Appendix

A common sampling technique often used in this setting is the Metropolis-Hastings algorithm, which is a Markov chain Monte Carlo (MCMC) method. The M-H acceptance probability for moving from state x to state x' is shown below, where each *state* is a DBN.

$$\alpha(x, x') = \min\left\{1, \frac{p(D|x')}{p(D|x)} \times \frac{p(x' \to x)}{p(x \to x')}\right\} = \min\left\{1, \quad \underbrace{\frac{p(D|x')}{p(D|x)}}_{likelihood\ ratio} \quad \times \quad \underbrace{\frac{p(m')p(x|x', m')}{p(m)p(x'|x, m)}}_{proposal\ ratio}\right\}$$

where m is the move type that allows for a transition from state x to x' and m' is the reverse move type for a transition from state x' back to state x. The proposal ratio can be split into two terms: one is the ratio of the proposal probabilities for move types and the other is the ratio of selecting a particular state given the current state and the move type. The choice of scoring metric determines the likelihoods, and often $p(m')$ and $p(m)$ are chosen *a priori* to be simple to calculate or to actually cancel out.

Move type m	Proposal probability	$\frac{p(m')}{p(m)}$	$\frac{p(x\|x',m')}{p(x'\|x,m)}$	
(M_1) add edge to G_1	P_a	$\frac{P_d}{P_a}$	$\frac{(E_1+1)^{-1}}{(P_{max}n-E_1)^{-1}} = \frac{P_{max}n-E_1}{E_1+1}$	KNKT, KNUT, UNUT
(M_2) delete edge from G_1	P_d	$\frac{P_a}{P_d}$	$\frac{(P_{max}n-E_1+1)^{-1}}{E_1^{-1}} = \frac{E_1}{P_{max}n-E_1+1}$	KNKT, KNUT, UNUT
(M_3) add edge to Δg_i	P_{ae}	$\frac{P_{de}}{P_{ae}}$	$\frac{m^{-1}(S_i+1)^{-1}}{m^{-1}(S_{max}-S_i)^{-1}} = \frac{S_{max}-S_i}{S_i+1}$	KNKT, KNUT, UNUT
(M_4) delete edge from Δg_i	P_{de}	$\frac{P_{ae}}{P_{de}}$	$\frac{m^{-1}(S_{max}-S_i+1)^{-1}}{m^{-1}S_i^{-1}} = \frac{S_i}{S_{max}-S_i+1}$	KNKT, KNUT, UNUT
(M_5) move edge from Δg_i to Δg_j	P_{me}	1	$\frac{(m(m-1))^{-1}(S_j+1)^{-1}}{(m(m-1))^{-1}S_i^{-1}} = \frac{S_i}{S_j+1}$	KNKT, KNUT, UNUT
(M_6) locally shift t_i	P_{st}	1	$\frac{(2d+1)^{-1}}{(2d+1)^{-1}} = 1$	KNUT, UNUT
(M_7) merge Δg_i and Δg_{i+1}	P_m	$\frac{P_s}{P_m}$	$\frac{(m-1)^{-1}2(S_i+S_{i+1})^{-1}\binom{S_i+S_{i+1}}{S_i}^{-1}}{(m-1)^{-1}} = \frac{2}{(S_i+S_{i+1})\binom{S_i+S_{i+1}}{S_i}}$	UNUT
(M_8) split Δg_i	P_s	$\frac{P_m}{P_s}$	$\frac{(m-1)^{-1}}{(m-1)^{-1}(S_i/2)^{-1}\binom{S_i}{x}^{-1}} = (S_i/2)\binom{S_i}{x}$	UNUT
(M_9) create new Δg_i	P_{ag}	$\frac{P_{dg}}{P_{ag}}$	$\frac{(m+1)^{-1}}{(N-m)^{-1}n^{-2}} = \frac{(N-m)n^2}{m+1}$	UNUT
(M_{10}) delete Δg_i	P_{dg}	$\frac{P_{ag}}{P_{dg}}$	$\frac{(N-m-1)^{-1}n^{-2}}{m^{-1}} = \frac{m}{(N-m-1)n^2}$	UNUT

Table 1: E_1 is the total number of edges in G_1, S_{max} is the maximum number of transitions allowed in a single transition time, P_{max} is the maximum parent set size, and S_i is the number of edge changes in the set Δg_i. The proposal ratio is the product of the last two columns. The KNKT setting uses moves $(M_1)-(M_5)$, KNUT uses moves $(M_1)-(M_6)$, and UNUT uses moves $(M_1)-(M_{10})$, in each case with the proposal probabilities appropriately normalized to add to 1.

All of the F1-measures for the nsDBNs learned under the UNUT setting are shown in Table 2 below.

A

λ_m \ λ_e	1.0	2.0	3.0	4.0	5.0
1.0	0.434	0.942	0.947	0.974	**0.992**
2.0	0.676	0.956	0.955	0.991	0.991
5.0	0.921	0.955	0.973	0.973	0.991
10.0	0.926	0.955	0.966	0.983	0.979
50.0	0.880	0.881	0.904	0.892	0.881

B

λ_m \ λ_e	2.0	3.0	5.0
1.0	0.949	0.951	0.947
2.0	0.938	0.952	0.936
5.0	**0.953**	0.946	0.940

Table 2: F1-measure for different values of λ_e and λ_m under the UNUT setting. F1-measures over 0.9 are shaded in light gray and the best score is shown in bold. **A:** F1-measures for the nine variable dataset defined by the nsDBN in Figure 1A. **B:** F1-measures for the large 100 variable dataset.