Gibbs sampling, CFTP-BC uses the following modified sampler \mathcal{G} to draw label X_v , at each voxel v , from the conditional PMF $P(X_v|x_{-v})$ conditioned on all other label values x_{-v} .

1. Draw label l uniformly from the label set $\mathcal{L} := \{1, \dots, L\}$. Draw $u \sim U(0, 1)$.
2. If $u < P(X_v = l|x_{-v})$, set $X_v := l$ and terminate; otherwise, iterate.

Provably, $\forall l$, the probability of \mathcal{G} terminating with $X_v = l$ is $P(X_v = l|x_{-v})$.

For the Gibbs sampler relying on \mathcal{G} , the bounding chain algorithm [5] efficiently tracks the states of coupled parallel chains (monotone or not). CFTP-BC uses this tracking strategy to detect coalescence. Consider a new kind of a Markov chain $\hat{\mathcal{M}}$ with state space $(2^\mathcal{L})^V$, where $2^\mathcal{L}$ is the set of subsets of \mathcal{L} . For $\hat{\mathcal{M}}$, each state, say, \hat{X} , contains a set of states $X \in \mathcal{L}^V$. $\hat{\mathcal{M}}$ is associated with a state sequence $\hat{X}^1, \hat{X}^2, \dots$ where the transition kernel $\hat{K}(\cdot, \cdot)$ on \hat{X} is defined in terms of the transition kernel $K(\cdot, \cdot)$ acting on each state $X \in \hat{X}$.

Definition 1. *Huber [5]: $\hat{\mathcal{M}}$ is a bounding chain for \mathcal{M} if there exists a coupling between $\hat{\mathcal{M}}$ and \mathcal{M} such that $X_v^t \in \hat{X}_v^t, \forall v, \implies X_v^{t+1} \in \hat{X}_v^{t+1}, \forall v$.*

Consider all coupled parallel chains \mathcal{M} running \mathcal{G} and visiting voxel v at time t . The bounding chain $\hat{\mathcal{M}}$ keeps track of the set $\hat{X}_v \subseteq \mathcal{L}$ of possible labels, at each v , across all chains \mathcal{M} at any given time; it initializes $\hat{X}_v := \mathcal{L}$ and

detects coalescence when $|\dot{X}_v| = 1, \forall v$. Each chain \mathcal{M} has its conditional PMFs $P(X_v | x_{-v})$, dependent on MRF-neighborhood configurations $x_{\mathcal{N}_v}$. For each label l , let the minimum and maximum of conditional probabilities $P(X_v = l | x_{-v})$, over all chains \mathcal{M} , be $P^{\min}(X_v = l | x_{-v})$ and $P^{\max}(X_v = l | x_{-v})$, computed over all possible neighborhood label configurations in the cross-product space $\dot{X}_{w_1} \times \dot{X}_{w_2} \times \dots$ over all $w_i \in \mathcal{N}_v$. Partition the set of all chains \mathcal{M} into equivalence classes, based on possible MRF-neighborhood label values $x_{\mathcal{N}_v}$, within which Gibbs samplers \mathcal{G} behave identically at voxel v . Now do the following at voxel v :

1. In the bounding chain $\dot{\mathcal{M}}$, initialize the set of possible labels $\dot{X}_v := \emptyset$.
2. Draw l uniformly from the label set \mathcal{L} . Draw $u \sim U(0, 1)$.
3. If $u > P^{\max}(X_v = l | x_{-v})$, then *no* chain \mathcal{M} has changed state. So, do nothing.
4. If $u \in [P^{\min}(X_v = l | x_{-v}), P^{\max}(X_v = l | x_{-v})]$, then some of the equivalence classes of chains \mathcal{M} have set $X_v \leftarrow l$. So, insert label l into set \dot{X}_v .
5. If $u < P^{\min}(X_v = l | x_{-v})$, then all chains \mathcal{M} set $X_v \leftarrow l$, indicating “local” coalescence that is a sufficient condition for every chain \mathcal{M} to have undergone *at least one* transition where sampler \mathcal{G} terminated. So, insert l into \dot{X}_v . Exit. $\dot{\mathcal{M}}$ avoids explicitly tracking a possibly exponential number of equivalence classes, but allows a possibly looser bound (larger $|\dot{X}_v|$) resulting from some chains \mathcal{M} running \mathcal{G} multiple times and including *all* sampled labels in \dot{X}_v .
6. Repeat from Step 2.

When, $\forall v$, set \dot{X}_v is a singleton, say, $\{\hat{x}_v\}$, then all Markov chains \mathcal{M} have coalesced to label image \hat{x} that is guaranteed to be a draw from the stationary PMF $Q(X)$. Ergodicity of \mathcal{M} ensures coalescence almost-surely in finite time.

Fill’s Algorithm (FA) for Perfect MCMC Sampling. A limitation of the CFTP strategy proposed in [10], including monotone-chain CFTP [10] and CFTP-BC [5], is that the CFTP running time M and the sampled state \hat{X} are *dependent* variables. M is unbounded whose order of magnitude is typically unknown a priori. So, some states x require a very long run from $-M$ to 0, with large unpredictable M . Impatient users who abort CFTP when M starts becoming large, add bias to the sampled states’ PMF. In contrast, FA [4] makes the sampled state independent of the running time; it relies on acceptance-rejection (AR) sampling. The FA in [4] works only for monotone \mathcal{M} , as below.

1. Choose a random time $T > 0$ and a random label image $X^T := z$.
2. Run a Markov chain \mathcal{M} from $T \rightarrow 0$, with initial $X^T := z$, reaching $X^0 := x$.
3. Let $S^T(x, z)$ be the event that a Markov chain starting at x ran for time T to reach z ; this occurs for some set of pseudo-random number sequences $\mathcal{U}^{x \rightarrow z}$. Let $C^T(z)$ be the event that coupled parallel chains ran for time T and coalesced in z ; this occurs for some set of pseudo-random number sequences $\mathcal{U}' \subseteq \mathcal{U}^{x \rightarrow z}$. With probability $P(C^T(z) | S^T(x, z))$, accept x as a draw from the stationary PMF $Q(X)$ and terminate; otherwise iterate from Step (1).

$P(C^T(z) | S^T(x, z))$ is computationally intractable, but AR decisions can be made by (i) simulating a $\square^{x \rightarrow z} \in \mathcal{U}^{x \rightarrow z}$ to ensure $S^T(x, z)$ occurs and (ii) tracking coupled parallel chains, transitioning as per $\square^{x \rightarrow z}$, to detect if $C^T(z)$ occurs.

Theorem 2. *Fill [4]: Fill’s algorithm, with constrained monotone chains, guarantees that the sampled state is from the stationary PMF $Q(X)$.*

This is true because the underlying AR sampler generates a proposal x from the T -step transition kernel $K^T(z, \cdot)$ and, knowing that $M_z K^T(z, \cdot)$ is an upper bound for the stationary PMF $Q(\cdot)$ for $M_z := Q(z)/P(C^T(z))$, accepts the proposed x with probability $Q(x)/(M_z K^T(z, x))$ that equals $P(C^T(z)|S^T(x, z))$.

Fill’s Algorithm with Bounding Chain (FA-BC). Previous works limit Fill’s algorithm to monotone chains that apply to a very small class of PMFs $Q(X)$; for monotone chains, detecting $C^T(z)$ constrained on $S^T(x, z)$ needs the tracking of only two extremal states. We generalize Fill’s algorithm to generic Bayesian MRFs by efficiently tracking constrained parallel *arbitrary* chains using a *novel constrained bounding chain* algorithm, as follows.

At time t and voxel v , for each label l , let $P^{\min}(X_v^t = l|x_{-v}^t)$ and $P^{\max}(X_v^t = l|x_{-v}^t)$ be defined as before. Let l^* be the label at voxel v for time $t + 1$ along the Markov chain path $x \rightarrow z$. At time t , let $P^*(X_v^t = l^*|x_{-v}^t)$ be the label probability conditioned on neighboring labels for the path $x \rightarrow z$. Clearly, $P^{\min}(X_v^t = l|x_{-v}^t) \leq P^*(X_v^t = l^*|x_{-v}^t) \leq P^{\max}(X_v^t = l|x_{-v}^t)$. Initialize $t := 0$, $x^0 := x$.

1. At time t , do the following at each voxel v :
 - (a) In the bounding chain \mathcal{M} , initialize the set of possible labels $\mathring{X}_v := \emptyset$.
 - (b) Draw l uniformly from the label set \mathcal{L} .
 - (c) If $l \neq l^*$, draw $u \sim U(P^*(X_v^t = l^*|x_{-v}^t), 1)$; otherwise draw $u \sim U(0, 1)$. This sampling strategy simulates $\Pi^{x \rightarrow z} \sim \mathcal{U}^{x \rightarrow z}$, ensuring that x^t transitions to x^{t+1} on the path $x \rightarrow z$, thereby leading to $S^T(x, z)$. The next steps track parallel coupled chains to detect if $C^T(z)$ occurs for $\Pi^{x \rightarrow z}$.
 - (d) If $u > P^{\max}(X_v^t = l|x_{-v}^t)$, then *no* chain \mathcal{M} changes state. Go to Step 1b.
 - (e) If $u \in [P^{\min}(X_v^t = l|x_{-v}^t), P^{\max}(X_v^t = l|x_{-v}^t)]$, then some chains \mathcal{M} accept label l . Insert l into \mathring{X}_v . Go to Step 1b.
 - (f) If $u < P^{\min}(X_v^t = l|x_{-v}^t)$, then all chains \mathcal{M} set $X_v = l$. Insert l into \mathring{X}_v . Go to Step 1a to process a new voxel.
2. Increment t by 1. If $t < T$, repeat Step 1. If $t = T$ and coalescence has occurred, i.e., $|\mathring{X}_v| = 1, \forall v$, then accept the initial x as a draw from $Q(X)$.

Theorem 3. *Our modification of the Fill’s algorithm, with constrained bounding chain, guarantees that the sampled state is from the stationary PMF $Q(X)$.*

Proof. We show that our random number generation scheme in Step 1c ensures $S^T(x, z)$ by simulating a $\Pi^{x \rightarrow z} \in \mathcal{U}^{x \rightarrow z}$. At time t and voxel v , let E^* be the event that, for the chain going from $x \rightarrow z$, the label at voxel v at time $t + 1$ is l^* . Let the $x \rightarrow z$ chain’s *unconstrained* modified Gibbs sampler be \mathcal{G}^* and the label probabilities be $P^*(X_v^t = l|x_{-v}^t)$. For E^* to occur, \mathcal{G}^* accepted label l^* in some iteration i . In any iteration, \mathcal{G}^* picked some l and some random u . Given E^* : if \mathcal{G}^* picked an $l \neq l^*$, then u must have been within $[P^*(X_v^t = l^*|x_{-v}^t), 1]$; otherwise u could have been anywhere within $[0, 1]$. Now consider *parallel coupled* chains, one starting at each possible state, running sampler \mathcal{G} for T transition steps. At iteration i , if \mathcal{G} picks $l \neq l^*$, then \mathcal{G} must pick u within $[P^*(X_v^t = l^*|x_{-v}^t), 1]$

because, otherwise, the chain started at x can incorrectly accept $l \neq l^*$ and E^* can fail to occur. At iteration i , if \mathcal{G} picks $l = l^*$, then \mathcal{G} can pick u within $[0, 1]$, leading to a non-zero probability for the chain started at x accepting l^* and leading to E^* . Steps 1d–1f track all chains, as in CFTP-BC, to detect $C^T(z)$ for the chosen $\square^{x \rightarrow z}$. The result then follows from Theorem 2. \square

Exact Sampling to Estimate Uncertainty in Segmentation. We apply our FA-BC perfect-MCMC sampler to estimate uncertainty in Bayesian segmentation that models the *label image* prior as a *hidden MRF* X with the Potts model. We use FA-BC (i) during parameter estimation via EM, in the E step for Monte Carlo sampling label image X from its posterior, and (ii) after parameter estimation, to estimate uncertainty by sampling label maps from the posterior, given optimal parameters, and measuring their variability per voxel. We apply to 4 classic segmentation problems, in brain magnetic resonance imaging (MRI), with different likelihood models: (i) EM segmentation of tissues with mild lesions, with a Gaussian mixture model (GMM) for the intensities. (ii) EM segmentation of tumor, with a 2-component GMM for the tumor and non-tumor intensity patches on multimodal MRI, (iii) multiatlas segmentation of subcortical structures, and (iv) multiatlas segmentation of 4 lobes. Both (iii) and (iv) use a basic voxelwise nonparametric label-likelihood model for proof-of-concept, as follows. Let the multiatlas database $\mathcal{D} := \{z^j, s^j\}_{j=1}^J$ have template MRI images z^j paired with label images s^j . At voxel i , the observed-image patch $y_{\mathcal{N}_i}$ has likelihood $P(y_{\mathcal{N}_i} | X_i = l, \mathcal{D}) := \sum_{j=1}^J \mathbf{1}_l(s_i^j) G(\check{y}_{\mathcal{N}_i}; \check{z}_{\mathcal{N}_i}^j, \sigma^2 \mathbf{I}) / \sum_{j=1}^J \mathbf{1}_l(s_i^j)$ where $\mathbf{1}_l(a) = 1$ if $l = a$ (0 otherwise), \mathbf{I} is the identity matrix, σ^2 the Gaussian kernel variance, and $\check{y}_{\mathcal{N}_i}$ and $\check{z}_{\mathcal{N}_i}^j$ are normalized patches with mean 0 and variance 1.

3 Results and Discussion

We show results on simulated data and on 4 classic brain-MRI analyses for 3 methods: (i) ours, (ii) approximate Gumbel perturbation model (aGPM) [1], (iii) Gibbs sampler with limited burn-in. For posterior-sampled label images (sample size 10^3), we compute mean and standard deviation (SD) per voxel (for the multi-category case, we generalize SD by square-root of unalikeability).

Validation on Simulated Data. The aGPM approximation of the true sampling PMF (in [9], which is intractable) can be severe (Fig. 1(a)–(c)), leading to a strong bias in the empirical mean estimate near edges (Fig. 1(d)). Our empirical mean estimate (Fig. 1(d)) is much closer to ground truth.

Results on Clinical Brain MRI. For many segmentation tasks, typical maximum-a-posteriori (MAP) segmentations can be very misleading by failing to expose regions with high uncertainty, e.g., (i) in subcortical structures, the hippocampus tail region (Fig. 2), (ii) in tumor, the edema regions (Fig. 3), (iii) in tissues, regions with mild lesions in white matter (Fig. 4). In these cases, the empirical means and SDs resulting from posterior-sampled label images are far more informative than the MAP estimate. However, in *all* these cases, unlike

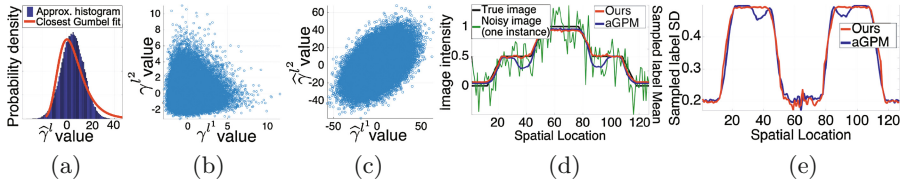


Fig. 1. Validation on Simulated Data: 128-voxel 1D image, 2 labels. Differences between *ideal* Gumbel perturbations γ in [9] (intractable for label-image sampling) and their tractable *approximations* $\hat{\gamma}$ in aGPM [1]: (a) For a label image l , empirical histogram for $\hat{\gamma}^l := \sum_{i=1}^{128} \gamma_i^{l_i}$, as per aGPM’s notation, is almost Gaussian (central limit theorem), deviating significantly from Gumbel. (b)–(c) For label images l^1 and l^2 , scatter between aGPM draws $\hat{\gamma}^{l^1}$ and $\hat{\gamma}^{l^2}$ (both using *same* sample for γ_i) deviates from that between Gumbel draws γ^{l^1} and γ^{l^2} . (d)–(e) Sample mean and SD (voxelwise) of label images drawn from hidden-MRF posterior for aGPM [1] and our FA-BC sampler, averaged over multiple simulated image instances with different noise instances.

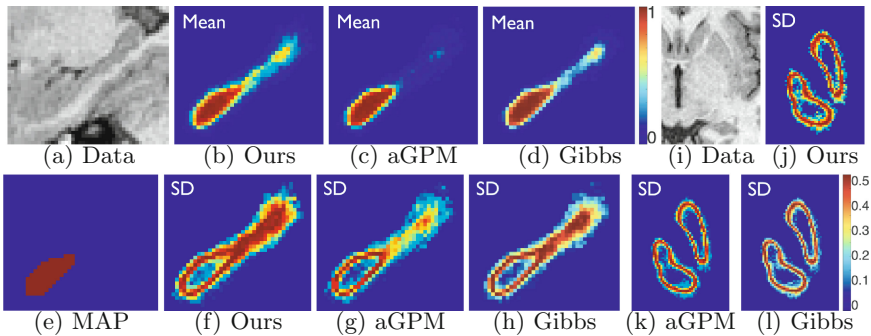


Fig. 2. Clinical brain MRI: Multiatlas segmentation, subcortical structures

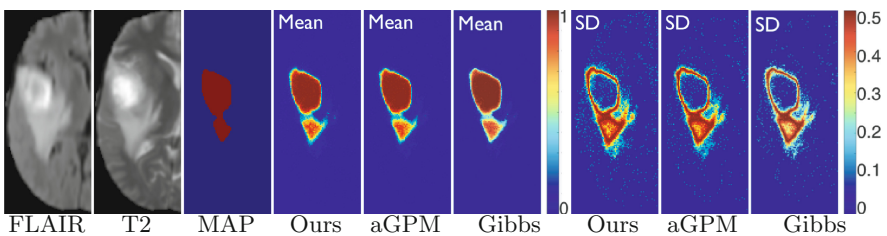


Fig. 3. Clinical multimodal brain MRI: Tumor segmentation.

our approach, both aGPM and Gibbs significantly underestimate the label SDs. Our FA-BC clearly improves over CFTP-BC [5] (Fig. 4) when, unlike our FA-BC, for large values of the smoothness parameter (say, β) in the Potts-MRF model, CFTP-BC takes far too many transition steps T and computation times, or virtually fails to terminate (for $\beta > 0.66$). For multiatlas hippocampus

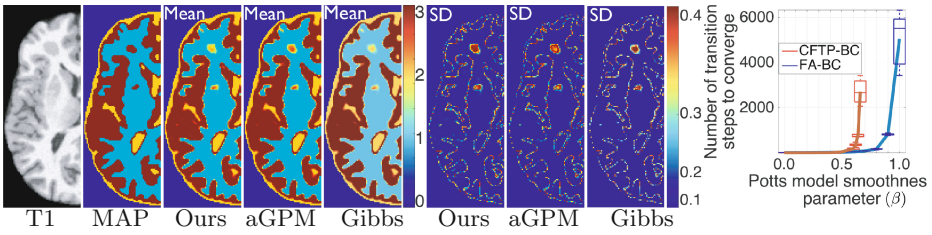


Fig. 4. Clinical brain MRI, simulated mild lesion: Tissue segmentation.

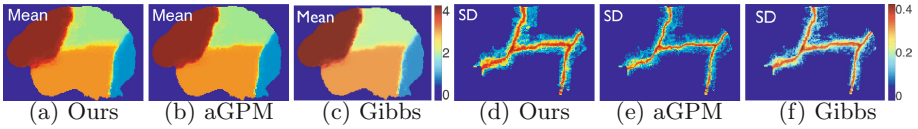


Fig. 5. Clinical brain MRI: Multiatlas segmentation of lobes.

segmentation (Fig. 2(b)–(d)), compared to our method, aGPM and Gibbs severely underestimate the label means as well. For tissue segmentation (Fig. 4), within the mild lesion with intensities between those of gray and white matter, our label mean is halfway between the label values of gray and white matter and indicates a greater uncertainty. In contrast, aGPM (or Gibbs) labels the lesion more confidently as gray (or white) matter, which is undesirable. For multiatlas multimodal-MRI tumor segmentation (Fig. 3), tissue segmentation with mild lesions (Fig. 4), and lobe segmentation (Fig. 5), aGPM and Gibbs severely underestimate label SDs, unlike our method that theoretically and practically guarantees sampled label images from the true posterior.

Computation Times: Gibbs’s convergence time varies severely with the MRF model and the data, making it very difficult to predict burn-in. With a safe-side burn-in of 5000, as per the plot in Fig. 4, our FA-BC is 10–20× faster.

Conclusion. We introduced a new framework for uncertainty estimation in segmentation relying on *perfect MCMC sampling* of label images from their posteriors, defined using a *generic MRF model*. Our FA-BC extended Fill’s algorithm to use a bounding-chain scheme, *improving theoretically and practically* over the state of the arts for (i) uncertainty estimation, e.g., aGPM and naive Gibbs, and (ii) perfect sampling, e.g., CFTP-BC, for analyzing simulated data and clinical brain MRI (segmenting tissues, subcortical structures, tumor, lobes).

References

1. Alberts, E., et al.: Uncertainty quantification in brain tumor segmentation using CRFs and random perturbation models. In: IEEE International Symposium on Biomedical Imaging, pp. 428–431 (2016)
2. Awate, S., Whitaker, R.: Multiatlas segmentation as nonparametric regression. *IEEE Trans. Med. Imaging* **33**(9), 1803–17 (2014)
3. Fan, A.C., Fisher, J.W., Wells, W.M., Levitt, J.J., Willsky, A.S.: MCMC curve sampling for image segmentation. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007. LNCS, vol. 4792, pp. 477–485. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75759-7_58
4. Fill, J.: An interruptible algorithm for perfect sampling via Markov chains. *Ann. Appl. Prob.* **8**(1), 131–62 (1998)
5. Huber, M.: Perfect sampling using bounding chains. *Ann. Appl. Prob.* **14**(2), 734–753 (2004)
6. Kohli, P., Torr, P.: Measuring uncertainty in graph cut solutions. *Comp. Vis. Image Underst.* **112**, 30–8 (2008)
7. Kybic, J.: Bootstrap resampling for image registration uncertainty estimation without ground truth. *IEEE Trans. Image Process.* **19**(1), 64–73 (2010)
8. Le, M., Unkelbach, J., Ayache, N., Delingette, H.: Sampling image segmentations for uncertainty quantification. *Med. Image Anal.* **34**, 42–51 (2016)
9. Papandreou, G., Yuille, A.: Perturb-and-MAP random fields: using discrete optimization to learn and sample from energy models. In: International Conference on Computer Vision (2011)
10. Propp, J., Wilson, D.: Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Struct. Algorithms* **9**(1), 223–52 (1996)