Supplementary Materials for "A Bayesian hidden Potts mixture model for analyzing lung cancer pathology images"

QIWEI LI[1], XINLEI WANG[2], FAMING LIANG[3], FALIU YI[1], YANG XIE[1], ADI GAZDAR[4,5], GUANGHUA XIAO[1,*]

[1]Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas, U.S.A.

[2]Department of Statistics, Southern Methodist University, Dallas, Texas, U.S.A.

[3]Department of Statistics, Purdue University, West Lafayette, Indiana, U.S.A.

[4]Department of Pathology, UT Southwestern Medical Center, Dallas, Texas, U.S.A.

[5]Hamon Center for Therapeutic Oncology Research, UT Southwestern Medical Center, Dallas, Texas, U.S.A.

**email:* guanghua.xiao@utsouthwestern.edu*

1. A Modified Hidden Potts Model

In this section, we discuss a modified hidden Potts model, as a complement to the hidden Potts model that is presented in Section 2.2 of the article. We consider the type of cell i depends on both of its adjacent spins and its location. For example, Equation (2.4) can be modified to

$$\Pr(x_i', y_i', z_i = q|\boldsymbol{P}, d, \tau) \propto \exp\left(d \sum_{\{(l,w):l\leqslant x_i'<l+1, w\leqslant y_i'<w+1\}} I(p_{lw} = q)\right) \exp\left(-\tau \min_{\{(l,w):p_{lw}=q\}}\{\text{dist}((x_i', y_i'), (l, w)\}\right),$$

where the additional term is an exponential decay with parameter τ and the Euclidean distance between cell i and its nearest spin q denoted by $\min_{\{(l,w):p_{lw}=q\}}\{\text{dist}((x_i', y_i'), (l, w)\}$. A larger value of τ makes a cell type more likely to be the same with its closest spin, while a smaller value leads to Equation (2.4), which assumes that the cell types are independent and identically distributed conditional on their four adjacent spins. Supplementary Figure S3 shows how the choices of τ influence the spatial distribution of the cell types. The location-dependent hidden Potts model is interesting, especially when the lattice size is considerably small. However, compared to the original version, it is more computationally intensive.

2. Full Details of MCMC Algorithm

Update of d: We update the projection parameter d by using a random walk Metroplis-Hastings (RWMH) algorithm. We first propose a new d^* from $\text{Ga}(d^2/\tau_d, d/\tau_d)$ and then accept the proposed value with probability $\min(1, r_d)$ and the Hastings ratio is

$$r_d = \frac{\prod_{i=1}^n \Pr(x_i', y_i', z_i|\boldsymbol{P}, d^*)}{\prod_{i=1}^n \Pr(x_i', y_i', z_i|\boldsymbol{P}, d)} \frac{\text{Ga}(d^*; a_d, b_d)}{\text{Ga}(d; a_d, b_d)},$$

where the form of $\Pr(x_i', y_i', z_i|\boldsymbol{P}, d)$ is given by Equation (2.4) in the main manuscript. Note that the proposal density ratio equals to 1 for this random walk Metropolis update.

Update of $\boldsymbol{P}$: We first update each p_{lw} in the AOI that corresponds to $\delta_{lw} = 1$ by using Gibbs sampling. Then we repeat the same step to update each p_{lw} in the background area

that corresponds to $\delta_{lw} = 0$. The probability of assigning class q to p_{lw} is proportional to

$$\pi(p_{lw} = q|\delta_{lw} = 1, \cdot) \propto \Pr(p_{lw} = q, \boldsymbol{P}_{-l,-w}|\boldsymbol{\theta}) \prod_{\{i:l-1\leqslant x_i'<l+1, w-1\leqslant y_i'<w+1\}} \Pr(x_i', y_i', z_i|\boldsymbol{P}, d)$$

$$\propto \sum_{(l',w')\in \mathrm{Nei}_{(l,w)}} \sum_{q=1}^{Q}\sum_{q'=1}^{Q} \theta_{qq'} I(p_{lw} \neq p_{l'w'}, p_{lw} = q, p_{l'w'} = q', \delta_{l'w'} = 1)$$

$$\prod_{\{i:l-1\leqslant x_i'<l+1, w-1\leqslant y_i'<w+1\}} \Pr(x_i', y_i', z_i|\boldsymbol{P}, d),$$

or

$$\pi(p_{lw} = q|\delta_{lw} = 0, \cdot) \propto \sum_{(l',w')\in \mathrm{Nei}_{(l,w)}} \sum_{q=1}^{Q}\sum_{q'=1}^{Q} \theta_{0qq'} I(p_{lw} \neq p_{l'w'}, p_{lw} = q, p_{l'w'} = q', \delta_{l'w'} = 0)$$

$$\prod_{\{i:l-1\leqslant x_i'<l+1, w-1\leqslant y_i'<w+1\}} \Pr(x_i', y_i', z_i|\boldsymbol{P}, d),$$

where $\Pr(x_i', y_i', z_i|\boldsymbol{P}, d)$ is given by Equation (2.4) in the main manuscript.

Update of $\boldsymbol{\theta}$: We update each $\theta_{qq'}, q = 1, \ldots, Q-1, q' = q+1, \ldots, Q, q \neq q'$ by using a double Metropolis-Hastings (DMH) algorithm. We first propose a new $\theta_{qq'}^*$ from $\mathrm{N}(\theta_{qq'}, \tau_\theta^2)$ and then simulate an auxiliary variable $\boldsymbol{P}^*$ starting from $\boldsymbol{P}$ based on the new $\boldsymbol{\theta}^*$, where all the elements are the same as $\boldsymbol{\theta}$, excluding $\theta_{qq'}$. Note that, from $\boldsymbol{P}$ to $\boldsymbol{P}^*$, we only need to update those spins that correspond to $\delta_{lw} = 1$. The proposed value $\theta_{qq'}^*$ will be accepted with probability $\min(1, r_\theta)$ and the Hastings ratio is

$$r_\theta = \frac{\Pr(\boldsymbol{P}^*|\boldsymbol{\theta})}{\Pr(\boldsymbol{P}|\boldsymbol{\theta})} \frac{\Pr(\boldsymbol{P}|\boldsymbol{\theta}^*)}{\Pr(\boldsymbol{P}^*|\boldsymbol{\theta}^*)} \frac{\mathrm{N}(\theta_{qq'}^*; \mu, \sigma^2)}{\mathrm{N}(\theta_{qq'}; \mu, \sigma^2)},$$

where the form of $\Pr(\boldsymbol{P}|\boldsymbol{\theta})$ is given by Equation (2.2) in the main manuscript. As a result, the normalizing constants can be canceled. Note that the proposal density ratio equals to 1 for this random walk Metropolis update on $\theta_{qq'}$.

Update of $\boldsymbol{\theta}_0$: We update each $\theta_{0qq'}, q = 1, \ldots, Q-1, q' = q+1, \ldots, Q, q \neq q'$ by using a DMH algorithm. We first propose a new $\theta_{0qq'}^*$ from $\mathrm{N}(\theta_{0qq'}, \tau_\theta^2)$ and then simulate an auxiliary variable $\boldsymbol{P}^*$ starting from $\boldsymbol{P}$ based on the new $\boldsymbol{\theta}_0^*$, where all the elements are the same as $\boldsymbol{\theta}_0$, excluding $\theta_{0qq'}$. Note that, from $\boldsymbol{P}$ to $\boldsymbol{P}^*$, we only need to update those spins that correspond to $\delta_{lw} = 0$. The proposed value $\theta_{0qq'}^*$ will be accepted with probability $\min(1, r_{\theta_0})$ and the

Hasting ratio is

$$r_{\theta_0} = \frac{\Pr(\boldsymbol{P}^*|\boldsymbol{\theta}_0)}{\Pr(\boldsymbol{P}|\boldsymbol{\theta}_0)} \frac{\Pr(\boldsymbol{P}|\boldsymbol{\theta}_0^*)}{\Pr(\boldsymbol{P}^*|\boldsymbol{\theta}_0^*)} \frac{\mathrm{N}(\theta_{0qq'}^*; \mu_0, \sigma_0^2)}{\mathrm{N}(\theta_{0qq'}; \mu_0, \sigma_0^2)},$$

where the form of $\Pr(\boldsymbol{P}|\boldsymbol{\theta})$ is given by Equation (2.2) in the main manuscript. As a result, the normalizing constants can be canceled. Note that the proposal density ratio equals to 1 for this random walk Metropolis update on $\theta_{0qq'}$.

Update of $\boldsymbol{\Delta}$: We update $\boldsymbol{\Delta}$ via an *add-delete* algorithm. Note that we perform the following steps 20 times within each iteration. We first select an l from $\{1, \ldots, L\}$ and a w from $\{1, \ldots, W\}$ at random and then change its value $\delta_{lw}^* = 1 - \delta_{lw}$. Then an auxiliary variable $\boldsymbol{P}^*$ starting from $\boldsymbol{P}$ based on the new $\boldsymbol{\Delta}^*$ is simulated. All the elements of $\boldsymbol{\Delta}^*$ are the same as $\boldsymbol{\Delta}$, excluding δ_{lw}. Note that, from $\boldsymbol{P}$ to $\boldsymbol{P}^*$, we update those spins that correspond to $\delta_{lw}^* = 1$ based on $\boldsymbol{\theta}$ and update those spins that correspond to $\delta_{lw}^* = 0$ based on $\boldsymbol{\theta}_0$. We finally accept the proposed values with probability $\min(1, r_\delta)$ and the Hastings ratio is

$$r_\delta = \frac{\Pr(\boldsymbol{P}^*|\boldsymbol{\Delta}, \boldsymbol{\theta}_0, \boldsymbol{\theta})}{\Pr(\boldsymbol{P}|\boldsymbol{\Delta}, \boldsymbol{\theta}_0, \boldsymbol{\theta})} \frac{\Pr(\boldsymbol{P}|\boldsymbol{\Delta}^*, \boldsymbol{\theta}_0, \boldsymbol{\theta})}{\Pr(\boldsymbol{P}^*|\boldsymbol{\Delta}^*, \boldsymbol{\theta}_0, \boldsymbol{\theta})} \frac{\Pr(\delta_{lw}^*|\boldsymbol{\Delta}^*)}{\Pr(\delta_{lw}|\boldsymbol{\Delta})},$$

where the forms of $\Pr(\boldsymbol{P}|\boldsymbol{\Delta}, \boldsymbol{\theta}_0, \boldsymbol{\theta})$ and $\Pr(\delta_{lw}|\boldsymbol{\Delta})$ are given by Equation (2.8) and Equation (2.9) in the main manuscript, respectively. As a result, the normalizing constants in Equation (2.8) can be canceled.

3. Scalability Test

The scalability test was conducted in `R` with `Rcpp` package on a Mac PC with 2.60GHz CPU and 16GB memory. Within each MCMC iteration, we need to update d, $\boldsymbol{P}$, $\boldsymbol{\theta}$, $\boldsymbol{\theta_0}$, and $\boldsymbol{\Delta}$. The total number of parameters is $1 + LW + Q(Q-1)/2 + Q(Q-1)/2 + LW$. Therefore, the time complexity is at least $O(LW + Q^2)$. For small numbers of class, Q, it reduces to $O(LW)$.

We considered three settings of Q (i.e $Q = 2$, $Q = 3$, and $Q = 4$) to generate the simulated

datasets. All the generative models were based on a 50-by-50 lattice. We considered the true structure of $\boldsymbol{\Delta}$ the same as scenario 2, which is shown in Figure 2(d) of the article. The AOI is composed of four rectangles with equivalent size. For the case of $Q = 2$, we set $\theta = 0.5$ and $\theta_0 = -1.0$; For the case of $Q = 3$, we set $\boldsymbol{\theta} = (\theta_{12}, \theta_{13}, \theta_{23}) = (0.5, 0.7, 1.0)$ and $\boldsymbol{\theta}_0 = (\theta_{012}, \theta_{013}, \theta_{023}) = (-1.0, -0.5, -1.5)$; For $Q = 4$, we set $\boldsymbol{\theta} = (\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}) = (0.1, 0.3, 0.5, 0.7, 0.9, 1.1)$ and $\boldsymbol{\theta}_0 = (\theta_{012}, \theta_{013}, \theta_{014}, \theta_{023}, \theta_{024}, \theta_{034}) = (-1.0, -0.75, -0.5, -0.25, -1.0, -1.5)$. The hidden spins $\boldsymbol{P}$ were simulated using the Gibbs sampler, running $100,000$ iterations with random starting configurations. The observed points $(\boldsymbol{x}, \boldsymbol{y})$ were generated from a homogeneous Poisson point process with a constant intensity $\lambda = 4$ over the planar space $[1, 50]^2$. Their classes $\boldsymbol{z}$ were assigned according to adjacent spins. Specifically, for point i, its class was drawn from a multinomial distribution $\text{Mn}(\phi_1, \ldots, \phi_Q)$. The parameters $(\phi_1, \ldots, \phi_Q)$ were inferred from a Dirichlet distribution $\text{Dir}(0.1 + \tilde{n}_{i1}, \ldots, 0.1 + \tilde{n}_{iQ})$, where $\tilde{n}_{iq}$ denotes the number of adjacent spins that belong to class q. We repeated the above steps to generate 10 independent datasets for each setting of Q.

For each dataset, we applied the proposed model with the same hyperparameter and algorithm settings as described in Section 4.1 of the article and the lattice size chosen from 20×20, 30×30, 40×40, or 50×50. We then ran an MCMC chain with $10,000$ iterations. Figure S12 shows the runtime of the MCMC algorithm as a function of the lattice size $L \times W$. Our observations are three-fold. First, the algorithm is fast for the small lattice size. For example, the runtime of fitting a 20×20 lattice size is as short as 60, 90, and 120 seconds for the three cases, respectively. Second, the actual runtime increases approximately quadratically in L or W, which is consistent with the theoretical lower bound $O(LW)$. Third, the runtime difference is relatively small among different settings of Q. This is because the time complexity is dominated by the lattice size $L \times W$, especially when Q is small. We further

fit a linear regression for the runtime (per $10,000$ iterations) versus $(LW)^2$ and Q, which has an R^2 of 0.976. It also indicates the runtime depends mainly on $(LW)^2$, followed by Q in order of significance. The estimated model was time $= -640.6 + 0.0004459(LW)^2 + 231.7Q$. This analysis implied that, computationally speaking,our method can be applied to datasets with a large number of Q.

References

Yi, F., Huang, J., Yang, L., Xie, Y., and Xiao, G. (2017). Automatic extraction of cell nuclei from H&E-stained histopathological images. *Journal of Medical Imaging*, **4**(2), 027502.

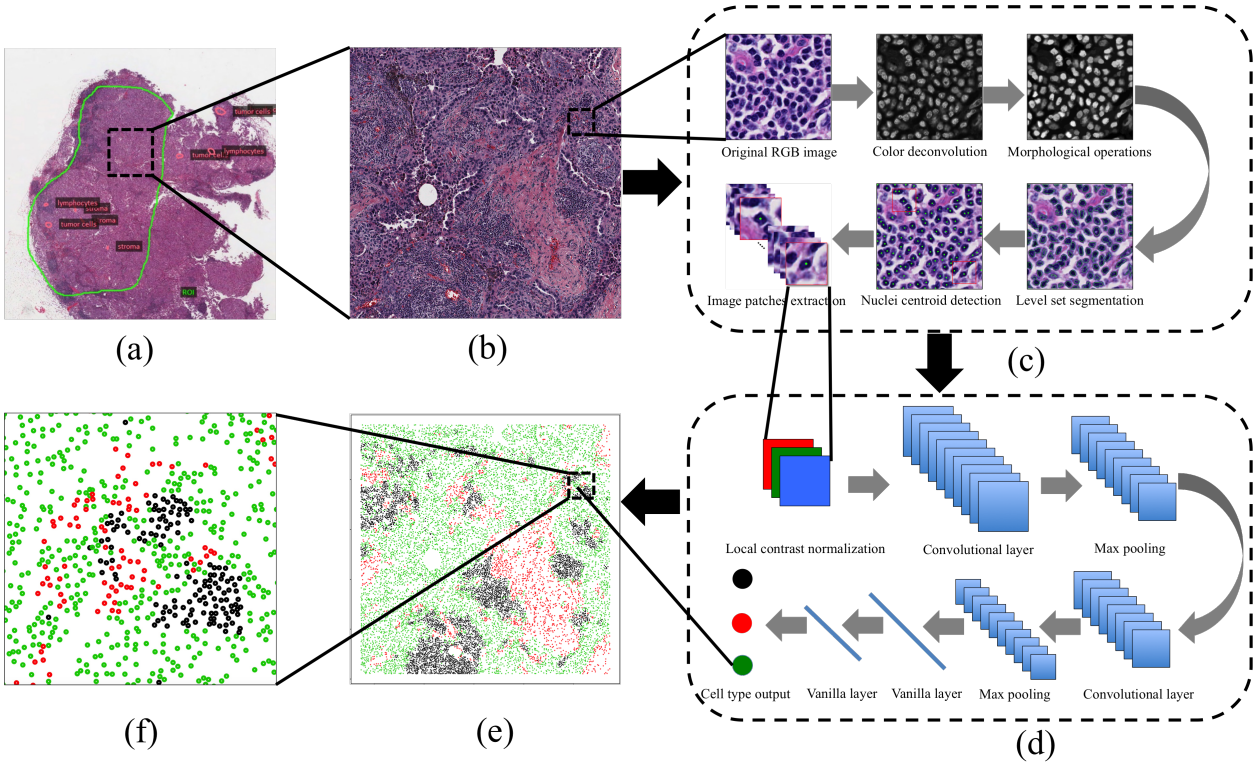


Figure S1: Illustration of the ConvPath pipeline: (a) A part of the whole pathological imaging slide (the median size is $24,244 \times 19,261$ pixel) from a lung cancer patient and the tumor region (within the green border) manually labeled by a pathologist; (b) A randomly chosen sample image (5000×5000 pixel) in the tumor region for image analysis; (c) The image processing pipeline for cell detection and segmentation (Yi *et al.*, 2017); (d) The convolutional neural network (CNN) that predicts the type of each individual cell (https://qbrc.swmed.edu/projects/cnn/); (e) The cell distribution map corresponding to the sample image as shown in (b), where the cell locations were detected by (c) and the cell types were predicted by (d), with black, red, and green representing lymphocyte, stromal, and tumor cells, respectively; (f) The cell distribution map corresponding to the sub-region of the sample image.

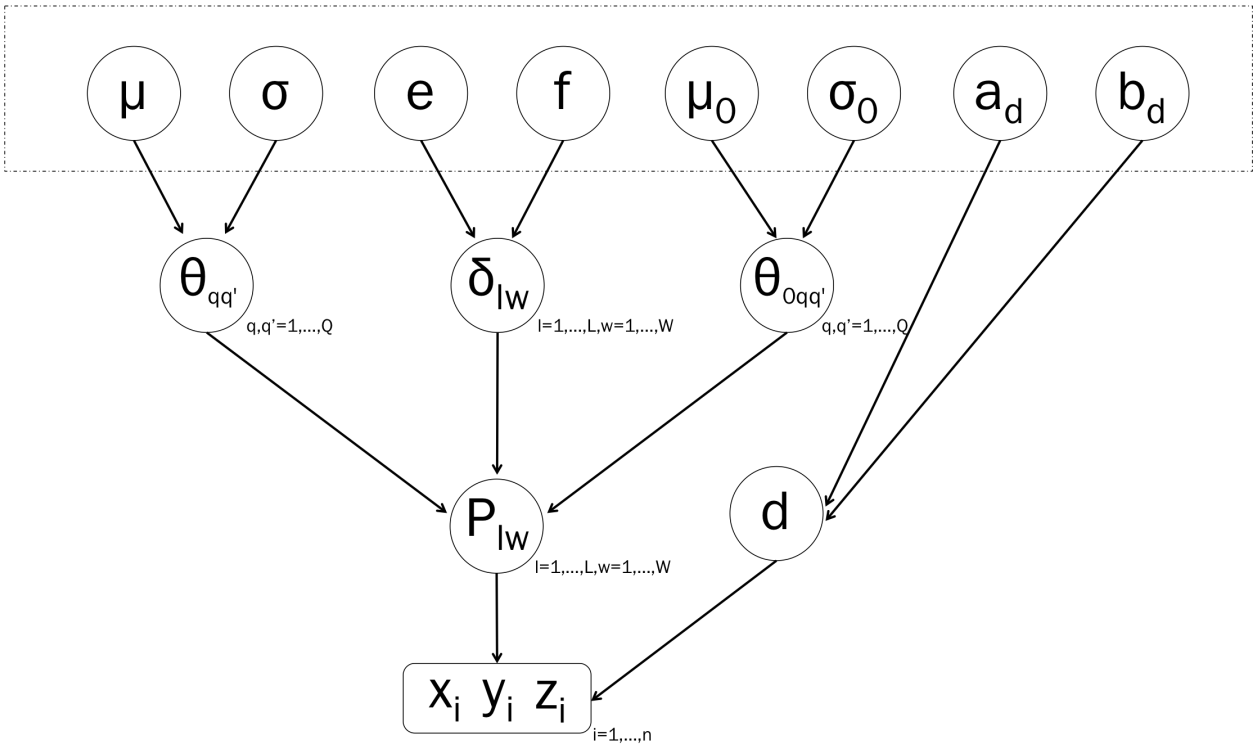
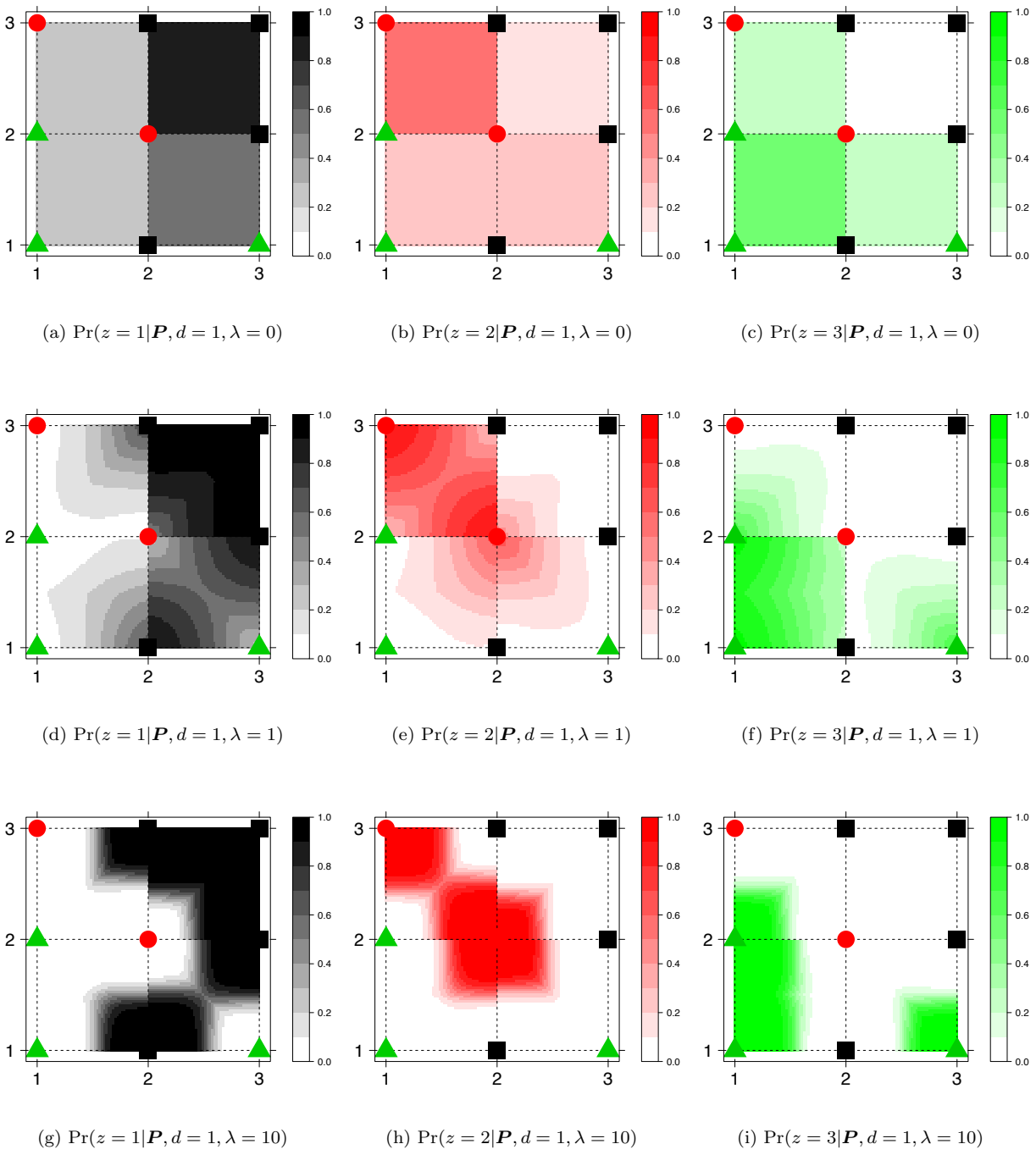


Figure S2: The graphical formulation of the proposed Bayesian hidden Potts mixture model. Each node in a circle refers to a parameters of the model. Node in a rectangle is observable data. Circle nodes in the dashed block are fixed hyperparameters. The link between two nodes represents a direct probabilistic dependence.

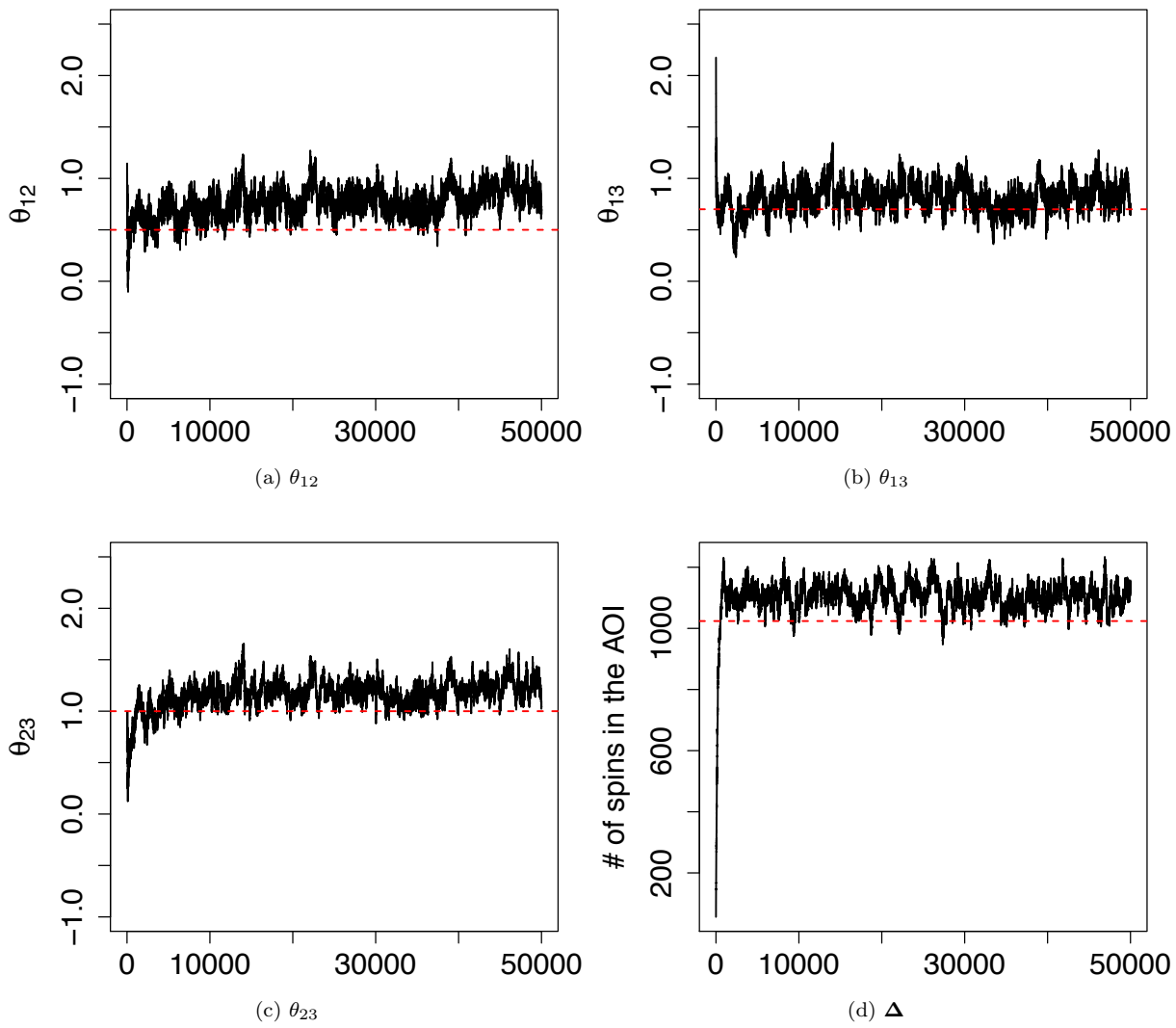


(a) $\Pr(z = 1|\boldsymbol{P}, d = 1, \lambda = 0)$ (b) $\Pr(z = 2|\boldsymbol{P}, d = 1, \lambda = 0)$ (c) $\Pr(z = 3|\boldsymbol{P}, d = 1, \lambda = 0)$

(d) $\Pr(z = 1|\boldsymbol{P}, d = 1, \lambda = 1)$ (e) $\Pr(z = 2|\boldsymbol{P}, d = 1, \lambda = 1)$ (f) $\Pr(z = 3|\boldsymbol{P}, d = 1, \lambda = 1)$

(g) $\Pr(z = 1|\boldsymbol{P}, d = 1, \lambda = 10)$ (h) $\Pr(z = 2|\boldsymbol{P}, d = 1, \lambda = 10)$ (i) $\Pr(z = 3|\boldsymbol{P}, d = 1, \lambda = 10)$

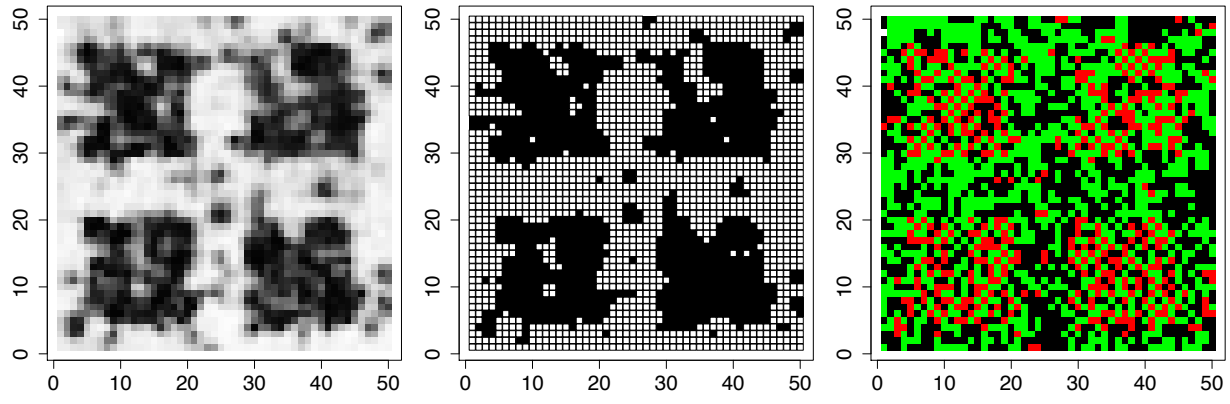
Figure S3: The heatmaps of the probability $\Pr(z = q|\boldsymbol{P}, d, \lambda)$ under different choices of d and λ, in the illustrated 3-by-3 lattice as shown in Figure 1 of the article. The points represent the hidden spins in the lattice. Black, red, and green indicate class $q = 1$, 2, and 3, respectively.



(a) θ_{12} (b) θ_{13}

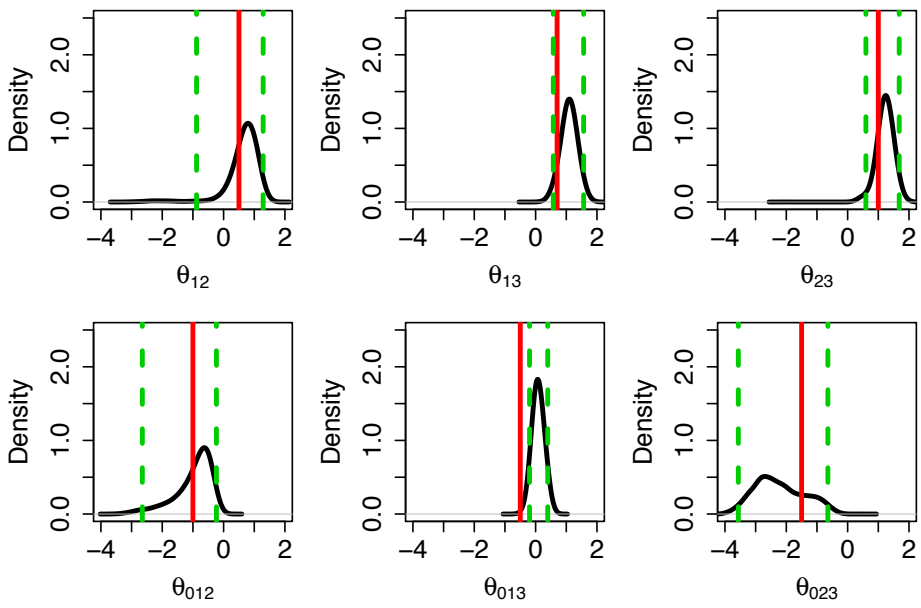
(c) θ_{23} (d) $\boldsymbol{\Delta}$

Figure S4: A simulated dataset generated from the log Gaussian Cox process and $\boldsymbol{\Delta}$ shown in Figure 2(d) of the article (scenario 2): The trace plots of MCMC samples of the interaction parameters (a) θ_{12}, (b) θ_{13}, (c) θ_{23}, and (d) the number of spins in the AOI (i.e. $\sum_{l=1}^{L}\sum_{w=1}^{W}\delta_{lw}$). The red horizontal lines indicate the true values.

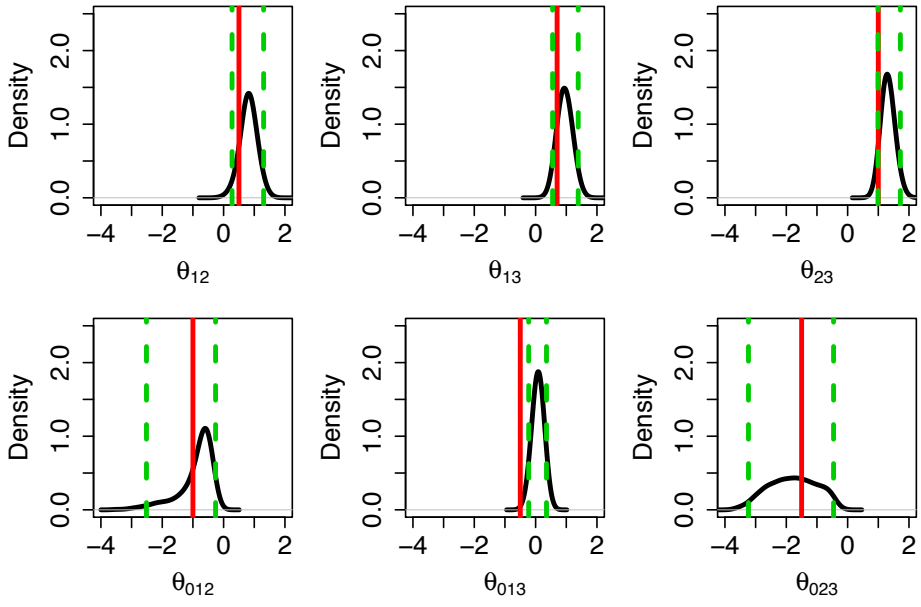


(a) θ_{12} (b) θ_{13} (c) θ_{23}

Figure S5: A simulated dataset generated from the log Gaussian Cox process and $\boldsymbol{\Delta}$ shown in Figure 2(d) of the article (scenario 2): (a) The heatmap of the marginal posterior probabilities $\Pr(\delta_{lw} = 1|\cdot)$'s; (b) The map of estimated $\hat{\boldsymbol{\Delta}}$'s corresponding to the median model when choosing $c = 0.5$; (c) The estimated hidden spins $\hat{\boldsymbol{P}}_{\text{MP}}$.

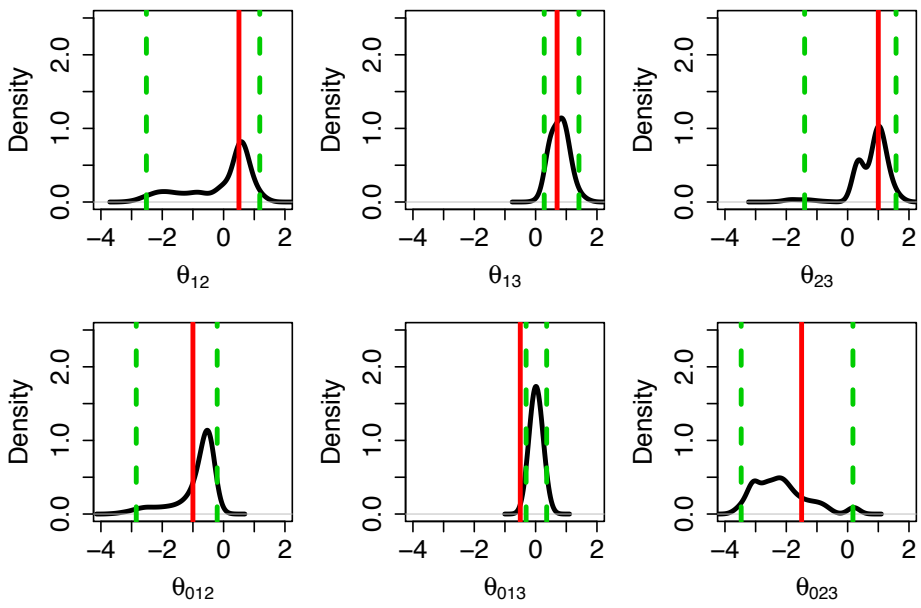


(a) Scenario 1

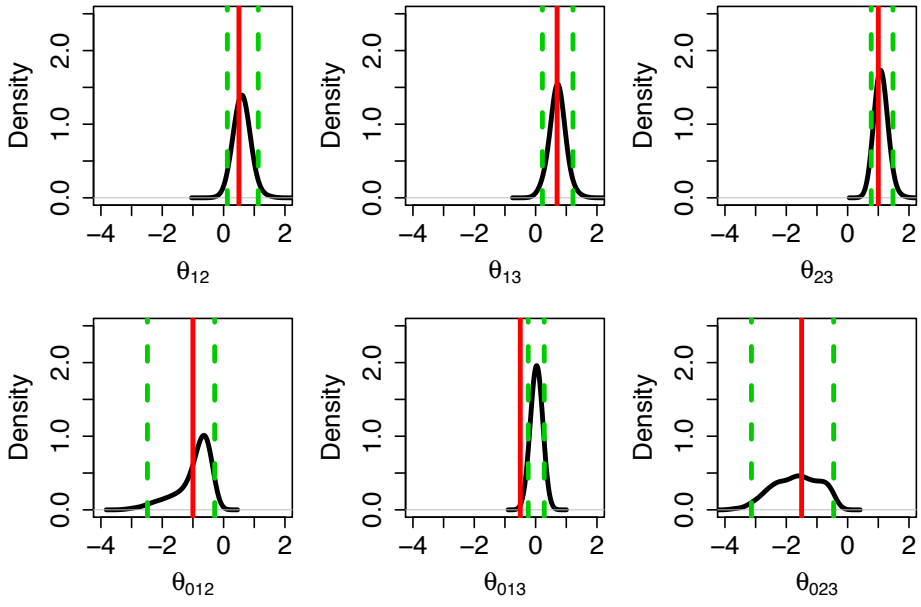


(b) Scenario 2

Figure S6: Simulated datasets generated from the homogeneous Poisson process: The density plots of MCMC samples of the interaction parameters θ_{12}, θ_{13}, θ_{23}, θ_{012}, θ_{013}, and θ_{023}, under (a) scenario 1 and (b) scenario 2. The red vertical lines indicate the true values and the intervals between two green dashed lines indicate the 95% credible intervals obtained by the proposed method.

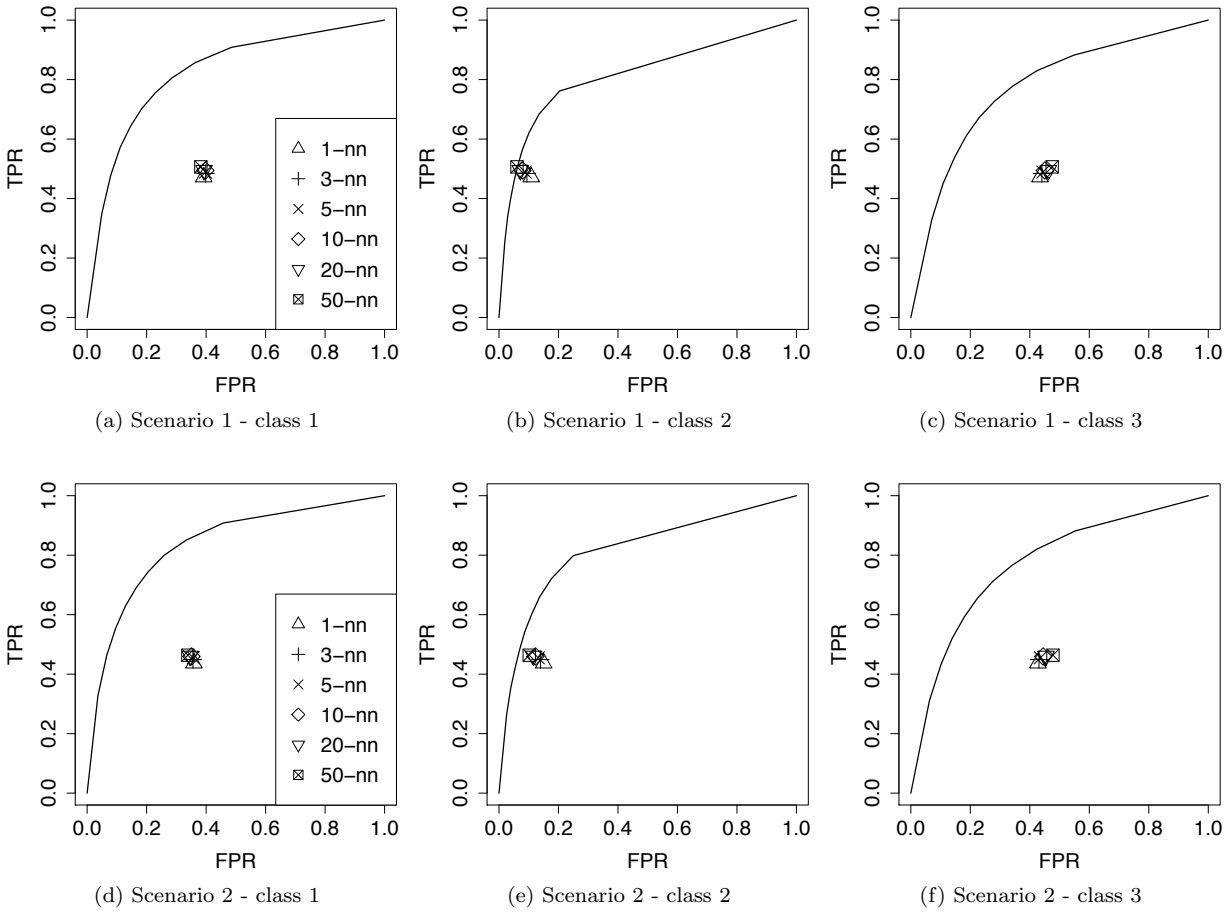


(a) Scenario 1



(b) Scenario 2

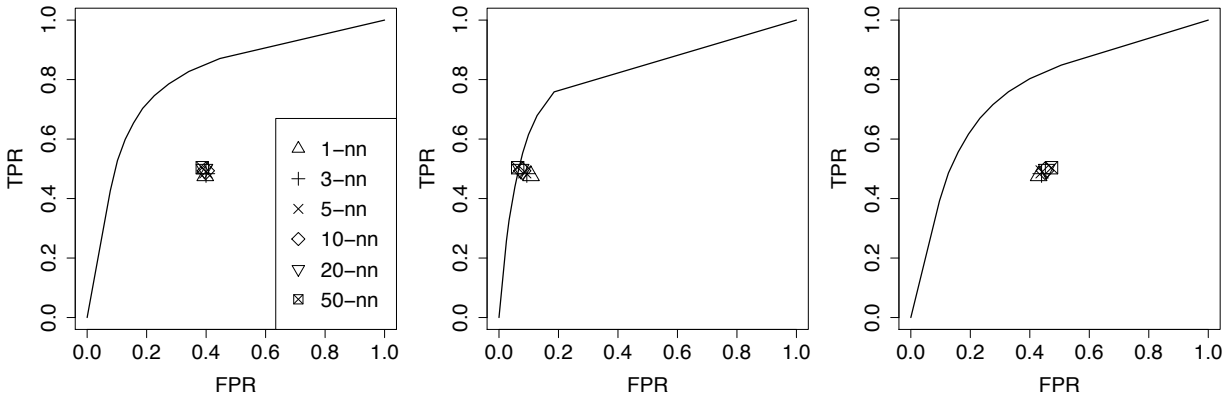
Figure S7: Simulated datasets generated from the log Gaussian Cox process: The density plots of MCMC samples of the interaction parameters θ_{12}, θ_{13}, θ_{23}, θ_{012}, θ_{013}, and θ_{023}, under (a) scenario 1 and (b) scenario 2. The red vertical lines indicate the true values and the intervals between two green dashed lines indicate the 95% credible intervals obtained by the proposed method.



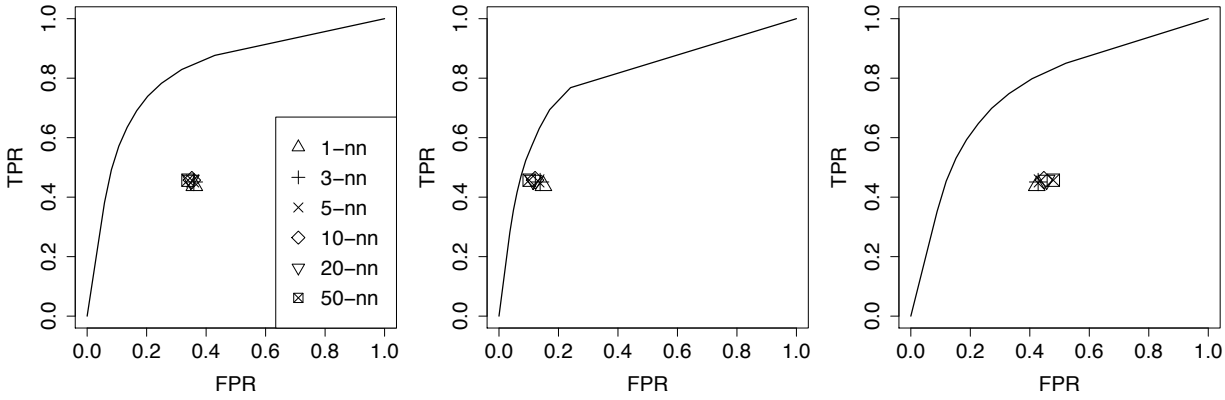
(a) Scenario 1 - class 1 (b) Scenario 1 - class 2 (c) Scenario 1 - class 3

(d) Scenario 2 - class 1 (e) Scenario 2 - class 2 (f) Scenario 2 - class 3

Figure S8: Simulated datasets generated from the homogeneous Poisson process: The average ROC curves for different values of the threshold on the posterior probabilities of inclusion on $p_{wl} = q, w = 1, \ldots, W, l = 1, \ldots, L$, for each scenario and class q. The points under the ROC curves indicate the average false positive rates and true positive rates by assigning each spin p_{wl} based on its k-nearest neighbors of cells.

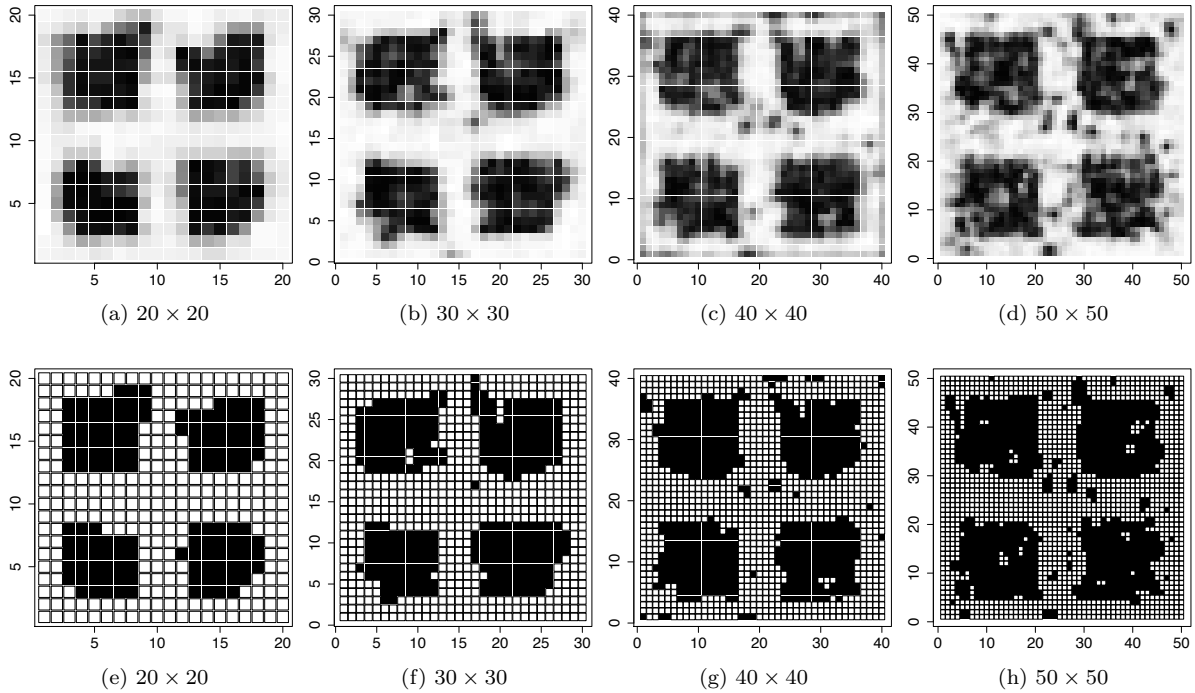


(a) Scenario 1 - class 1 (b) Scenario 1 - class 2 (c) Scenario 1 - class 3



(d) Scenario 2 - class 1 (e) Scenario 2 - class 2 (f) Scenario 2 - class 3

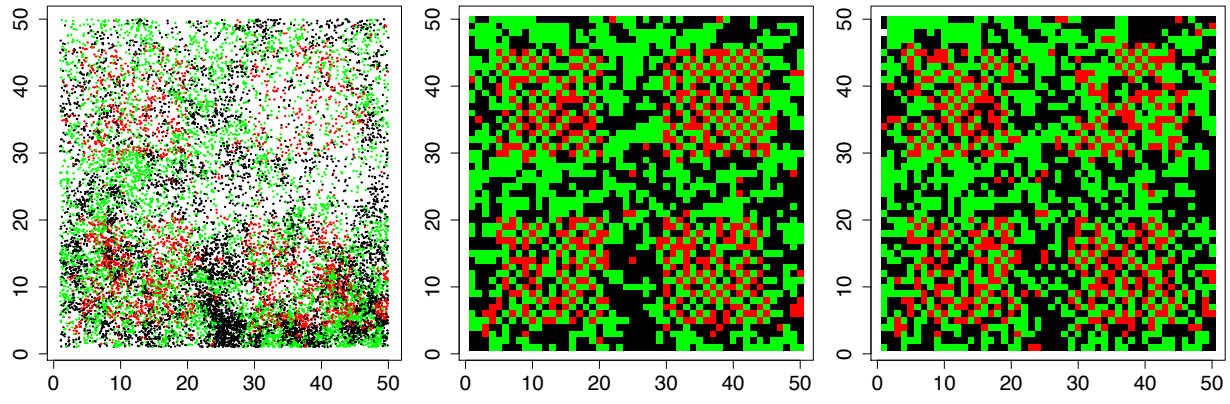
Figure S9: Simulated datasets generated from the log Gaussian Cox process: The average ROC curves for different values of the threshold on the posterior probabilities of inclusion on $p_{wl} = q, w = 1, \ldots, W, l = 1, \ldots, L$, for each scenario and class q. The points under the ROC curves indicate the average false positive rates and true positive rates by assigning each spin p_{wl} based on its k-nearest neighbors of cells.



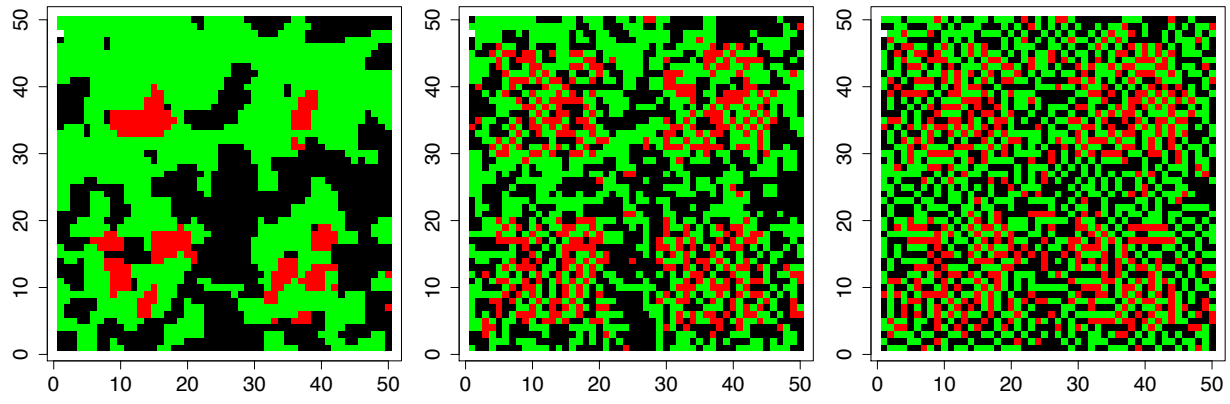
(a) 20×20 (b) 30×30 (c) 40×40 (d) 50×50

(e) 20×20 (f) 30×30 (g) 40×40 (h) 50×50

Figure S10: Sensitivity analysis on the choices of lattice size, on a simulated dataset generated from the homogeneous Poisson process and $\mathbf{\Delta}$ shown in Figure 2(d) of the article (scenario 2): (a), (b), (c), and (d) The heatmaps of the marginal posterior probabilities $\Pr(\delta_{lw} = 1|\cdot)$'s obtained by models with different lattice sizes; (e), (f), (g), and (h) The maps of estimated $\hat{\mathbf{\Delta}}$'s corresponding to the median models when choosing $c = 0.5$.

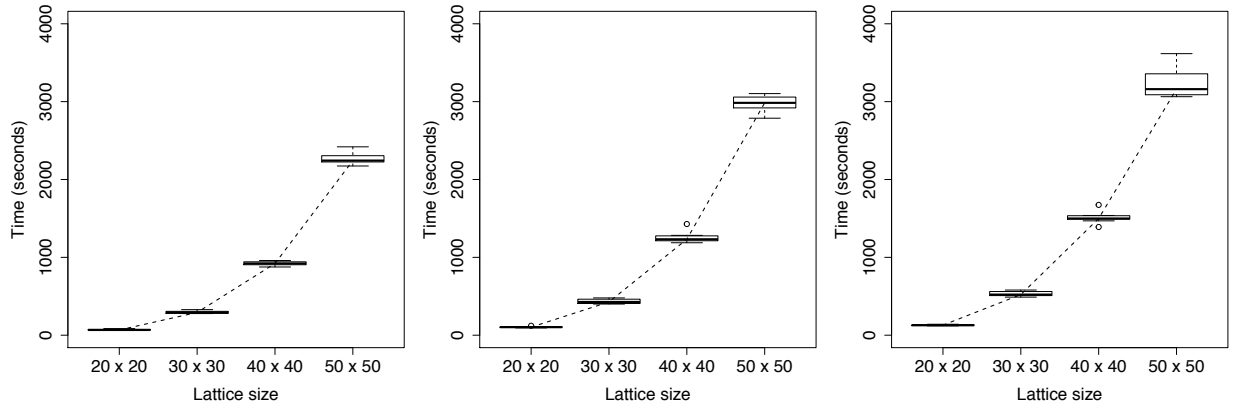


(a) The observed dot distribution map (b) The true map of the hidden spins $\boldsymbol{P}$ (c) $d \sim \text{Ga}(0.001, 0.001)$



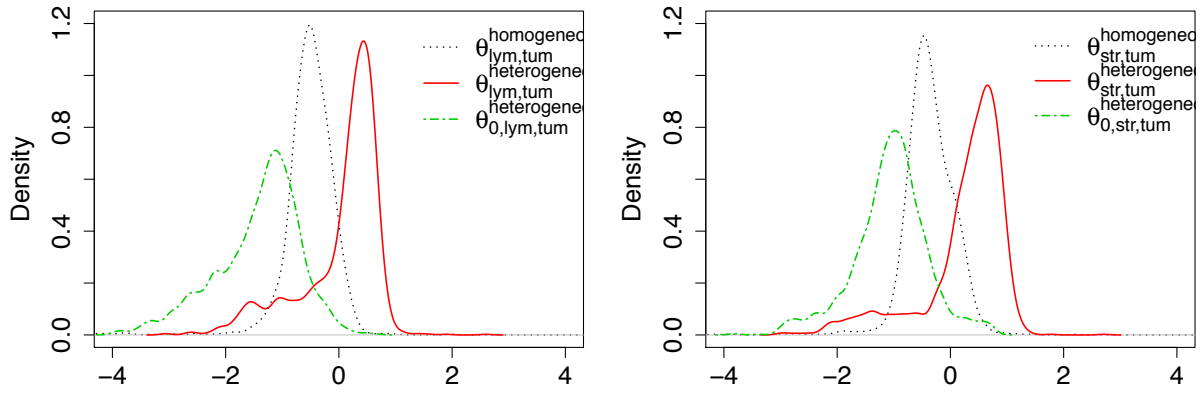
(d) $d = 0.1$ (e) $d = 1$ (f) $d = 10$

Figure S11: Sensitivity analysis on the choices of the projection parameter d, on a simulated dataset generated from the log Gaussian Cox process and $\boldsymbol{\Delta}$ shown in Figure 2(d) of the article (scenario 2): (a) The observed dot distribution map; (b) The true map of the 50-by-50 hidden spins $\boldsymbol{P}$; (c) The map of the estimated 50-by-50 hidden spins $\hat{\boldsymbol{P}}_{\text{MP}}$ by placing a weakly informative prior on d; (c), (d), and (e): The maps of the estimated 50-by-50 hidden spins $\hat{\boldsymbol{P}}_{\text{MP}}$ by fixing the value of d to 0.1, 1, and 10, respectively.



(a) $Q = 2$ (b) $Q = 3$ (c) $Q = 4$

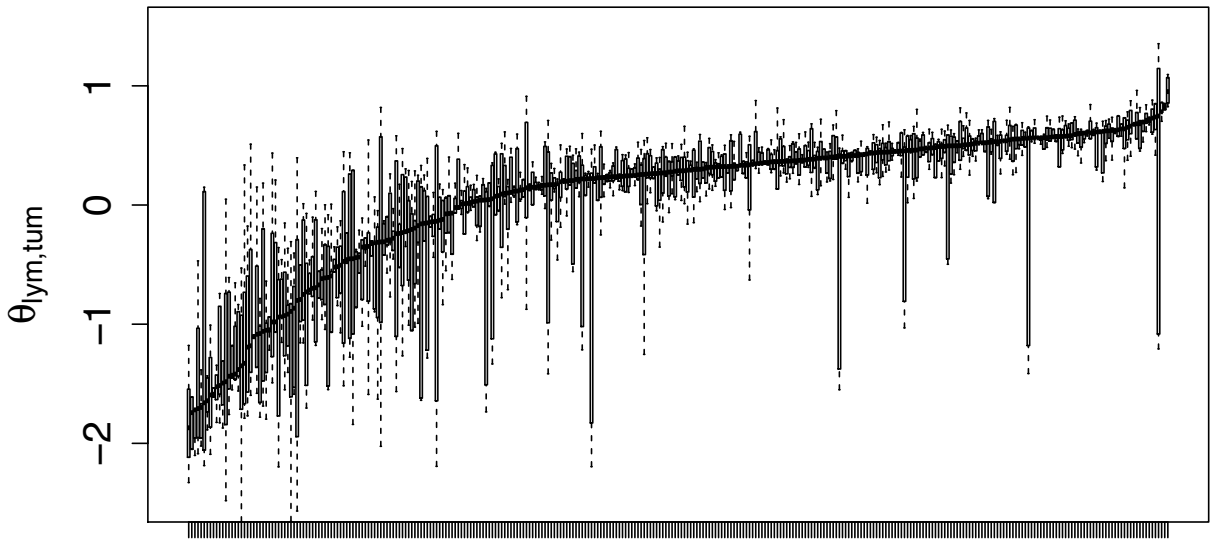
Figure S12: The boxplots of the runtime in seconds (per $10,000$ MCMC iterations) in terms of the lattice size $L \times W$, over multiple simulated datasets generated from the homogeneous Poisson process, $\mathbf{\Delta}$ shown in Figure 2(d) of the article (scenario 2), and the number of classes (a) $Q = 2$, (b) $Q = 3$, and (c) $Q = 4$.



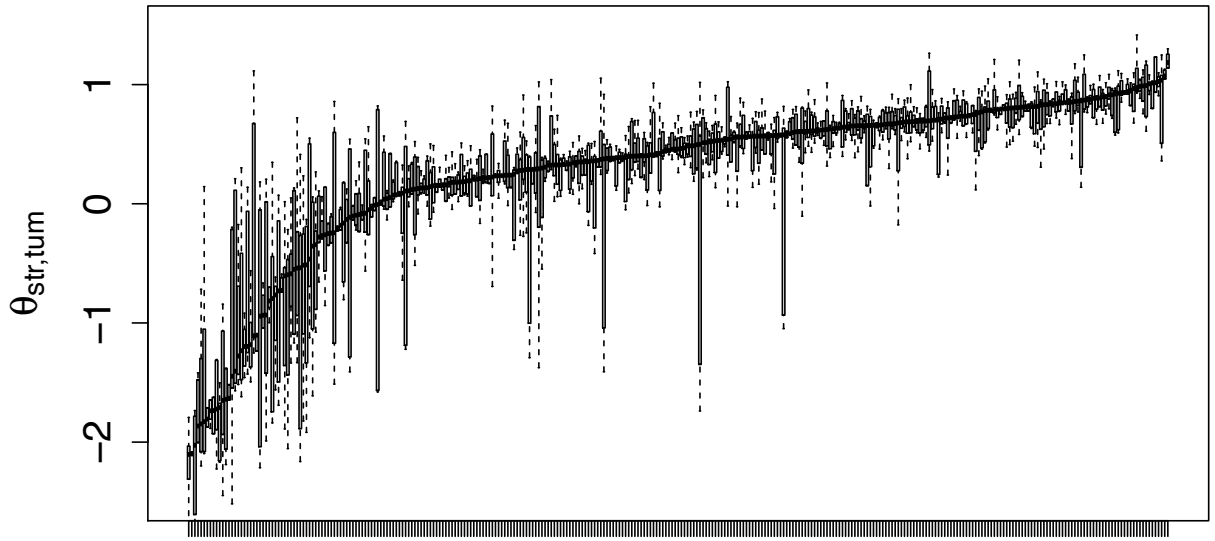
(a) Lymphocyte and tumor

(b) Stroma and tumor

Figure S13: Lung cancer case study: (a) The density plots of $\hat{\theta}_{\text{lym,tum}}$ (red) and $\hat{\theta}_{0,\text{lym,tum}}$ (green) by fitting the $1,585$ sample images to the hidden Potts mixture model, and $\hat{\theta}_{\text{lym,tum}}$ (black) by fitting the datasets to the homogeneous hidden Potts model; (b) The density plots of $\hat{\theta}_{\text{str,tum}}$ (red) and $\hat{\theta}_{0,\text{str,tum}}$ (green) by fitting the $1,585$ sample images to the hidden Potts mixture model, and $\hat{\theta}_{\text{lym,tum}}$ (black) by fitting the datasets to the homogeneous hidden Potts model.

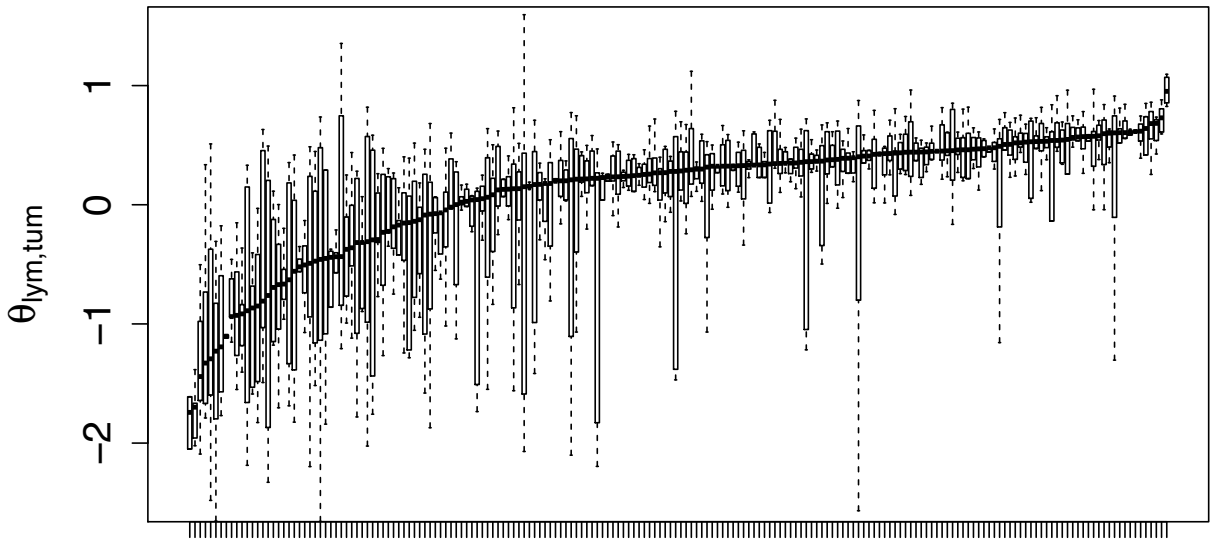


(a) Lymphocyte and tumor

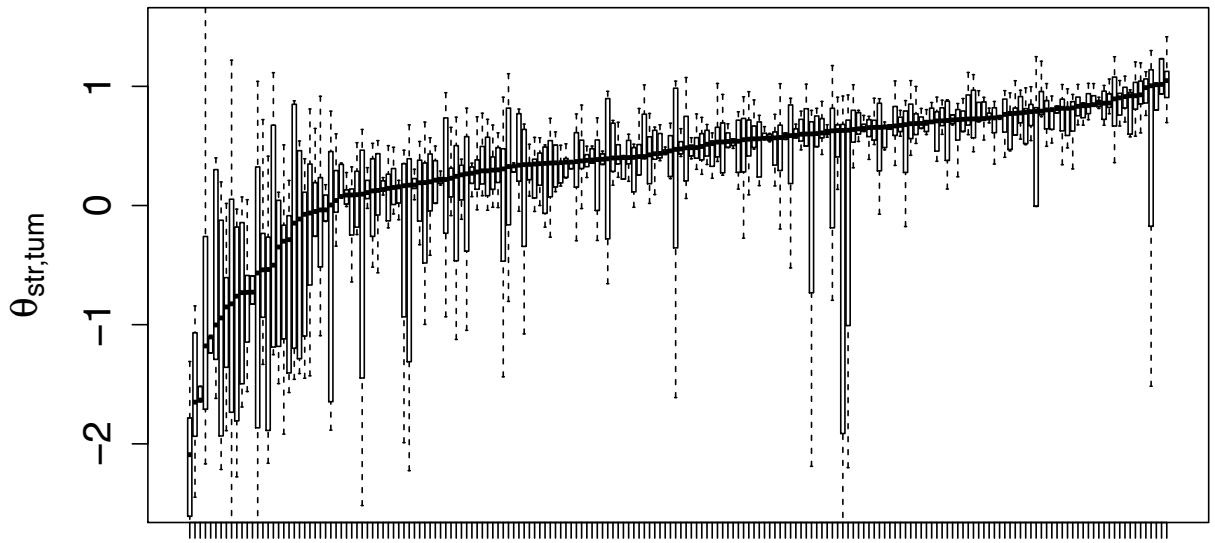


(b) Stroma and tumor

Figure S14: Lung cancer case study: The boxplots of (a) $\hat{\theta}_{\text{lym,tum}}$ and (b) $\hat{\theta}_{\text{str,tum}}$ for each of the 317 pathological imaging slides, sorted by the median values.



(a) Lymphocyte and tumor



(b) Stroma and tumor

Figure S15: Lung cancer case study: The boxplots of (a) $\hat{\theta}_{\text{lym,tum}}$ and (b) $\hat{\theta}_{\text{str,tum}}$ for each of the 205 lung cancer patients from the National Lung Screening Trial (NLST), sorted by the median values.

Table S1: Suggested values for hyperparameters e and f of the MRF prior to avoid the phase transition problem.

Expected prop. of the AOI $\exp(e)/(1+\exp(e))$	e	Maximum allowed value of $f_{\max}$
1%	-4.60	2.30
2%	-3.89	1.90
5%	-2.94	1.40
10%	-2.20	1.10
15%	-1.73	0.80
20%	-1.39	0.80
25%	-1.10	0.70
30%	-0.85	0.50
35%	-0.62	0.50
40%	-0.41	0.30

Table S2: Means and standard deviations (in parentheses) of the areas under the ROC curves (AUCs) on the posterior probabilities of inclusion on $\boldsymbol{\Delta}$ and the FPRs, recalls (TPRs), precisions, and F-1 scores of the median models, when choosing $e = -2.94$ and $f = 1.4$, over multiple simulated datasets generated from each point process and each setting of $\boldsymbol{\Delta}$.

	Homogeneous Poisson process		Log Gaussian Cox process	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
AUC	0.924 (0.171)	0.950 (0.009)	0.903 (0.090)	0.948 (0.010)
FPR	0.111 (0.164)	0.119 (0.028)	0.128 (0.077)	0.126 (0.030)
Recall (TPR)	0.841 (0.138)	0.883 (0.023)	0.804 (0.120)	0.889 (0.015)
Precision	0.941 (0.127)	0.915 (0.018)	0.937 (0.047)	0.911 (0.019)
F-1 score	0.887 (0.135)	0.898 (0.015)	0.862 (0.090)	0.900 (0.010)

Table S3: Sensitivity analysis on the choices of hyperparameters e and f of the MRF prior: Means and standard deviations (in parentheses) of the precisions, recalls (TPRs), and F-1 scores of the median models under different values of e and f, over multiple simulated datasets generated from the homogeneous Poisson process and $\mathbf{\Delta}$ shown in Figure 2(d) of the article (scenario 2). The values of e are chosen corresponding to the expected proportion of the AOI equal to 1%, 2%, 5%, and 10%, respectively. The values of f are chosen the corresponding maximum allowed values $f_{\max}$, as shown in Table S1.

Values of (e, f)	$(-4.60, 2.30)$	$(-3.89, 1.90)$	$(-2.94, 1.40)$	$(-2.20, 1.10)$	$(-1.73, 0.80)$
Recall	0.831 (0.134)	0.866 (0.022)	0.883 (0.023)	0.875 (0.020)	0.880 (0.036)
Precision	0.885 (0.133)	0.917 (0.017)	0.915 (0.018)	0.914 (0.014)	0.896 (0.026)
F1-score	0.857 (0.133)	0.890 (0.015)	0.898 (0.015)	0.894 (0.012)	0.887 (0.013)

Table S4: Sensitivity analysis on the choices of hyperparameters e and f of the MRF prior: Means and standard deviations (in parentheses) of the areas under the ROC curves (AUCs) on the posterior probabilities of inclusion on $\boldsymbol{\Delta}$ under different values of e and f, over multiple simulated datasets generated from the homogeneous Poisson process and $\boldsymbol{\Delta}$ shown in Figure 2(d) of the article (scenario 2). The values of e are chosen corresponding to the expected proportions of the AOI equal to 1%, 2%, 5%, and 10%, respectively. The values of f are chosen to one, one half, one quarter, and zero of the corresponding maximum allowed values $f_{\max}$, as shown in Table S1.

Values of (e, f)	$(-4.60, 0.00)$	$(-3.89, 0.00)$	$(-2.94, 0.00)$	$(-2.20, 0.00)$	$(-1.73, 0.00)$
AUC	0.619 (0.020)	0.667 (0.025)	0.718 (0.028)	0.723 (0.036)	0.671 (0.067)
Values of (e, f)	$(-4.60, 0.58)$	$(-3.89, 0.48)$	$(-2.94, 0.35)$	$(-2.20, 0.28)$	$(-1.73, 0.20)$
AUC	0.637 (0.021)	0.688 (0.031)	0.742 (0.037)	0.737 (0.063)	0.693 (0.115)
Values of (e, f)	$(-4.60, 1.15)$	$(-3.89, 0.95)$	$(-2.94, 0.70)$	$(-2.20, 0.55)$	$(-1.73, 0.40)$
AUC	0.709 (0.027)	0.791 (0.041)	0.841 (0.047)	0.838 (0.092)	0.729 (0.163)
Values of (e, f)	$(-4.60, 2.30)$	$(-3.89, 1.90)$	$(-2.94, 1.40)$	$(-2.20, 1.10)$	$(-1.73, 0.80)$
AUC	0.937 (0.011)	0.945 (0.012)	0.950 (0.009)	0.948 (0.008)	0.941 (0.009)