SUPPLEMENT TO "EXTENDING THE LATENT MULTINOMIAL MODEL WITH COMPLEX ERROR PROCESSES AND DYNAMIC MARKOV BASES"

By Simon J Bonner[*], Matthew R Schofield[†], Patrik Noren[‡] and Steven J Price[§]

University of Western Ontario[*], University of Otago[†], North Carolina State University[‡], and University of Kentucky[§]

Supplement A: Proof of Convergence. To prove that chains generated from Algorithm 2 from the manuscript converge to the correct distribution, we need to satisfy three conditions: 1) that Step 1 produces chains which converge to $\pi(\boldsymbol{\phi}, \boldsymbol{p}|\boldsymbol{z})$ for any $\boldsymbol{z}$ such that $\boldsymbol{z} = \boldsymbol{B}\boldsymbol{x}$ for some $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{n}}$, 2) that Step 2 produces chains which converge to $\pi(\boldsymbol{x}|\boldsymbol{n}, \boldsymbol{\phi}, \boldsymbol{p}, \alpha)$ for any $\boldsymbol{\phi}$ and $\boldsymbol{p}$ in the parameter space, and 3) that the joint posterior distribution satisfies the positivity condition of (Robert and Casella, 2010, pg. 345). Sampling from $\pi(\boldsymbol{\phi}, \boldsymbol{p}|\boldsymbol{z})$ is equivalent to sampling from the posterior distribution for a simple CJS model. This is now standard (Link et al., 2002, see e.g.), and so we assume that Condition 1 is satisfied. It is also simple to show that the positivity constraint is satisfied given that the prior distributions for $\boldsymbol{\phi}$ and $\boldsymbol{p}$ are positive over all of $(0,1)^T \times (0,1)^{T-1}$, as assumed in Section 5 of the original manuscript. It remains to show that Condition 2 holds.

We assume here that $\mathcal{F}_{\boldsymbol{n}}$ contains at least two elements. The fibre always contains at least one element with no errors which we denote by $\boldsymbol{x}^{\emptyset}$. The entries of this element are

$$x_{\boldsymbol{\nu}}^{\emptyset} = \begin{cases} n_{\boldsymbol{\nu}} & \text{if } \boldsymbol{\nu} \text{ is observable} \\ 0 & \text{otherwise} \end{cases}.$$

Cases in which $\mathcal{F}_{\boldsymbol{n}} = \{x^{\emptyset}\}$ arise when no errors could have occurred, for example, if no individuals were ever recaptured. These situations are easily identified and there is no need to sample from the joint posterior of both $\boldsymbol{x}$ and $\boldsymbol{\theta}$ in such cases since $\boldsymbol{x} = x^{\emptyset}$ with probability one.

Some useful results that are easy to prove are:

1. that any configuration of the latent error histories within the fibre has positive probability under the conditional posterior for all values of the parameters in the parameter space,

LEMMA 1. *If $x \in \mathcal{F}_n$ then $\pi(x|n, \phi, p, \alpha) > 0$ for all values of ϕ and p in the parameter space.*

2. that the local sets within the dynamic Markov basis are symmetric,

LEMMA 2. *Let $x \in \mathcal{F}_n$. If $b^+ \in \mathcal{M}_1(x)$ then $-b^+ \in \mathcal{M}_2(x+b^+)$ and if $b^- \in \mathcal{M}_2(x)$ then $-b^- \in \mathcal{M}_1(x + b^-)$.*

3. that all proposals remain inside $\mathcal{F}_n$,

LEMMA 3. *Let $x \in \mathcal{F}_n$. If $b \in \mathcal{M}(x) = \mathcal{M}_1(x) \bigcup \mathcal{M}_2(x)$ then $x+b \in F_n$.*

4. that there is a unique element $x^{\emptyset}$ in $\mathcal{F}_n$ with no errors.

LEMMA 4. *Suppose that $x^{\emptyset} \in \mathcal{F}_n$. Then $e_t(x^{\emptyset}) = 0 \; \forall t = 2, \ldots, T$ if and only if*

$$x_{\nu}^{\emptyset} = \begin{cases} n_{\nu} & \text{if } \nu \text{ is observable} \\ 0 & \text{otherwise} \end{cases}.$$

First we establish irreducibility. Proposition 1 implies that there is a path connecting any two elements in the fibre while Proposition 2 implies that each step, and hence the entire path, has positive probability under the transition kernel. Together, these show that that the chains are irreducible.

PROPOSITION 1. *For any distinct $x_1, x_2 \in \mathcal{F}_n$ there exists a sequence of moves $b_1, \ldots, b_L$ such that:*

1. $b_{L'} \in \mathcal{M}\left(x_1 + \sum_{l=1}^{L'-1} b_l\right)$ *for all* $L' = 1, \ldots, L$
2. $x_1 + \sum_{l=1}^{L'} b_l \in \mathcal{F}_n$ *for all* $L' = 1, \ldots, L-1$, *and*
3. $x_2 = x_1 + \sum_{l=1}^{L} b_l,$

where we take $x_1 + \sum_{l=1}^{0} b_l = x_1$.

PROOF. Our proof follows by (reverse) induction on the number of errors. Suppose that $e_t(x_1) > 0$ for some t. Then $\mathcal{X}_{2t}(x_1)$ and $\mathcal{X}_{3t}(x_1)$ are both non-empty and $\exists b_{11}^- \in \mathcal{M}_2(x_1)$. Then $e_t(x_1+b_{11}^-) = e_t(x_1)-1$ and $x_1+b_{11}^- \in \mathcal{F}_n$ by Lemma 3. Repeating this procedure $L_1 = \sum_{t=2}^{T} e_t(x_1)$ times, we find $b_{11}^-, \ldots, b_{1L_1}^-$ such that

1. $b_{1L'}^- \in \mathcal{M}_2(x_1 + \sum_{l=1}^{L'-1} b_{1l}^-)$ for $L' = 1, \ldots, L_1$,
2. $x_1 + \sum_{l=1}^{L'} b_{1l}^- \in \mathcal{F}_n$ for all $L' = 1, \ldots, L_1$, and

3. $e_t(\boldsymbol{x}_1 + \sum_{l=1}^{L_1} \boldsymbol{b}_{1l}^-) = 0$.

It follows from Lemma 4 that $\boldsymbol{x}_1 + \sum_{l=1}^{L_1} \boldsymbol{b}_{1l}^- = \boldsymbol{x}^\emptyset$. By the same argument, $\exists \boldsymbol{b}_{21}^-, \ldots, \boldsymbol{b}_{2L_2}^-$ such that

1. $\boldsymbol{b}_{2l}^- \in \mathcal{M}_2(\boldsymbol{x}_2 + \sum_{l=1}^{L'-1} \boldsymbol{b}_{2l}^-)$ for $L' = 1, \ldots, L_2$,,
2. $\boldsymbol{x}_2 + \sum_{l=1}^{L'} \boldsymbol{b}_{2l}^- \in \mathcal{F}_{\boldsymbol{n}}$ for all $L' = 1, \ldots, L_2$, and
3. $\boldsymbol{x}_2 + \sum_{l=1}^{L_2} \boldsymbol{b}_{2l}^- = \boldsymbol{x}^\emptyset$.

Moreover, $-\boldsymbol{b}_{2,L_2-l+1}^- \in M_1(\boldsymbol{x}^\emptyset + \sum_{l=0}^{L'-1} -\boldsymbol{b}_{2,L_2-l}^-)$ for all $L' = 1, \ldots, L_2$ by Lemma 2. Then the sequence $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$ where $L = L_1 + L_2$, $\boldsymbol{b}_l = \boldsymbol{b}_{1l}^-$ for $l = 1, \ldots, L_1$, and $\boldsymbol{b}_{L_1+l} = -\boldsymbol{b}_{2,L_2-l+1}^-$ for $l = 1, \ldots, L_2$ satisfies the conditions of the proposition. Note that half of this argument suffices if either $\boldsymbol{x}_1 = \boldsymbol{x}^\emptyset$ or $\boldsymbol{x}_2 = \boldsymbol{x}^\emptyset$. $\qquad\square$

PROPOSITION 2. *Let $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{n}}$. If $\boldsymbol{b} \in \mathcal{M}(\boldsymbol{x})$ then $P(\boldsymbol{x}^{(k+1)} = \boldsymbol{x}+\boldsymbol{b}|\boldsymbol{x}^{(k)} = \boldsymbol{x}) > 0$.*

PROOF. Suppose that $\boldsymbol{b} \in \mathcal{M}_1(\boldsymbol{x})$ and let $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{b}$. Then $-\boldsymbol{b} \in \mathcal{M}_2(\boldsymbol{x}')$ by Lemma 2. Direct calculation of equations (5) and (6) shows that both $q(\boldsymbol{x}'|\boldsymbol{x}) > 0$ and $q(\boldsymbol{x}|\boldsymbol{x}') > 0$. Combined with Lemma 1 it follows that $r(\boldsymbol{x}, \boldsymbol{x}'|\boldsymbol{\phi}^{(k)}, \boldsymbol{p}^{(k)}, \alpha)$ (defined in Step 2, Substep iii of Algorithm 2) is positive and hence that $P(\boldsymbol{x}^{(k+1)} = \boldsymbol{x}'|\boldsymbol{x}^{(k)} = \boldsymbol{x}) = q(\boldsymbol{x}'|\boldsymbol{x}) \cdot r(\boldsymbol{x}, \boldsymbol{x}'|\boldsymbol{\phi}^{(k)}, \boldsymbol{p}^{(k)}, \alpha) > 0$. A similar argument shows that $P(\boldsymbol{x}^{(k+1)} = \boldsymbol{x} + \boldsymbol{b}|\boldsymbol{x}^{(k)} = \boldsymbol{x}) > 0$ for all $\boldsymbol{b} \in \mathcal{M}_2(\boldsymbol{x})$. $\qquad\square$

We establish aperiodicity by showing that there is positive probability of holding at $\boldsymbol{x}^\emptyset$.

PROPOSITION 3. *If $\boldsymbol{x}^{(k)} = \boldsymbol{x}^\emptyset$ then $P(\boldsymbol{x}^{(k+1)} = \boldsymbol{x}_0|\boldsymbol{x}^{(k)} = \boldsymbol{x}_0) \geq .5$.*

PROOF. The set $\mathcal{M}_2(\boldsymbol{x}^\emptyset)$ is empty since there are no errors to remove from $\boldsymbol{x}^\emptyset$. However, Algorithm 2 still proposes to draw a move from $\mathcal{M}_2(\boldsymbol{x}^\emptyset)$ with probability .5. When this occurs we set $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$ so that $P(\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^\emptyset|\boldsymbol{x}^{(k)} = \boldsymbol{x}^\emptyset) \geq .5$. $\qquad\square$

This shows that $\boldsymbol{x}^\emptyset$ is an aperiodic state and hence that the entire chain is aperiodic (Cinlar, 1975, pg. 125)

Since the fibre is finite, irreducibility and aperiodicty are sufficient to ensure that the chains have a unique stationary distribution which is also the limiting distribution (see Cinlar, 1975, Corollary 2.11). That this distribution is equal to the target distribution is guaranteed by the detailed balance

condition of the MH algorithm which holds under Proposition 4 (Liu, 2004, pg. 111).

PROPOSITION 4.   *If $q(\boldsymbol{x}'|\boldsymbol{x}) > 0$ then $q(\boldsymbol{x}|\boldsymbol{x}') > 0$ for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{F_n}$.*

PROOF. Suppose that $q(\boldsymbol{x}'|\boldsymbol{x}) > 0$. Then either $\boldsymbol{x}' - \boldsymbol{x} \in \mathcal{M}_1(\boldsymbol{x})$ or $\boldsymbol{x}' - \boldsymbol{x} \in \mathcal{M}_2(\boldsymbol{x})$. If $\boldsymbol{x}' - \boldsymbol{x} \in \mathcal{M}_1(\boldsymbol{x})$ then $\boldsymbol{x} - \boldsymbol{x}' \in \mathcal{M}_2(\boldsymbol{x}')$ by Lemma 2 and $q(\boldsymbol{x}|\boldsymbol{x}') > 0$. Similarly, if $\boldsymbol{x}' - \boldsymbol{x} \in \mathcal{M}_2(\boldsymbol{x})$ then $\boldsymbol{x} - \boldsymbol{x}' \in \mathcal{M}_1(\boldsymbol{x}')$ and $q(\boldsymbol{x}|\boldsymbol{x}') > 0$.   □

This completes our proof that the Markov chains produced by Algorithm 2 have unique limiting distribution $\pi(\boldsymbol{x}, \boldsymbol{\phi}, \boldsymbol{p}|\boldsymbol{n}, \alpha)$ so that realisations from the tail of a converged chain can be used to approximate properties of the joint posterior distribution of $\boldsymbol{x}$, $\boldsymbol{\phi}$, and $\boldsymbol{p}$.

Supplement B: Model $M_{t\alpha}$.   Here we show how the model described by Link et al. (2010) can be fit into the extended model using a dynamic Markov basis to sample from the posterior distribution. Model $M_{t\alpha}$ extends the standard closed population model with time dependent capture probabilities by allowing for individuals to be misidentified. Specifically, $M_{t\alpha}$ assumes that individuals are identified correctly with probability α on each capture and that errors do not duplicate observed marks in that one marked individual cannot be identified as another marked individual and that the same error can never occur twice. This means that each error creates a new recorded history with a single non-zero entry. For example, suppose that $T = 3$ and that individual i is captured on the first occasion, recaptured and misidentified on the second occasion, and not captured on the third occasion. The true capture history for this individual, including errors, is 120, where the event 2 denotes that the individual was captured and misidentified. The individual would then contribute two recorded histories, 100 and 010, to the observed data set.

Extended Formulation.   As with the CJS/BRE model, we cast model $M_{t\alpha}$ into the new framework by 1) identifying the sets of observable capture histories, latent error histories, and latent capture histories, 2) constructing the linear constraints matrices for the corresponding count vectors, and 3) identifying the components of the likelihood function. For an experiment with T occasions, the set of observable capture histories contains all 2^{T-1} histories in $\{0, 1\}^T$ excluding the unobservable zero history, the set of latent error histories includes all 3^T histories in $\{0, 1, 2\}^T$, and the set of latent capture histories includes all 2^T histories in $\{0, 1\}^T$. The matrix $\boldsymbol{A}$ is defined

exactly as in Link et al. (2010): $A_{ij} = 1$ if the ith observable history is observed from the jth latent error history and $A_{ij} = 0$ otherwise. Similarly, $B_{kj} = 1$ if the kth latent capture history has the same pattern of captures as the jth latent error history. Mathematically, let $\boldsymbol{\omega}_i$, $\boldsymbol{\nu}_j$, and $\boldsymbol{\xi}_k$ represent the ith, j^{th}, and k^{th} observable history, latent error history, and latent capture history for some implicit orderings. Then

$$
A_{ij} = \begin{cases} 1 & \text{if } \omega_{it} = I(\nu_{jt} = 1) \text{ for all } t = 1, \ldots, T \\ & \text{or if } \boldsymbol{\omega}_i = \boldsymbol{\delta}_t \text{ and } \nu_{jt} = 2 \text{ for some } t \in \{1, \ldots, T\} \\ \\ 0 & \text{otherwise} \end{cases}
$$

and

$$
B_{kj} = \begin{cases} 1 & \xi_{kt} = I(\nu_{jt} = 1) + I(\nu_{jt} = 2) \text{ for all } t = 1, \ldots, T \\ \\ 0 & \text{otherwise} \end{cases} .
$$

Here $\boldsymbol{\delta}_t$ represents t^{th} column of the $T \times T$ identity matrix (i.e., the vector with a single 1 in the tth entry). Finally, the two components of the likelihood function are

$$
\pi(\boldsymbol{z}|\boldsymbol{p}) = \frac{N!}{\prod_{\boldsymbol{\xi} \in \mathcal{Z}} z_{\boldsymbol{\xi}}!} \prod_{\boldsymbol{\xi} \in \mathcal{Z}} \left[\prod_{k=1}^{K} p_k^{I(\xi_k=1)} (1 - p_k)^{I(\xi_k=0)} \right]^{z_{\boldsymbol{\xi}}}
$$

and

$$
\pi(\boldsymbol{x}|\boldsymbol{z}, \alpha) = \frac{\prod_{\boldsymbol{\xi} \in \mathcal{Z}} z_{\boldsymbol{\xi}}!}{\prod_{\boldsymbol{\nu} \in \mathcal{X}} x_{\boldsymbol{\nu}}!} \prod_{\boldsymbol{\nu} \in \mathcal{X}} \left[\prod_{k=1}^{K} \alpha^{I(\nu_k=1)} (1 - \alpha)^{I(\nu_k=2)} \right]^{x_{\boldsymbol{\nu}}}
$$

Here $N = \sum_{\boldsymbol{\xi} \in \mathcal{Z}} z_{\boldsymbol{\xi}}^T$ represents the true population size. Note that the product of these two contributions exactly reconstructs the single likelihood function for $M_{t\alpha}$ defined by Link et al. (2010, eqns 6 and 7).

Dynamic Moves. We can again generate a dynamic Markov basis by selecting from a set of moves which add or remove errors from the current configuration. An extra step is also needed on each iteration of the algorithm to update the count of the unobserved individuals, $\boldsymbol{\nu_0}$.

Let $\chi_{\nu t}(\boldsymbol{x}) = \{\boldsymbol{\nu} : \nu_t = \nu \text{ and } x_{\boldsymbol{\nu}} > 0\}$ be the set of latent error histories with event ν on occasion t and positive counts in $\boldsymbol{x}$. As with the CJS/BRE model, we define $\mathcal{M}(\boldsymbol{x})$ as the union of two sets: $\mathcal{M}_1(\boldsymbol{x})$ containing moves

which add errors and $\mathcal{M}_2(\boldsymbol{x})$ containing moves which remove errors. Each move in the dynamic Markov basis modifies the counts of three latent error histories. Moves in $\mathcal{M}_1(\boldsymbol{x})$ are defined by sampling one history $\boldsymbol{\nu}_0 \in \chi_{0t}(\boldsymbol{x})$ for some t and are denoted by $\boldsymbol{b}(\boldsymbol{\nu}_0)$. The other two latent error histories, $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$, are defined by setting $\boldsymbol{\nu}_1 = \delta_t$ and $\boldsymbol{\nu}_2 = \boldsymbol{\nu}_0 + 2\delta_t$. The move $\boldsymbol{b}(\boldsymbol{\nu}_0)$ is then defined by setting

$$b_{\boldsymbol{\nu}}(\boldsymbol{\nu}_0) = \begin{cases} -1 & \text{if } \boldsymbol{\nu} = \boldsymbol{\nu}_0 \text{ or } \boldsymbol{\nu} = \boldsymbol{\nu}_1 \\ 1 & \text{if } \boldsymbol{\nu} = \boldsymbol{\nu}_2 \end{cases}.$$

Similarly, moves in $\mathcal{M}_2(\boldsymbol{x})$, denoted by $\boldsymbol{b}(\boldsymbol{\nu}_2)$, are defined by sampling a history $\boldsymbol{\nu}_2 \in \chi_{2t}(\boldsymbol{x})$ for some t, setting

$$\boldsymbol{\nu}_1 = \delta_t$$

and

$$\nu_{0s} = \begin{cases} 0 & s = t \\ \nu_{2s} & \text{otherwise} \end{cases} \quad t = 1, \ldots, T.$$

The move $\boldsymbol{b}(\boldsymbol{\nu}_2)$ is then defined by setting

$$b_{\boldsymbol{\nu}}(\boldsymbol{\nu}_2) = \begin{cases} -1 & \text{if } \boldsymbol{\nu} = \boldsymbol{\nu}_2 \\ 1 & \text{if } \boldsymbol{\nu} = \boldsymbol{\nu}_0 \text{ or } \boldsymbol{\nu} = \boldsymbol{\nu}_1 \end{cases}.$$

If we assume that the decision to add or remove an error are each chosen with probability .5 and that histories are sampled uniformly from $\chi_{0.} = \bigcup_{t=1}^{T} \chi_{0t}$ and $\chi_{2.} = \bigcup_{t=1}^{T} \chi_{2t}$, when adding or removing an error respectively, then the proposal densities for the moves $\boldsymbol{x}' = \boldsymbol{x}^{(k-1)} + \boldsymbol{b}(\boldsymbol{\nu}_0)$ and $\boldsymbol{x}' = \boldsymbol{x}^{(k-1)} + \boldsymbol{b}(\boldsymbol{\nu}_2)$ are given by

$$q(\boldsymbol{x}'|\boldsymbol{x}^{(k-1)}) = \frac{.5}{\#\chi_{0.}(\boldsymbol{x}^{(k-1)}) \cdot \#\{t : \nu_{0t} = 1\}}$$

and

$$q(\boldsymbol{x}'|\boldsymbol{x}^{(k-1)}) = \frac{.5}{\#\chi_{2.}(\boldsymbol{x}^{(k-1)}) \cdot \#\{t : \nu_{2t} = 2\}}.$$

As in the algorithm for the CJS/BRE, we retain $x^{(k-1)}$ with probability 1 if an empty set is encountered in one of these processes or if the selected move reduces any of the counts in $\boldsymbol{x}$ below zero. Details of the full algorithm for sampling from the posterior distribution of model $M_{t\alpha}$ are given in Algorithm 3.

Initialise $\boldsymbol{p}^{(0)}$ and α.
Initialise $\boldsymbol{x}^{(0)}$ so that $\boldsymbol{n} = \boldsymbol{A}\boldsymbol{x}^{(0)}$ and set $\boldsymbol{z}^{(0)} = \boldsymbol{B}\boldsymbol{x}^{(0)}$.
Set $k = 1$.

1. Update $\boldsymbol{p}$ and α conditional on $\boldsymbol{x}^{(k-1)}$ and $\boldsymbol{z}^{(k-1)}$. Call the results $\boldsymbol{p}^{(k)}$ and $\alpha^{(k)}$.

2. Update $\boldsymbol{x}$ and $\boldsymbol{z}$ conditional on and $\boldsymbol{p}^{(k)}$ and $\alpha^{(k)}$ as follows.

   i) With probability .5 sample $\boldsymbol{b}$ from $\mathcal{M}_1(\boldsymbol{x}^{(k-1)})$. If $\mathcal{M}_1(\boldsymbol{x}^{(k-1)}) = \emptyset$ then set $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(k-1)}$ and continue to step v).
   Otherwise sample $\boldsymbol{b}$ from $\mathcal{M}_2(\boldsymbol{x}^{(k-1)})$. If $\mathcal{M}_2(\boldsymbol{x}^{(k-1)}) = \emptyset$ then set $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(k-1)}$ and continue to step v).

   ii) Set $\boldsymbol{x}^{\mathrm{prop}} = \boldsymbol{x}^{(k-1)} + \boldsymbol{b}$. If $x_j' < 0$ for any $j = 1, \ldots, J$ set $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(k-1)}$ and continue to step v).

   iii) Calculate the Metropolis acceptance probability:

   $$r(\boldsymbol{x}, \boldsymbol{x}^{\mathrm{prop}}|\boldsymbol{p}^{(k)}, \alpha) = \min\left\{1, \frac{\pi(\boldsymbol{x}^{\mathrm{prop}}|\boldsymbol{n}, \boldsymbol{p}^{(k)}, \alpha)}{\pi(\boldsymbol{x}^{(k-1)}|\boldsymbol{n}, \boldsymbol{p}^{(k)}, \alpha)} \cdot \frac{q(\boldsymbol{x}^{(k-1)}|\boldsymbol{x}^{\mathrm{prop}})}{q(\boldsymbol{x}^{\mathrm{prop}}|\boldsymbol{x}^{(k-1)})}\right\}.$$

   iv) Set $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{\mathrm{prop}}$ with probability $r(\boldsymbol{x}, \boldsymbol{x}^{\mathrm{prop}}|\boldsymbol{\phi}^{(k)}, \boldsymbol{p}^{(k)}, \alpha)$ and $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(k-1)}$ otherwise.

   v) Set $\boldsymbol{z}^{(k)} = \boldsymbol{B}\boldsymbol{x}^{(k)}$

3. Increment k.

Algorithm 3: Proposed algorithm for sampling from the posterior distribution of Model $M_{t\alpha}$ using the dynamic Markov basis.

References.

CINLAR, E. (1975). *Introduction to Stochastic Processes.* Prentice-Hall, Englewood Cliffs, NJ.

LINK, W. A., CAM, E., NICHOLS, J. D. and COOCH, E. G. (2002). Of BUGS and Birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. *Journal of Wildlife Management* **66** 277–291.

LINK, W. A., YOSHIZAKI, J., BAILEY, L. L. and POLLOCK, K. H. (2010). Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. *Biometrics* **66** 178–185.

LIU, J. S. (2004). *Monte Carlo Strategies in Scientific Computing.* Springer, New York, NY.

ROBERT, C. P. and CASELLA, G. (2010). *Monte Carlo Statistical Methods*, 2nd ed. Springer-Verlag, New York.

DEPARTMENT OF STATISTICAL AND ACTUARIAL SCIENCES
UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO N6A 5B7
E-MAIL: sbonner6@uwo.ca

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF OTAGO
DUNEDIN 90504, NEW ZEALAND
E-MAIL: mschofield@maths.otago.ac.nz

DEPARTMENT OF MATHEMATICS
NORTH CAROLINA STATE UNIVERSITY
BOX 8205
RALEIGH, NC 27695
E-MAIL: pgnoren2@ncsu.edu

DEPARTMENT OF FORESTRY
UNIVERSITY OF KENTUCKY
LEXINGTON, KENTUCKY 40546-0073 USA
E-MAIL: steven.price@uky.edu