Properties of the Maximum Likelihood Estimator in Spatial Autoregressive Models

Supplementary Material

Grant Hillier[a] and Federico Martellosio[b,*]

[a]CeMMAP and Department of Economics, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[b]School of Economics, University of Surrey, Guildford, Surrey, GU2 7XH, UK

12 September 2013

Abstract

Section 8 presents proofs of two results asserted in the main text, and two additional lemmas used in the paper. Section 9 gives further details for the applications discussed in the main paper.

8 Additional Proofs and Results

8.1 Proof of Lemma 3.4

Let ω_t, $t = 1, ..., T$, denote the distinct (possibly complex) eigenvalues of W, ordered arbitrarily, let $e_t = e_t(W)$ denote the t-th elementary symmetric function in the T distinct eigenvalues of W, and let $e_{t,j}$ be that with the j-th eigenvalue omitted. The polynomial

$$\prod_{t=1}^{T}(1-\lambda\omega_t) = \sum_{t=0}^{T}(-\lambda)^t e_t$$

is a generating function for the e_t, and we have accordingly $e_0 = 1$, and $e_r = 0$ for $r > T$. Correspondingly, the polynomial

$$\prod_{\substack{t=1 \\ t\neq j}}^{T}(1-\lambda\omega_t) = \sum_{t=0}^{T-1}(-\lambda)^t e_{t,j}$$

*Corresponding author. Tel: +44 (0) 1483 683473
E-mail addresses: ghh@soton.ac.uk (G. Hillier), f.martellosio@surrey.ac.uk (F. Martellosio)

is a generating function for the $e_{t,j}$, and it can easily be checked (by equating coefficients of suitable powers of λ) that

$$\omega_j e_{t-1,j} = e_t - e_{t,j}, \tag{8.1}$$

for $t = 1, ..., T-1$, and

$$\omega_j e_{T-1,j} = e_T. \tag{8.2}$$

We can therefore write the first-order condition (see equation (3.6) as

$$n\left(b - a\lambda\right)\sum_{t=0}^{T}(-\lambda)^t e_t - \left(a\lambda^2 - 2b\lambda + c\right)\sum_{j=1}^{T}\left(n_j\omega_j \sum_{\substack{t=0\\ t\neq j}}^{T-1}(-\lambda)^t e_{t,j}\right) = 0, \tag{8.3}$$

where $a := y'W'M_XWy$, $b := y'W'M_Xy$, and $c := y'M_Xy$. We now show that the polynomial equation (8.3) has degree T. Using (8.2) and $\sum_{j=1}^{T} n_j = n$, the coefficient of λ^{T+1} is

$$na(-1)^{T+1}e_T + (-1)^T a \sum_{j=1}^{T} n_j\omega_j e_{T-1,j} = 0.$$

On the other hand, the coefficient of λ^T is

$$a(-1)^T\left(ne_{T-1} - \sum_{j=1}^{T} n_j\omega_j e_{T-2,j}\right) + nb(-1)^{T-1}e_T,$$

which, on using (8.1), reduces to

$$a(-1)^T\left(\sum_{j=1}^{T} n_j e_{T-1,j}\right) + nb(-1)^{T-1}e_T.$$

This will a.s. not vanish: the term e_T can vanish if one eigenvalue is zero, but at least one term in the sum in the first term will not vanish, since only one eigenvalue can be zero.

8.2 Proof of Proposition 3.8

For simplicity, assume that all densities exist. We need to show that the distribution of the maximal invariant $v = y(y'y)^{-1/2} \in \mathcal{S}^{n-1}$ is invariant under scale-mixtures of the distribution of y. Let $f(y)$ denote the density of $y \in \mathbb{R}^n$, and let $q := (y'y)^{1/2} >$

0. We may transform $y \to (q, v)$, setting $y = qv$. The volume element (Lebesgue measure) (dy) on $\mathbb{R}^n$ decomposes as

$$(dy) = q^{n-1} dq (v' dv)$$

where $(v'dv)$ denotes (unnormalized) invariant measure on $\mathcal{S}^{n-1}$ (see Muirhead (1982), Theorem 2.1.14 for a more general version of this result). The measure on $\mathcal{S}^{n-1}$ induced by the density $f(y)$ for y is therefore defined, for any subset $\mathcal{A}$ of $\mathcal{S}^{n-1}$, by

$$\Pr(v \in \mathcal{A}) = \int_{\mathcal{A}} \left\{ \int_{q>0} q^{n-1} f(qv) dq \right\} (v'dv).$$

Now let κ be a random scalar independent of y with density $p(\kappa)$ on $\mathbb{R}^+$. The density of $y^* := \kappa y$ is then given by the mixture

$$g(y^*) := \int_{\kappa>0} \kappa^{-n} f(y^*/\kappa) p(\kappa) d\kappa$$

The measure induced by $g(\cdot)$ for $v(y^*) = v(y)$ is therefore

$$\begin{aligned}
\int_{q>0} q^{n-1} g(qv) dq &= \int_{q>0} \int_{\kappa>0} q^{n-1} \kappa^{-n} f(qv/\kappa) p(\kappa) d\kappa dq \\
&= \int_{q>0} q^{n-1} f(qv) dq
\end{aligned}$$

on transforming to $(q/\kappa, \kappa)$ and integrating out κ. That is, for any (proper) density $p(\cdot)$, $g(\cdot)$ induces the same measure on $\mathcal{S}^{n-1}$ as does $f(\cdot)$, as claimed.

8.3 The case $T = 3$

In several of the examples we discuss in the text W has just three distinct eigenvalues, so $T = 3$ (namely, the pure (unbalanced) Group Interaction model with $u = 2$, and the Group Interaction model with $u = 2$ and constant mean). In that case Theorem 3 implies that the cdf of $\hat{\lambda}_{\mathrm{ML}}$ can, under conditions given in the text, be represented in the form

$$\Pr(\hat{\lambda}_{\mathrm{ML}} \leq z) = \Pr(d_{11} q_{n_1} + d_{22} q_{n_2} + d_{33} q_{n_3} \leq 0), \tag{8.4}$$

where q_{n_i}, $i = 1, 2, 3$ are independent $\chi^2_{n_i}$ random variables. Proposition 4.1 implies that $d_{11} < 0$ and $d_{33} >$ for all $z \in \Lambda$, while d_{22} changes sign at a point $z = z_2$ in Λ, with $d_{22} > 0$ for $z < z_2$, and conversely. For $z > z_2$, therefore, for fixed values of (q_{n_1}, q_{n_2}), we may write the conditional cdf as

$$\Pr(\hat{\lambda}_{\mathrm{ML}} \leq z | q_{n_1}, q_{n_2}) = \Pr(q_{n_3} \leq \psi_1 q_{n_1} + \psi_2 q_{n_2}) = \mathcal{G}_{n_3}(\psi_1 q_{n_1} + \psi_2 q_{n_2}), \tag{8.5}$$

with coefficient functions

$$\psi_t = \psi_t(z;\lambda) = -\frac{d_{tt}}{d_{33}}, t = 1, 2, \tag{8.6}$$

that are both positive for $z_2 < z < 1$. Likewise, for $z < z_2$ we may write

$$\Pr(\hat{\lambda}_{\text{ML}} \le z | q_{n_2}, q_{n_3}) = 1 - \Pr(q_{n_1} \le \tilde{\psi}_2 q_{n_2} + \tilde{\psi}_3 q_{n_3}) = 1 - \mathcal{G}_{n_1}(\tilde{\psi}_2 q_{n_2} + \tilde{\psi}_3 q_{n_3}). \tag{8.7}$$

with

$$\tilde{\psi}_t = \psi_t(z;\lambda) = -\frac{d_{tt}}{d_{11}}, t = 2, 3, \tag{8.8}$$

again both positive.

The corresponding conditional densities are, for $z > z_2$,

$$\text{pdf}_{\hat{\lambda}_{\text{ML}}}(z|q_{n_1}, q_{n_2}) = \frac{\exp\{-\frac{1}{2}(q_{n_1} + q_{n_2})\}}{2^{\frac{n_3}{2}}\Gamma(\frac{n_3}{2})}(\psi_1' q_{n_1} + \psi_2' q_{n_2})(\psi_1 q_{n_1} + \psi_2 q_{n_2})^{\frac{n_3}{2}-1},$$

and, for $z < z_2$,

$$\text{pdf}_{\hat{\lambda}_{\text{ML}}}(z|q_{n_2}, q_{n_3}) = -\frac{\exp\{-\frac{1}{2}(q_{n_2} + q_{n_3})\}}{2^{\frac{n_1}{2}}\Gamma(\frac{n_1}{2})}(\tilde{\psi}_2' q_{n_2} + \tilde{\psi}_3' q_{n_3})(\tilde{\psi}_2 q_{n_2} + \tilde{\psi}_3 q_{n_3})^{\frac{n_1}{2}-1},$$

(the dash here denotes differentiation with respect to z). To obtain the unconditional densities we simply need to the average these expressions with respect to the conditioning variates.

The following Lemma gives a general result that can be used to obtain the density of $\hat{\lambda}_{\text{ML}}$ in each case. Recall that $\mathcal{G}_v$ denotes the cdf of a χ^2_v random variable.

Lemma 8.1. *Let $a(z)$ and $c(z)$ be positive functions of z on some interval Λ_0. Let $q_1 \sim \chi^2_\alpha$, $q_2 \sim \chi^2_\beta$ be independent, and assume that the conditional cdf of the random variable w, given (q_1, q_2), is defined by*

$$\Pr(w \le z | q_1, q_2) = \mathcal{G}_\gamma(a(z)q_1 + c(z)q_2)$$

Then, denoting the density of w at $w = z$ when the parameters are (α, β, γ) by $\text{pdf}_w(z; \alpha, \beta, \gamma)$, we have:

(i) for $\gamma = 2s + 2$, $s = 0, 1, ...,$ the unconditional density is

$$\text{pdf}_w(z; \alpha, \beta, 2s{+}2) = \frac{1}{2(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}}\left[\frac{\alpha\dot{a}}{1+a}h^{(s)}_{\alpha+2,\beta}(\varphi_1, \varphi_2) + \frac{\beta\dot{c}}{1+c}h^{(s)}_{\alpha,\beta+2}(\varphi_1, \varphi_2)\right], \tag{8.9}$$

where $\varphi_1 := a/(1+a)$, $\varphi_2 := c/(1+c)$, *and we have defined functions*

$$h^{(s)}_{\alpha,\beta}(x_1,x_2) := \frac{1}{s!}\sum_{i=0}^{s}\binom{s}{i}x_1^i x_2^{s-i}\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i},$$

(with $h^{(0)}_{\alpha,\beta}(x_1,x_2) := 1$*). When* $\gamma = 2$ $(s = 0)$ *this reduces to the very simple form*

$$\mathrm{pdf}_w(z;\alpha,\beta,2) = \frac{\alpha\frac{\dot{a}}{1+a} + \beta\frac{\dot{c}}{1+c}}{2(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}}.$$

(ii) for $\gamma = 1$ *the unconditional density is*

$$\mathrm{pdf}_w(z;\alpha,\beta,1) = \frac{B\left(\frac{\alpha+\beta+1}{2},\frac{1}{2}\right)}{2\pi(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}}\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{\left(\frac{\alpha+\beta+2}{2}\right)_s} \times\left[\frac{\alpha\dot{a}h^{(s)}_{\alpha+2,\beta}(1-\varphi_1,1-\varphi_2)}{(1+a)} + \frac{\beta\dot{c}h^{(s)}_{\alpha,\beta+2}(1-\varphi_1,1-\varphi_2)}{(1+c)}\right]. \tag{8.10}$$

(iii) for $\gamma = 2s+1$, $s = 0, 1, ...,$ *the unconditional density is*

$$\mathrm{pdf}_w(z;\alpha,\beta,2s{+}1) = \frac{1}{\left(\frac{1}{2}\right)_s}\sum_{i=0}^{s}\binom{s}{i}a^i c^{s-i}\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i}\mathrm{pdf}_w(z;\alpha{+}2i,\beta{+}2(s{-}i);1). \tag{8.11}$$

An ordinary generating function for the functions $h^{(s)}_{\alpha,\beta}(x_1,x_2)$ *is*

$$g_{\alpha,\beta}(\theta) := \sum_{s=0}^{\infty}\theta^s h^{(s)}_{\alpha,\beta}(x_1,x_2) = (1-\theta x_1)^{-\frac{\alpha}{2}}(1-\theta x_2)^{-\frac{\beta}{2}}. \tag{8.12}$$

Proof of Lemma 8.1. The conditional density of w given (q_1, q_2) is

$$\begin{aligned}\mathrm{pdf}_w(z|q_1,q_2) &= \dot{\mathcal{G}}_\gamma(aq_1 + cq_2)\\ &= \frac{\exp\{-\frac{1}{2}(aq_1+cq_2)\}}{2^{\frac{\gamma}{2}}\Gamma(\frac{\gamma}{2})}(\dot{a}q_1 + \dot{c}q_2)(aq_1+cq_2)^{\frac{\gamma}{2}-1}\end{aligned}$$

and the unconditional density is the expectation of this expression with respect to $q_1 \sim \chi^2(\alpha)$ and $q_2 \sim \chi^2(\beta)$.

(i) If $\gamma = 2s + 2$, the last term in the conditional density can be expanded Binomially and integrated term by term. A simple calculation gives

$$\begin{aligned}
\mathrm{pdf}_w(z;\alpha,\beta,2s+2) &= \frac{1}{2^{s+\frac{\alpha+\beta+2}{2}} s!\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})} \int_{q_1>0}\int_{q_2>0} \exp\{-\frac{1}{2}((1+a)q_1 + (1+c)q_2)\} \\
&\quad \times (\dot{a}q_1 + \dot{c}q_2) \sum_{i=0}^{s} \binom{s}{i} a^i c^{s-i} q_1^{i+\frac{\alpha}{2}-1} q_2^{s-i+\frac{\beta}{2}-1} (dq_1)(dq_2) \\
&= \frac{\alpha\dot{a}}{2s!(1+a)^{\frac{\alpha}{2}+1}(1+c)^{\frac{\beta}{2}}} \sum_{i=0}^{s} \binom{s}{i} \varphi_1^i \varphi_2^{s-i} \left(\frac{\alpha}{2}+1\right)_i \left(\frac{\beta}{2}\right)_{s-i} \\
&\quad + \frac{\beta\dot{c}}{2s!(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}+1}} \sum_{i=0}^{s} \binom{s}{i} \varphi_1^i \varphi_2^{s-i} \left(\frac{\alpha}{2}\right)_i \left(\frac{\beta}{2}+1\right)_{s-i} \\
&= \frac{1}{2(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}} \left[\frac{\alpha\dot{a}}{1+a} h^{(s)}_{\alpha+2,\beta}(\varphi_1,\varphi_2) + \frac{\beta\dot{c}}{1+c} h^{(s)}_{\alpha,\beta+2}(\varphi_1,\varphi_2)\right].
\end{aligned}$$

The result for $\gamma = 2$ is the case $s = 0$, when $h^{(0)}_{\alpha+2,\beta} = h^{(0)}_{\alpha,\beta+2} = 1$.

(ii) Case (ii). When $\gamma = 1$ the last term involved in the conditional density is

$$(aq_1 + cq_2)^{-\frac{1}{2}} = \frac{1}{\sqrt{2\pi}} \int_{x>0} \exp\{-\frac{1}{2}x(aq_1 + cq_2)\} x^{-\frac{1}{2}} dx.$$

Using this to evaluate the expectation with respect to (q_1, q_2) (and reversing the order of integration) gives the following integral expression for the unconditional density:

$$\frac{1}{2\pi}\int_{x>0} x^{-\frac{1}{2}} \left[\alpha\dot{a}(1+a+ax)^{-(\frac{\alpha}{2}+1)}(1+c+cx)^{-\frac{\beta}{2}} + \beta\dot{c}(1+a+ax)^{-\frac{\alpha}{2}}(1+c+cx)^{-(\frac{\beta}{2}+1)}\right] dx.$$

Each term is of the form

$$\begin{aligned}
&\frac{1}{2\pi}\int_{x>0} x^{-\frac{1}{2}}(1+a+ax)^{-\frac{\alpha}{2}}(1+c+cx)^{-\frac{\beta}{2}} dx \\
&= \frac{1}{2\pi}(1+a)^{-\frac{\alpha}{2}}(1+c)^{-\frac{\beta}{2}} \int_{x>0} x^{-\frac{1}{2}}(1+\varphi_1 x)^{-\frac{\alpha}{2}}(1+\varphi_2 x)^{-\frac{\beta}{2}} dx.
\end{aligned}$$

Transform to $b := x/(1+x)$, so the integral becomes

$$\begin{aligned}
&\int_{0<b<1} b^{-\frac{1}{2}}(1-b)^{\frac{\alpha+\beta-1}{2}-1}(1-\frac{b}{1+a})^{-\frac{\alpha}{2}}(1-\frac{b}{1+c})^{-\frac{\beta}{2}}db \\
&= B\left(\frac{\alpha+\beta-1}{2},\frac{1}{2}\right)\sum_{s,i=0}^{\infty}\frac{\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_s}{s!i!(1+a)^i(1+c)^s}\frac{\left(\frac{1}{2}\right)_{i+s}}{\left(\frac{\alpha+\beta}{2}\right)_{i+s}} \\
&= B\left(\frac{\alpha+\beta-1}{2},\frac{1}{2}\right)\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{s!\left(\frac{\alpha+\beta}{2}\right)_s}\sum_{i=0}^{s}\binom{s}{i}\frac{\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i}}{(1+a)^i(1+c)^{s-i}} \\
&= B\left(\frac{\alpha+\beta-1}{2},\frac{1}{2}\right)\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{\left(\frac{\alpha+\beta}{2}\right)_s}h^{(s)}_{\alpha,\beta}(1-\varphi_1,1-\varphi_2).
\end{aligned}$$

(The validity of the series expansions used for the Bessel functions $(1-b/(1+a))^{-\alpha/2}$ and $(1-b/(1+c))^{-\beta/2}$, as well as of the term-by term integration involved, are consequences of the fact that $\varphi_1,\varphi_2 \in (0,1)$).

For the first term, replace α by $\alpha+2$, giving

$$\frac{\alpha\dot{a}B\left(\frac{\alpha+\beta+1}{2},\frac{1}{2}\right)}{2\pi(1+a)^{\frac{\alpha}{2}+1}(1+c)^{\frac{\beta}{2}}}\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{\left(\frac{\alpha+\beta+2}{2}\right)_s}h^{(s)}_{\alpha+2,\beta}(1-\varphi_1,1-\varphi_2)$$

and for the second replace β by $\beta+2$, giving

$$\frac{\beta\dot{c}B\left(\frac{\alpha+\beta+1}{2},\frac{1}{2}\right)}{2\pi(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}+1}}\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{\left(\frac{\alpha+\beta+2}{2}\right)_s}h^{(s)}_{\alpha,\beta+2}(1-\varphi_1,1-\varphi_2).$$

We therefore have, when $\gamma=1$,

$$\begin{aligned}
\mathrm{pdf}_w(z;\alpha,\beta,1) &= \frac{B\left(\frac{\alpha+\beta+1}{2},\frac{1}{2}\right)}{2\pi(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}}\sum_{s=0}^{\infty}\frac{\left(\frac{1}{2}\right)_s}{\left(\frac{\alpha+\beta+2}{2}\right)_s} \\
&\quad\times\left[\frac{\alpha\dot{a}h^{(s)}_{\alpha+2,\beta}(1-\varphi_1,1-\varphi_2)}{(1+a)}+\frac{\beta\dot{c}h^{(s)}_{\alpha,\beta+2}(1-\varphi_1,1-\varphi_2)}{(1+c)}\right].
\end{aligned}$$

(iii) For the case $\gamma = 2s+1$ the final term in the conditional density is $(aq_1 +$

$cq_2)^{s-\frac{1}{2}}$. Expanding the term $(aq_1+cq_2)^s$ we have

$$\begin{aligned}\mathrm{pdf}_w(z;\alpha,\beta,2s+1) &= \frac{1}{\sqrt{2}\Gamma(s+\frac{1}{2})}\sum_{i=0}^{s}\binom{s}{i}a^i c^{s-i}\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i}\\ &\times\int_{q_1>0}\int_{q_2>0}\frac{\exp\{-\frac{1}{2}(aq_1+cq_2)\}}{2^{s+\frac{\alpha+\beta}{2}}\Gamma(\frac{\alpha}{2}+i)\Gamma(\frac{\beta}{2}+s-i)}\\ &\times q_1^{\frac{\alpha}{2}+i-1}q_2^{\frac{\beta}{2}+s-i-1}(\dot{a}q_1+\dot{c}q_2)(aq_1+cq_2)^{-\frac{1}{2}}dq_1dq_2\\ &= \frac{1}{\left(\frac{1}{2}\right)_s}\sum_{i=0}^{s}\binom{s}{i}a^i c^{s-i}\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i}\mathrm{pdf}_w(z;\alpha+2i,\beta+2(s-i);1)\end{aligned}$$

as stated.

Generating Function. Straightforward manipulation of the double series using the device of "summing by diagonals" gives

$$\begin{aligned}\sum_{s=0}^{\infty}\theta^s h_{\alpha,\beta}^{(s)}(x_1,x_2) &= \sum_{s=0}^{\infty}\frac{\theta^s}{s!}\sum_{i=0}^{s}\binom{s}{i}x_1^i x_2^{s-i}\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_{s-i}\\ &= \sum_{i,s=0}^{\infty}\frac{\theta^{s+i}}{i!s!}x_1^i x_2^s\left(\frac{\alpha}{2}\right)_i\left(\frac{\beta}{2}\right)_s\\ &= (1-\theta x_1)^{-\frac{\alpha}{2}}(1-\theta x_2)^{-\frac{\beta}{2}}.\end{aligned}$$

□

Under an additional restriction on the functions $a(z)$ and $c(z)$ these results can be expressed quite simply in terms of the Gaussian hypergeometric function.

Lemma 8.2. *If, in the same context as Lemma 8.1, $a(1+c)\leq 2c(1+a)$ for all $z\in\Lambda_0$, then the results in Lemma 8.1 can be written more simply as*

$$\begin{aligned}\mathrm{pdf}_w(z;\alpha,\beta,\gamma) = \frac{B\left(\frac{\alpha+\beta}{2},\frac{\gamma}{2}\right)a^{\frac{\gamma+\beta}{2}}}{c^{\frac{\beta}{2}}(1+a)^{\frac{\alpha+\beta+\gamma}{2}}}&\left[\frac{\alpha\dot{a}}{(\alpha+\beta)a}{}_2F_1\left(\frac{\alpha+\beta+\gamma}{2},\frac{\alpha+2}{2},\frac{\alpha+\beta+2}{2},\eta\right)\right.\\ &\left.+\frac{\beta\dot{c}}{(\alpha+\beta)c}{}_2F_1\left(\frac{\alpha+\beta+\gamma}{2},\frac{\alpha}{2},\frac{\alpha+\beta+2}{2},\eta\right)\right], \quad (8.13)\end{aligned}$$

where

$$\eta := 1-\frac{a(1+c)}{c(1+a)}.$$

Proof of Lemma 8.2. The conditional density of w given (q_1, q_2) is

$$\mathrm{pdf}_w(z|q_1, q_2) = \frac{\exp\{-\frac{1}{2}(aq_1 + cq_2)\}}{2^{\frac{\gamma}{2}}\Gamma(\frac{\gamma}{2})}(\dot{a}q_1 + \dot{c}q_2)(aq_1 + cq_2)^{\frac{\gamma}{2}-1}.$$

Multiply by the joint density of (q_1, q_2) and transform to $x_1 := (1 + a)q_1$ and $x_2 := (1 + c)q_2$ to get

$$\mathrm{pdf}_w(z, x_1, x_2) = \frac{\exp\{-\frac{1}{2}(x_1 + x_2)\}}{2^{\frac{\gamma+\alpha+\beta}{2}}\Gamma(\frac{\gamma}{2})\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}} x_1^{\frac{\alpha}{2}-1} x_2^{\frac{\beta}{2}-1} \times \left(\frac{\dot{a}}{1+a}x_1 + \frac{\dot{c}}{1+c}x_2\right)\left(\frac{a}{1+a}x_1 + \frac{c}{1+c}x_2\right)^{\frac{\gamma}{2}-1}.$$

Note that if $\gamma = 2$ the last term is not present and, on integrating out x_1, x_2, we obtain the result given in Lemma 8.1. For the general case, transforming to $q := x_1 + x_2$ and $b := x_1/q$, and integrating out q, we obtain

$$\mathrm{pdf}_w(z, b) = \frac{\Gamma(\frac{\alpha+\beta+\gamma}{2})\left(\frac{c}{1+c}\right)^{\frac{\gamma}{2}-1}}{\Gamma(\frac{\gamma}{2})\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})(1+a)^{\frac{\alpha}{2}}(1+c)^{\frac{\beta}{2}}} \times \left[\frac{\dot{a}b}{1+a} + \frac{\dot{c}(1-b)}{1+c}\right] b^{\frac{\alpha}{2}-1}(1-b)^{\frac{\beta}{2}-1}(1-\eta b)^{\frac{\gamma}{2}-1}.$$

Integrating out b in the last line is a standard result, provided $|\eta| < 1$, and gives two terms (ignoring the constant in the first line for the moment):

$$\frac{\dot{a}}{1+a}\frac{\Gamma(\frac{\beta}{2})\Gamma(\frac{\alpha+2}{2})}{\Gamma(\frac{\alpha+\beta+2}{2})}\left(\frac{a(1+c)}{c(1+a)}\right)^{\frac{\gamma+\beta}{2}-1} {}_2F_1\left(\frac{\alpha+\beta+\gamma}{2}, \frac{\beta}{2}; \frac{\alpha+\beta+2}{2}; \eta\right)$$

and

$$\frac{\dot{c}}{1+c}\frac{\Gamma(\frac{\beta+2}{2})\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{\alpha+\beta+2}{2})}\left(\frac{a(1+c)}{c(1+a)}\right)^{\frac{\gamma+\beta}{2}} {}_2F_1\left(\frac{\alpha+\beta+\gamma}{2}, \frac{\beta+2}{2}; \frac{\alpha+\beta+2}{2}; \eta\right).$$

Multiplying by the remaining constant gives the result stated. □

9 Additional Results for the Examples

9.1 Balanced Group Interaction Model

Further details can be found in Hillier and Martellosio (2013).

9.2 Group Interaction Model

In the unbalanced Group Interaction model with $u = 2$ the following parameters and functions in Lemma 8.1 occur: on the interval $-(m_1 - 1) < z < z_2$, $(\alpha, \beta, \gamma) = (r, n_2, n_1)$ and $(a(z), c(z)) = (\psi_{22}, \psi_{23})$, and on the interval $z_2 < z < 1$, $(\alpha, \beta, \gamma) = (n_1, n_2, r)$ and $(a(z), c(z)) = (\psi_{11}, \psi_{12})$. The density is simplest when both r and $n_1 = r_1(m_1 - 1)$ are even, and we restrict attention to this situation for brevity. The formulae make use of the functions $h^{(s)}_{\alpha,\beta}(x_1, x_2)$ defined in Lemma 8.1. The two components of the density are given in the next result.

Proposition 9.1. *In the unbalanced Group interaction model with* $\varepsilon \sim \mathrm{SMN}(0, \sigma^2 I_n)$, $r = 2s_1 + 2$, *and* $n_1 = 2s_2 + 2$ $(s_1, s_2 = 0, 1, ...)$, *the density of* $\hat{\lambda}_{\mathrm{ML}}$ *is given, for* $-(m_1 - 1) < z < z_2$, *by*

$$\mathrm{pdf}_{\hat{\lambda}_{\mathrm{ML}}}(z; \lambda) = \frac{\frac{r\dot{\psi}_{22}}{1+\psi_{22}} h^{(s_1)}_{r+2,n_2}(\varphi_{11}, \varphi_{12}) + \frac{n_2\dot{\psi}_{23}}{1+\psi_{23}} h^{(s_1)}_{r,n_2+2}(\varphi_{11}, \varphi_{12})}{2(1+\psi_{22})^{\frac{r}{2}}(1+\psi_{23})^{\frac{n_2}{2}}}, \tag{9.1}$$

and for $z_2 < z < 1$ *by*

$$\mathrm{pdf}_{\hat{\lambda}_{\mathrm{ML}}}(z; \lambda) = -\frac{\frac{n_1\dot{\psi}_{11}}{1+\psi_{11}} h^{(s_2)}_{n_1+2,n_2}(\varphi_{21}, \varphi_{22}) + \frac{n_2\dot{\psi}_{12}}{1+\psi_{12}} h^{(s_2)}_{n_1,n_2+2}(\varphi_{21}, \varphi_{22})}{2(1+\psi_{11})^{\frac{n_1}{2}}(1+\psi_{12})^{\frac{n_2}{2}}}, \tag{9.2}$$

where the dot as usual denotes differentiation with respect to z, $\varphi_{11} := \psi_{22}/(1+\psi_{22})$, $\varphi_{12} := \psi_{23}/(1+\psi_{23})$, $\varphi_{21} := \psi_{11}/(1+\psi_{11})$, *and* $\varphi_{22} := \psi_{12}/(1+\psi_{12})$.

If $s_1 = s_2 = 0$ (i.e., $r = 2$ and $m_1 = 3$) these formulae become even simpler, reducing to

$$\begin{aligned} \mathrm{pdf}_{\hat{\lambda}_{\mathrm{ML}}}(z; \lambda) &= \frac{\frac{2\dot{\psi}_{22}}{1+\psi_{22}} + \frac{(m_2-1)\dot{\psi}_{23}}{1+\psi_{23}}}{2(1+\psi_{22})(1+\psi_{23})^{\frac{m_2-1}{2}}}, \text{ for } -2 < z < z_2, \\ &= -\frac{\frac{2\dot{\psi}_{11}}{1+\psi_{11}} + \frac{(m_2-1)\dot{\psi}_{12}}{1+\psi_{12}}}{2(1+\psi_{11})(1+\psi_{12})^{\frac{m_2-1}{2}}}, \text{ for } z_2 < z < 1, \end{aligned} \tag{9.3}$$

with $z_2 = -2n/(n + 3(m_2 - 3)) < 0$.

Figures 6 and 7 complement Figure 2 in the paper. They are obtained using equation (9.3), Lemma 8.1 and Proposition 9.1. Each of the three rows of Figure 6 displays $\mathrm{pdf}_{\hat{\lambda}_{\mathrm{ML}}}(z; \lambda)$ for a fixed value of m_1 and varying n, while Figure 7 displays $\mathrm{pdf}_{\hat{\lambda}_{\mathrm{ML}}}(z; \lambda)$ for fixed n and varying m_1. The parameter space is $\Lambda = (-(m_1 - 1), 1)$. For convenience, all densities are plotted on $(-2, 1) \subset \Lambda$. Recall that as long as the model is unbalanced (i.e., $m_1 < m_2$), there is a point $z_2 \in \Lambda$

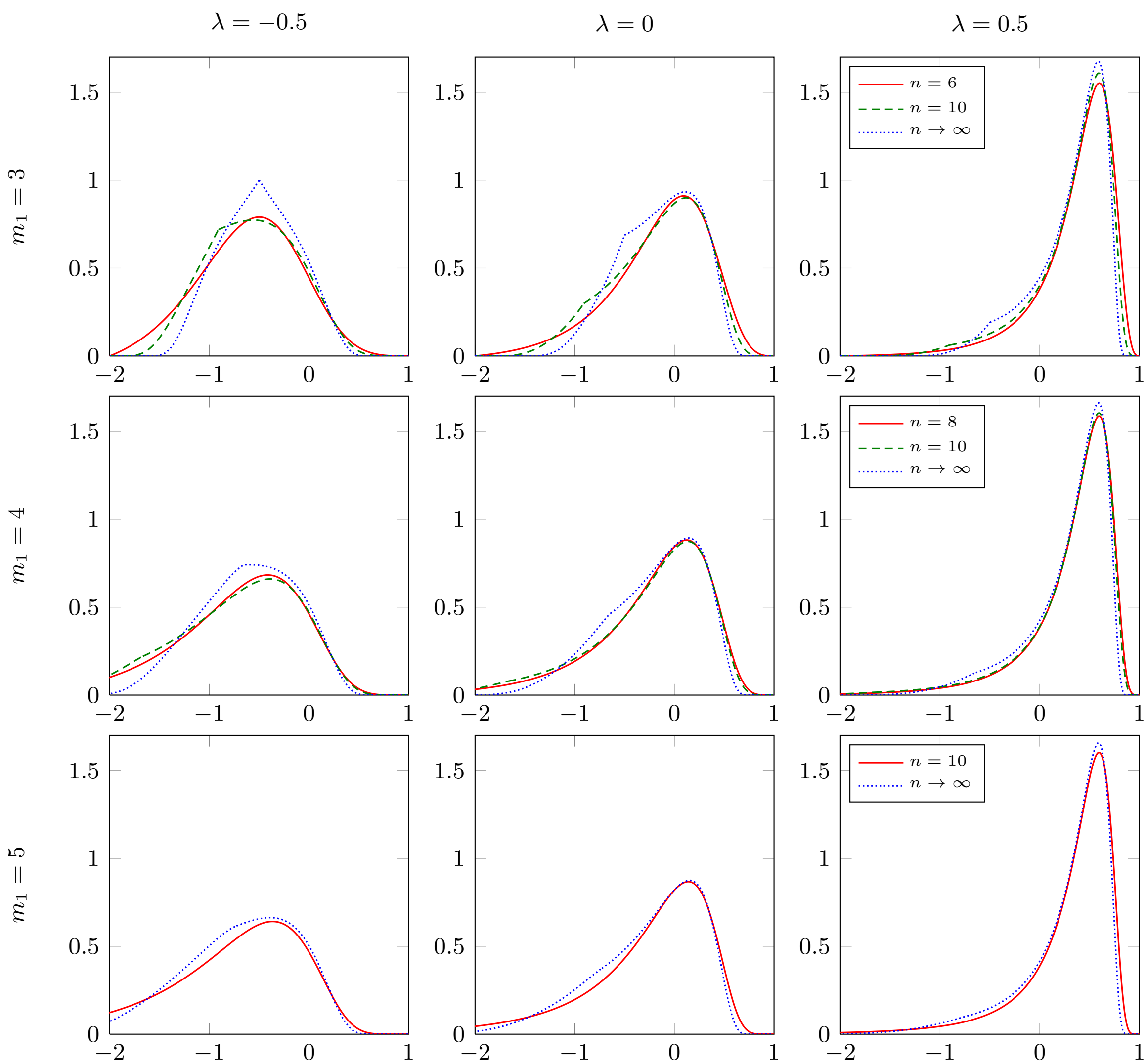

Figure 6: Density of $\hat{\lambda}_{\text{ML}}$ for the Gaussian pure Group Interaction model with two groups.

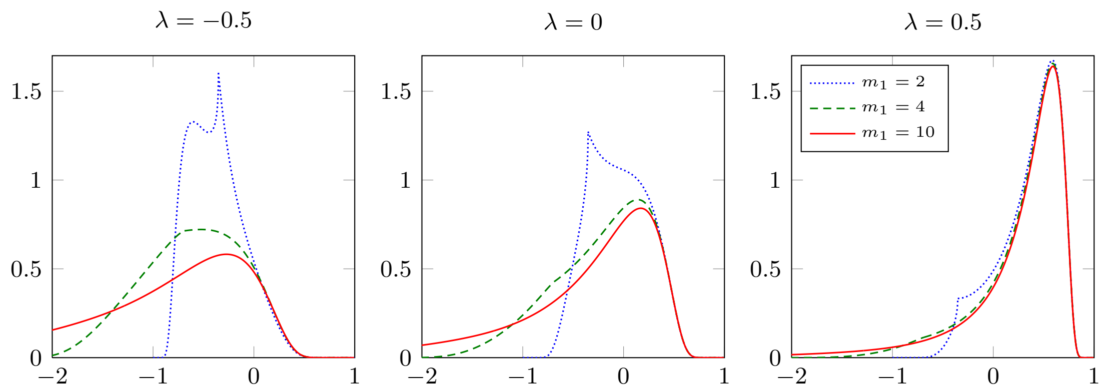

Figure 7: Density of $\hat{\lambda}_{\text{ML}}$ for the Gaussian pure row-standardized Group Interaction model with two groups and $n = 25$.

where the density of $\hat{\lambda}_{\text{ML}}$ in nonanalytic, whatever the sample size n. Graphically, nonanalyticity is clearly visible only for small m_1; at $m_1 = 6$ it becomes already difficult to notice.

The approach used here to obtain the density from the cdf by first conditioning on some of the variables involved, and treating each component of the density separately, can be extended to the general model with $u > 2$. However, there will then be $T - 1$ intervals in which a separate computation needs to be done, and the process obviously becomes very cumbersome. We do not pursue this further here; it is a subject for future research.

9.3 Asymptotics for the Pure Group Interaction Model

The representation of the cdf of $\hat{\lambda}_{\text{ML}}$ in Theorem 3 provides a useful starting point for deriving asymptotic properties of $\hat{\lambda}_{\text{ML}}$ under the mixed Gaussian assumption. Different asymptotic regimes are possible depending on how the m_t's and the r_t's are assumed to vary with n. According to equation (4.11) the cdf of $\hat{\lambda}_{\text{ML}}$ in a pure Group Interaction model is

$$\Pr(\hat{\lambda}_{\text{ML}} \leq z) = \Pr\left(\sum_{t=1}^{u} d_{tt}(x_t' x_t) + d_{u+1,u+1}(x_{u+1}' x_{u+1}) \leq 0\right).$$

From the definitions of the d_{tt} in the paper we find that, on dividing through by $d_{u+1,u+1}$, this becomes

$$\Pr(\hat{\lambda}_{\text{ML}} \leq z) = \Pr\left(\chi_r^2 - \sum_{t=1}^{u} \psi_t \chi^2_{r_t(m_t-1)} \leq 0\right). \tag{9.4}$$

where the coefficients ψ_t are given by

$$\psi_t := \left(-\frac{d_{tt}}{d_{u+1,u+1}}\right) = \left(\frac{(1-\lambda)(z+m_t-1)}{(1-z)(\lambda+m_t-1)}\right)^2 \frac{\sum_{i=1}^{u} r_i m_i \left(\frac{1-z}{z+m_t-1} + \frac{z}{z+m_i-1}\right)}{\sum_{i=1}^{u} r_i m_i \left(\frac{m_i-1}{z+m_i-1}\right)},$$

for $t = 1, ..., u$.

Assuming u, the number of different group sizes, is fixed, one can consider two types of asymptotic regime. The first, infill asymptotics, holds the r_t's fixed (hence also r), and assumes one or more of the m_t's produce the increased sample size. The second, fixed-domain asymptotics, holds the m_t's fixed and assumes an increase in one or more of the r_t's. This second case satisfies the assumptions in Lee (2004). Hence, it is already known that, under regularity conditions, $\hat{\lambda}_{\mathrm{ML}}$ is consistent and asymptotically normal. In the first case Lee's (2004) results leave the properties of $\hat{\lambda}_{\mathrm{ML}}$ open.

It is clear at once that in the first case, convergence will be to a random variable, because the first term in (9.4) will be unaffected. Precise details for this situation depend on exactly what is assumed about the behavior of the m_t's, but $\hat{\lambda}_{\mathrm{ML}}$ is clearly again inconsistent under infill asymptotics. In the second case the known results are easily deduced from the representation (9.4) by a characteristic function argument.