not exist due either to nonpositive definiteness or to singularity ($r \leq k$). Suppose that one can actually set some reasonable bounds on the posterior distribution of each of the k coefficient estimates in $\hat{\boldsymbol{\theta}}$. These bounds may be set according to empirical observation with similar models, as a Bayes-like prior assertion (Hathaway 1985; O'Leary and Rust 1986; McCullagh and Nelder 1989; Geyer 1991; Wolak 1991; Geyer and Thompson 1992; Dhrymes 1994, Sec. 5.11). Thus, we assume that $\boldsymbol{\theta} \in [\mathbf{g}, \mathbf{h}]$, where $\mathbf{g}$ is a $k \times 1$ vector of lower bounds and $\mathbf{h}$ is a $k \times 1$ vector of upper bounds.

The goal now is to draw samples from the distribution of $\hat{\boldsymbol{\theta}} : \hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, (-H)^{-1}) \propto e^{-T/2}$, truncated to be within $[\mathbf{g}, \mathbf{h}]$, and where $T = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\mathbf{H}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$. Note that the normal density does not include an expression for the variance–covariance matrix—only the inverse (i.e., the negative of the Hessian), which exists here. We thus decompose T as follows:

$$\begin{aligned} T &= (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\mathbf{H}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ &= (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\mathbf{U}'\mathbf{L}\mathbf{U}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}), \end{aligned} \tag{6.15}$$

where $\mathbf{U}'\mathbf{L}\mathbf{U}$ is the *spectral decomposition* of $\mathbf{H}$; $\text{rank}(\mathbf{H}) = r \leq k$; $\mathbf{H}$ has r non-zero eigenvalues, denoted $d_1, \ldots, d_r$; $\mathbf{U}$ is $k \times k$ and orthogonal and hence $(\mathbf{U})^{-1} = \mathbf{U}'$; and $\mathbf{L} = \text{diag}(\mathbf{L}_1, 0)$, where $\mathbf{L}_1 = \text{diag}(d_1, \ldots, d_r)$. Thus, the $\mathbf{L}$ matrix is a diagonal matrix with r leading values of eigenvalues and $n - r$ trailing zero values.

Now make the transformation $\mathbf{A} = \mathbf{U}(\hat{\boldsymbol{\theta}} - [\mathbf{h} + \mathbf{g}]/2)$, the density for which would normally be $\mathbf{A} \sim N(\mathbf{U}(\boldsymbol{\theta} - [\mathbf{h} + \mathbf{g}]/2), (-\mathbf{L})^{-1})$. This transformation centers the distribution of $\mathbf{A}$ at the middle of the bounds, and since $\mathbf{L}$ is diagonal, it factors into the product of independent densities. But this expression has two problems:

- $(-\mathbf{L})^{-1}$ does not always exist.
- $\mathbf{A}$ has complicated multivariate support (a hypercube not necessarily parallel with the axes of the elements of $\mathbf{A}$), which is difficult to draw random numbers from.

We now address these two problems. First, in place of $\mathbf{L}$, we use $\mathbf{L}^*$ defined such that $L_i^* = L_i$ if $L_i > 0$ and L_i^* is equal to some small positive value otherwise (where the subscript refers to the row and column of the diagonal element). Except for the specification of the support of A that we consider next, this transforms the density into

$$\begin{aligned} \mathbf{A} &\sim N(\mathbf{U}'\boldsymbol{\theta}, (-\mathbf{L}^*)^{-1}) \\ &= \prod_i N(U_i(\theta_i - [h_i + g_i]/2, -1/L_i^*). \end{aligned} \tag{6.16}$$

Typeset in TeX/LaTeX | Inline Equation

THEOREM 4 *Suppose* (1) *is disconjugate on the interval* $I \subseteq \mathbb{T}$ *and* (3) *holds. Let* $a \in I$*. Then there are solutions* $x_1, x_2, \dots, x_n$ *of* (1) *so that the Wronskians* $w_i := w(x_1, x_2, \dots, x_i)$ *satisfy*

$$w_i(t) > 0 \quad \textit{for } t \in (\sigma^{n-i-1}(a), \infty) \cap \mathbb{T}^{\kappa^i} \cap I, \ 1 \le i \le n-1. \tag{6}$$

Proof Let $x_1, x_2, \dots, x_n$ be the collection of solutions of (1) defined in (5). First observe that $w_1 = x_1$ has a zero of multiplicity $n-1$ at a, so that x_1 can neither equal zero nor change sign for $t > \sigma^{n-1}(a)$ since otherwise x_1 would have n generalized zeros. But $x_1^{\Delta^{n-1}}(a) = 1$, so $x(t) > 0$ for $t > \sigma(a)$. Next consider w_k for $1 < k < n$. If $w_k(t_0) = 0$ for some $t_0 > \sigma^{n-k-1}(a)$, then there are constant $c_1, c_2, \dots, c_k$, not all zero, so that $y := c_1x_1 + c_2x_2 + \cdots + c_kx_k$ has a zero of multiplicity k at t_0. But y has a zero of multiplicity $n-k$ at a giving y a total of at least n zeros. Since y is nontrivial, that contradicts the disconjugacy of L_n. Hence $w_k(t) \neq 0$ for $t > \sigma^{n-k}(a)$.

Next suppose there are points $t_1 < t_2$ in $I \cap (\sigma^{n-k-1}(a), \infty) \cap \mathbb{T}^{\kappa^k}$ so that $w_k(t_1)w_k(t_2) < 0$. Let $t_0 = \sup\{t \in I\colon\ w_k(t_1)w_k(s) > 0 \text{ for } t_1 \le s \le t\}$. Then $t_0 \le t_2$. Since $w_k(t_0) \neq 0$, $w_k(t_1)w_k(t_0) > 0$. If t_0 were right-dense, then we could find a sequence $s_m \to t_0^+$ so that $w_k(t_1)w_k(t_m) < 0$. But $w_k(t_1)w_k(s_m) \to w_k(t_1)w_k(t_0) > 0$, and that is impossible. Hence t_0 must be right-scattered. Next let x be the solution to the BVP $L_n[x] = 0$, $x^{\Delta^i}(a) = 0$ for $0 \le i \le n-k-1$, $x(t_0) = 1$, and $x^{\Delta^i}(\sigma(t_0)) = 0$ for $0 \le i \le k-1$. Then $x = c_1x_1 + c_2x_2 + \cdots + c_kx_k$ since x has a zero of multiplicity $n-k$ at a. Let W be the $(k+1) \times (k+1)$ matrix with first column $[x(t_0), x(\sigma(t_0)), \dots, x^{\Delta^{k-1}}(\sigma(t_0))]^T = [0, 0, \dots, 0, x^{\Delta^{k-1}}(\sigma(t_0))]$ and column $j+1$ equal to $[x_j(t_0), x_j(\sigma(t_0)), x_j^{\Delta}(\sigma(t_0)), \dots, x_j^{\Delta^{k-1}}(\sigma(t_0))]$ for $1 \le j \le k$. Then $0 = \det W = w_k(\sigma(t_0)) + (-1)^k x^{\Delta^{k-1}}(\sigma(t_0)) \det W_\sigma$ where W_σ is the $k \times k$ matrix with jth column $[x_j(t_0), x_j(\sigma(t_0)), x_j^{\Delta}(\sigma(t_0)), \dots, x_j^{\Delta^{k-2}}(\sigma(t_0))]^T$ for $1 \le j \le k$. For function y on $\mathbb{T}$ which is Δ-differentiable at t we have $y^\sigma(t) = y(t) + y^\Delta(t)\mu(t)$. Using this identity on the columns of W_σ we get that the jth column of W_σ has the form $[x_j(t_0), x_j(t_0) + x_j^{\Delta}(t_0)\mu(t_0), \dots, x_j^{\Delta^{k-2}} + x_j^{k-1}(t_0)\mu(t_0)]^T$ for $1 \le j \le k$. But $\mu(t_0) > 0$, so elementary row operations applied to W_σ give $\det W_\sigma = (\mu(t_0))^{k-1} w_k(t_0)$. Then $0 = \det W = w(\sigma(t_0)) + (-1)^k x^{\Delta^{k-1}}(\sigma(t_0)) \times (\mu(t_0))^{k-1} 2_\kappa(t_0)$. Now $w_k(t_0)w_k(\sigma(t_0)) < 0$, so $(-1)^k x^{\Delta^{k-1}}(\sigma(t_0))x(t_0) > 0$ since $x(t_0) = 1$. Then (4) holds with $\tau = \sigma(t_0)$, and x has a generalized zero for static traffic. Then we will consider static traffic only, with the additional restriction that all lightpaths have at most two hops.

3.1. DYNAMIC TRAFFIC

In this subsection, we consider dynamic traffic. The following is a simple result that comes from [Aggarwal et al., 1994] [2].

Theorem 16 *[Aggarwal et al., 1994] A network with no conversion has wavelength requirements of*

$$W \leq \min\{(L-1)H+1, (2L-1)\sqrt{M}-L+2\}$$

for dynamic traffic, where M is the number of links in the network and H is an upper bound on the length of any lightpath.

Proof. Suppose $(L-1)H+1 \leq (2L-1)\sqrt{M}-L+2$. Then $W = (L-1)H+1$ is sufficient because an arriving lightpath will share links with at most $(L-1)H$ other lightpaths. Therefore, we can find and use a wavelength unused by these other lightpaths.

Now suppose $(L-1)H+1 > (2L-1)\sqrt{M}-L+2$. Then $W = (2L-1)\sqrt{M}-L+2$ is also sufficient. In this case, there are $L\sqrt{M}$ wavelengths dedicated to *long* lightpaths that have at least $\sqrt{M}$ hops. The rest of the $(L-1)(\sqrt{M}-1)+1$ wavelengths are dedicated to *short* lightpaths that have at most $\sqrt{M}-1$ hops.

We can always find a wavelength for an arriving short lightpath because it has at most $\sqrt{M}-1$ hops and so will intersect with at most $(L-1)(\sqrt{M}-1)$ other lightpaths. We can always find a wavelength for an arriving long lightpath for the following reasons. Let K denote the number of long lightpaths including the arriving lightpath. Note that the average number of long lightpaths crossing any link is at least $\frac{K\sqrt{M}}{M}$ because each of the K lightpaths has at least $\sqrt{M}$ hops. Since one or more links have at least $\frac{K\sqrt{M}}{M}$ lightpaths, $\frac{K\sqrt{M}}{M} \leq L$. Thus, $K \leq L\sqrt{M}$, and we can give each long lightpath, and in particular the arriving one, its own wavelength.

The bound can be loose. For ring networks, $M = N$ and, assuming lightpaths take shortest-hop routes, $H \leq N/2$. The wavelength requirements then become $\min\{(L-1)N/2, (2L-1)\sqrt{N}-L+2\}$ which is $O(L\sqrt{N})$. This is larger than the $O(L \log N)$ wavelengths requirements in Theorem 10.

The next result shows that with "limited wavelength conversion" the number of wavelengths required for dynamic traffic is only dependent on L and not on H, N, or number of links. What we mean by limited conversion is that the wavelength degree is a constant d.