Gaussian Correlation Inequality

Let μ be a centered Gaussian measure on $\mathbb{R}^d$. The statement of the Gaussian correlation inequality is as follows.

Theorem 1 (Royen [5])**.** *For any convex symmetric sets K, L in $\mathbb{R}^d$,*

$$\mu(K \cap L) \geq \mu(K)\mu(L). \tag{1}$$

For $d = 2$ the result was proved by Pitt [4]. In the case when one of the sets K, L is a symmetric strip (which corresponds to $\min\{n_1, n_2\} = 1$ in Theorem 2 below) inequality (1) was established independently by Khatri [3] and Šidák [7]. Hargé [2] generalized the Khatri-Šidak result to the case when one of the sets is a symmetric ellipsoid. Some other partial results may be found in papers of Borell [1] and Schechtman, Schlumprecht and Zinn [6].

Since any symmetric convex set is a countable intersection of symmetric strips, it is enough to show (1) in the case when

$$K = \Big\{x \in \mathbb{R}^d \mid \forall_{1 \leq i \leq n_1} \ |\langle x, v_i \rangle| \leq t_i\Big\},$$
$$L = \Big\{x \in \mathbb{R}^d \mid \forall_{n_1 < i \leq n_1 + n_2} \ |\langle x, v_i \rangle| \leq t_i\Big\},$$

where v_i are vectors in $\mathbb{R}^d$ and t_i nonnegative numbers. If we set $n = n_1 + n_2$, $X_i := \langle v_i, G\rangle$, where G is the Gaussian random vector distributed according to μ, we obtain the following equivalent form of Theorem 1.

Theorem 2. *For any $t_1, \ldots, t_n > 0$,*

$$\begin{aligned}\mathbb{P}\big(|X_1| \leq t_1, \ldots, |X_n| \leq t_n\big) \geq \ &\mathbb{P}\big(|X_1| \leq t_1, \ldots, |X_{n_1}| \leq t_{n_1}\big) \\ &\times \mathbb{P}\big(|X_{n_1+1}| \leq t_{n_1+1}, \ldots, |X_n| \leq t_n\big).\end{aligned}$$

Royen established this result for a more general class of random vectors X such that $X^2 = (X_1^2, \ldots, X_n^2)$ has an n–variate gamma distribution (see [5] for details). We will emphasize important ideas in the proof of Theorem 2 by breaking it into several steps.

Notation. By $\mathcal{N}(0, C)$ we denote the centered Gaussian measure with the covariance matrix C. For a matrix $A = (a_{ij})_{i,j \leq n}$, we denote by $|A|$ the determinant of A. Given any subset $J \subset [n] := \{1, \ldots, n\}$, we denote by A_J the square matrix $(a_{ij})_{i,j \in J}$ and by $|J|$ the cardinality of J.

Interpolation. Without loss of generality we may and will assume that the covariance matrix C of X is strictly positive-definite. We may write C as

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where C_{ij} is the $n_i \times n_j$ matrix. *The idea of interpolation for Gaussian distributions is ubiquitous and, probably, this first step has been tried by everyone who spent time thinking about this problem.*

Let us consider an interpolation parameter $0 \le \tau \le 1$ and let

$$C(\tau) := \begin{bmatrix} C_{11} & \tau C_{12} \\ \tau C_{21} & C_{22} \end{bmatrix} \tag{2}$$

If we denote by $X(\tau)$ the Gaussian random vector with the distribution $\mathcal{N}(0, C(\tau))$ then Theorem 2 is equivalent to

$$\mathbb{P}\big(|X_1(1)| \le t_1, \ldots, |X_n(1)| \le t_n\big) \ge \mathbb{P}\big(|X_1(0)| \le t_1, \ldots, |X_n(0)| \le t_n\big).$$

This will be done by showing that the function

$$\tau \mapsto \mathbb{P}\big(|X_1(\tau)| \le t_1, \ldots, |X_n(\tau)| \le t_n\big)$$

is nondecreasing on $[0,1]$.

Using symmetry. Next important idea is to utilize the symmetry of the constraints $X_i(\tau) \in [-t_i, t_i]$. This is done by squaring and rewriting them as $Z_i(\tau) \in [0, s_i]$, where

$$Z_i(\tau) := \frac{1}{2} X_i(\tau)^2,\ s_1 := \frac{1}{2} t_i^2. \tag{3}$$

The factor $1/2$ here simplifies notation later. We would like to show that the function

$$\tau \mapsto \mathbb{P}\big(Z_1(\tau) \le s_1, \ldots, Z_n(\tau) \le s_n\big)$$

is nondecreasing on $[0,1]$. *Such a simple restatement is, in fact, quite a powerful way to use the symmetry of the problem.*

If $f(x,\tau)$ is the density of the random vector $Z(\tau)$ and $K = [0, s_1] \times \cdots \times [0, s_n]$,

$$\mathbb{P}\big(Z_1(\tau) \le s_1, \ldots, Z_n(\tau) \le s_n\big) = \int_K f(x,\tau)\, dx. \tag{4}$$

If we rewrite the covariance matrix $C(\tau)$ as

$$\begin{bmatrix} C_{11} & \tau C_{12} \\ \tau C_{21} & C_{22} \end{bmatrix} = \tau \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} + (1-\tau) \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix},$$

it is clear that it is uniformly strictly positive-definite over $\tau \in [0,1]$ when C is non-degenerate. In particular, the random vector $X(\tau) \sim \mathcal{N}(0, C(\tau))$ has the density

$$f_{C(\tau)}(x) = \frac{1}{\sqrt{|C(\tau)|(2\pi)^n}} \exp\Big(-\frac{1}{2}\langle C(\tau)^{-1} x, x\rangle\Big).$$

If we write

$$\mathbb{P}\big(Z_1(\tau) \le x_1, \dots, Z_1(\tau) \le x_n\big) = \int_{-\sqrt{2x_1}}^{\sqrt{2x_1}} \cdots \int_{-\sqrt{2x_n}}^{\sqrt{2x_n}} f_{C(\tau)}(x)\, dx,$$

then taking derivatives in $x_1, \dots, x_n$, we see that the density of $Z(\tau)$ on $(0,\infty)^n$ equals

$$f(x,\tau) = \frac{1}{\sqrt{|C(\tau)|(4\pi)^n}} \frac{1}{\sqrt{x_1 \cdots x_n}} \sum_{\varepsilon \in \{-1,1\}^n} \exp\Big(-\big\langle C(\tau)^{-1}\varepsilon\sqrt{x}, \varepsilon\sqrt{x}\big\rangle\Big), \tag{5}$$

where for $\varepsilon \in \{-1,1\}^n$ and $x \in (0,\infty)^n$ we set $\varepsilon\sqrt{x} := (\varepsilon_i\sqrt{x_i})_{i\le n}$. Using this explicit formula for the density, it is not difficult to show that

$$\frac{\partial}{\partial\tau}\mathbb{P}\big(Z_1(\tau) \le s_1, \dots, Z_n(\tau) \le s_n\big) = \frac{\partial}{\partial\tau}\int_K f(x,\tau)\, dx = \int_K \frac{\partial}{\partial\tau} f(x,\tau)\, dx, \tag{6}$$

i.e. one can interchange the derivative and integral.

Laplace transform. Next, we appeal to the following general principle,

"if you don't know what to do, use Laplace transform",

which deserves to be kept in mind as closely as another common rule

"if you don't know what to do, integrate by parts".

Using both principles, one can give an *explicit formula* for

$$\frac{\partial}{\partial\tau} f(x,\tau) \quad \text{and} \quad \int_K \frac{\partial}{\partial\tau} f(x,\tau)\, dx$$

by identifying the Laplace transform of the derivative,

$$\int_{(0,\infty)^n} e^{-\sum_{i=1}^n \lambda_i x_i} \frac{\partial}{\partial\tau} f(x,\tau)\, dx.$$

We start again by interchanging the derivative and integral,

$$\frac{\partial}{\partial\tau}\int_{(0,\infty)^n} e^{-\sum_{i=1}^n \lambda_i x_i} f(x,\tau)\, dx.$$

If we denote $\Lambda := \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ then, by a standard computation,

$$\int_{(0,\infty)^n} e^{-\Lambda x} f(x,\tau) dx = \mathbb{E}\exp\Big(-\frac{1}{2}\sum_{i\le n} \lambda_i X_i^2(\tau)\Big) = |I + \Lambda C(\tau)|^{-1/2}. \tag{7}$$

It is obvious (e.g. from the Leibniz formula for the determinant) that

$$|I+\Lambda C(\tau)| = 1 + \sum_{\emptyset\neq J\subset[n]} |(\Lambda C(\tau))_J| = 1 + \sum_{\emptyset\neq J\subset[n]} |C(\tau)_J| \prod_{j\in J} \lambda_j. \tag{8}$$

Fix $\emptyset \neq J \subset [n]$. Then $J = J_1 \cup J_2$, where $J_1 := [n_1] \cap J$, $J_2 := J \setminus [n_1]$ and

$$C(\tau)_J = \begin{bmatrix} C_{J_1} & \tau C_{J_1 J_2} \\ \tau C_{J_2 J_1} & C_{J_2} \end{bmatrix}.$$

If $J_1 = \emptyset$ or $J_2 = \emptyset$ then $C(\tau)_J = C_J$. Given a strictly positive-definite block matrix, we can write

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11}^{1/2} & 0 \\ 0 & A_{22}^{1/2} \end{bmatrix} \begin{bmatrix} I_{n_1} & A_{11}^{-1/2} A_{12} A_{22}^{-1/2} \\ A_{22}^{-1/2} A_{21} A_{11}^{-1/2} & I_{n_2} \end{bmatrix} \begin{bmatrix} A_{11}^{1/2} & 0 \\ 0 & A_{22}^{1/2} \end{bmatrix}$$

and the matrix in the middle can be converted by elementary transformations to

$$\begin{bmatrix} I_{n_1} - A_{11}^{-1/2} A_{12} A_{22}^{-1} A_{21} A_{11}^{-1/2} & 0 \\ 0 & I_{n_2} \end{bmatrix}.$$

First of all, this implies that the determinant

$$\begin{aligned} |C(\tau)_J| &= |C_{J_1}||C_{J_2}| \big| I_{|J_1|} - \tau^2 C_{J_1}^{-1/2} C_{J_1 J_2} C_{J_2}^{-1} C_{J_2 J_1} C_{J_1}^{-1/2} \big| \\ &= |C_{J_1}||C_{J_2}| \prod_{i\leq |J_1|} \big(1 - \tau^2 \mu_{J_1,J_2}(i)\big), \end{aligned} \tag{9}$$

where we denoted by $\mu_{J_1,J_2}(i)$, $1 \leq i \leq |J_1|$ the eigenvalues of

$$C_{J_1}^{-1/2} C_{J_1 J_2} C_{J_2}^{-1} C_{J_2 J_1} C_{J_1}^{-1/2}.$$

The above conversion by elementary transformations applied to $C(1)_J$ also shows that the eigenvalues $\mu_{J_1,J_2}(i)$ belong to $[0,1]$ so, for any $\emptyset \neq J \subset [n]$ and any $\tau \in [0,1]$,

$$a_J(\tau) := -\frac{\partial}{\partial \tau} |C(\tau)_J| \geq 0. \tag{10}$$

Therefore,

$$\begin{aligned} \frac{\partial}{\partial \tau} |I + \Lambda C(\tau)|^{-1/2} &= -\frac{1}{2} |I + \Lambda C(\tau)|^{-3/2} \sum_{\emptyset\neq J\subset[n]} \frac{\partial}{\partial \tau} |C(\tau)_J||\Lambda_J| \\ &= \frac{1}{2} |I + \Lambda C(\tau)|^{-3/2} \sum_{\emptyset\neq J\subset[n]} a_J(\tau) \prod_{j\in J} \lambda_j. \end{aligned} \tag{11}$$

We have thus shown that the Laplace transform of the derivative $\frac{\partial}{\partial\tau} f(x,\tau)$ is

$$\int_{(0,\infty)^n} e^{-\Lambda x} \frac{\partial}{\partial \tau} f(x,\tau)\, dx = \sum_{\emptyset \neq J \subset [n]} \frac{1}{2} a_J(\tau) |I + \Lambda C(\tau)|^{-3/2} \prod_{j \in J} \lambda_j. \tag{12}$$

Identifying the Laplace transform. To identify this Laplace transform, first of all, recall that we saw above that $|I + \Lambda C(\tau)|^{-1/2}$ is the Laplace transform of $Z(\tau)$, or the density $f(x,\tau)$, for which we also wrote down an explicit formula (5). As a result, if we consider the convolution

$$h_\tau(x) := f(x,\tau) \star f(x,\tau) \star f(x,\tau), \tag{13}$$

which is the density of the sum $Z^1(\tau) + Z^2(\tau) + Z^3(\tau)$ of three independent copies of $Z(\tau)$, then the Laplace transform of h_τ is equal to

$$\int_{(0,\infty)^n} e^{-\Lambda x} h_\tau(x)\, dx = |I + \Lambda C(\tau)|^{-3/2}. \tag{14}$$

From here, one can guess that the quantity

$$|I + \Lambda C(\tau)|^{-3/2} \prod_{j \in J} \lambda_j$$

in (12) is the Laplace transform of a function related to h_τ. In fact, the factor $\prod_{j \in J} \lambda_j$ comes out from integration by parts, as follows. For example, by a formal integration by parts gives

$$\int_0^\infty e^{-\Lambda x} \frac{\partial h_\tau}{\partial x_1}\, dx_1 = e^{-\Lambda x} h_\tau(x) \Big|_{x_1=0}^{x_1=\infty} + \lambda_1 \int_0^\infty e^{-\Lambda x} h_\tau(x)\, dx_1.$$

If we assume that the derivative $\frac{\partial h_\tau}{\partial x_1}$ is integrable, at infinity $h_\tau(x)$ grows sub-exponentially, and $\lim_{x_1 \downarrow 0} h_\tau(x) = 0$, then the first term on the right hand side disappears and, integrating in the other coordinates, we get

$$\int_{(0,\infty)^n} e^{-\Lambda x} \frac{\partial h_\tau}{\partial x_1}\, dx = |I + \Lambda C(\tau)|^{-3/2} \lambda_1.$$

If, for $J \subset [n]$, we introduce the notation

$$\partial_J h_\tau = \frac{\partial^{|J|} h_\tau}{\partial x_J} \tag{15}$$

then a similar formal computation, by induction, shows that

$$\int_{(0,\infty)^n} e^{-\Lambda x} \partial_J h_\tau(x)\, dx = |I + \Lambda C(\tau)|^{-3/2} \prod_{j \in J} \lambda_j, \tag{16}$$

assuming that all the derivatives of the form $\partial_J h_\tau$ are integrable, at infinity they grow sub-exponentially, and the limit

$$\lim_{x_i \downarrow 0} \partial_J h_\tau(x) = 0 \text{ for any } i \notin J. \tag{17}$$

We will discuss below why one can expect these properties to hold and how one can actually prove them, but for now let us assume them to be true. Then (12) implies that

$$\frac{\partial}{\partial \tau} f(x,\tau) = \sum_{\emptyset \neq J \subset [n]} \frac{1}{2} a_J(\tau) \partial_J h_\tau(x). \tag{18}$$

Finishing the proof. Assuming the above properties of $h_\tau(x)$ and its derivatives, by integration by parts,

$$\int_K \partial_J h_\tau(x)\, dx = \int_{\prod_{j \in J^c}[0,s_j]} h_\tau(s_J, x_{J^c})\, dx_{J^c} \geq 0,$$

where $J^c = [n] \setminus J$ and $y = (s_J, x_{J^c})$ if $y_i = s_i$ for $i \in J$ and $y_i = x_i$ for $i \in J^c$. This yields

$$\int_K \frac{\partial}{\partial \tau} f(x,\tau)\, dx \geq 0$$

and finishes the proof. What remains is to prove the above properties of h_τ.

Intuition behind the properties of $\mathbf{h_\tau(x)}$. Recall that

$$h_\tau(x) = f(x,\tau) \star f(x,\tau) \star f(x,\tau),$$

where $f(x,\tau)$ can be expressed as in (5). Notice that $f(x,\tau)$ has $1/\sqrt{x_i}$ singularity at zero along each coordinate. If the covariance $C(\tau)$ was identity, $f(x,\tau)$ would be a product

$$f(x,\tau) = \prod_{i \leq n} \frac{1}{\sqrt{\pi x_i}} e^{-x_i} \mathrm{I}(x_i > 0).$$

Each coordinate here has $\Gamma(1/2,1)$ distribution, and the k–fold convolution is the product of $\Gamma(k/2,1)$ distributions. In particular, for $k=3$ we would have

$$h_\tau(x) = \prod_{i \leq n} \frac{2}{\sqrt{\pi}} \sqrt{x_i} e^{-x_i} \mathrm{I}(x_i > 0).$$

It is easy to see that the derivatives of the form $\partial_J h_\tau$ are integrable and $\lim_{x_i \downarrow 0} \partial_J h_\tau(x) = 0$ for any $i \notin J$ because of the factor $\sqrt{x_i}$. The problem of course is that the covariance $C(\tau)$ is not identity, so the coordinates are correlated. We will see next that one can check these properties of h_τ easily by utilizing the properties of noncentral chi-squared distributions, but it would interesting to find a more straightforward way to obtain the properties of the above 3–fold convolution knowing the behaviour of the singularities at zero.

Noncentral chi-squared distributions. We recall that $f(x,\tau)$ is the density of

$$Z(\tau)=\frac{1}{2}\big(X_1(\tau)^2,\ldots,X_n(\tau)^2\big)$$

where $X(\tau)\sim\mathcal{N}(0,C(\tau))$, and we will obtain the properties of $h_\tau(x)$ by *'peeling off'* a standard Gaussian component with the identity covariance from $X(\tau)$ (this is where noncentral chi-squared distributions and their properties come into play). Namely, since $C(\tau)$ is strictly positive-definite, there exists a small enough $\delta>0$ such that $C(\tau)-\delta I_n$ is positive-definite. Consider independent Gaussian vectors

$$Y\sim\mathcal{N}\big(0,C(\tau)-\delta I_n\big),\ g\sim\mathcal{N}(0,I_n),$$

so that $X(\tau)$ is equal in distribution to $Y+\sqrt{\delta}g$. Let Y^ℓ and g^ℓ be their independent copies for $\ell=1,2,3$. If we let

$$y^\ell=\frac{Y^\ell}{\sqrt{\delta}},\ Z^\ell=\frac{\delta}{2}\Big(\big(y_1^\ell+g_1^\ell\big)^2,\ldots,\big(y_n^\ell+g_n^\ell\big)^2\Big) \tag{19}$$

then h_τ is the density of $Z^1+Z^2+Z^3$. One can rewrite this density as a mixture over the distribution of (y^1,y^2,y^3) of the conditional density of $Z^1+Z^2+Z^3$ given (y^1,y^2,y^3). For a fixed (y^1,y^2,y^3), the coordinates of $Z^1+Z^2+Z^3$ become independent, so we first need to compute the density of the conditional distribution of one coordinate

$$\big(y_i^1+g_i^1\big)^2+\big(y_i^2+g_i^2\big)^2+\big(y_i^3+g_i^3\big)^2. \tag{20}$$

This distribution is an example of *noncentral chi-squared distribution*, in this case with 3 degrees of freedom. Let us denote

$$g_i=(g_i^1,g_i^2,g_i^3),y_i=(y_i^1,y_i^2,y_i^3)\ \text{ and }\ \lambda_i=\|y_i\|^2. \tag{21}$$

First, let us note that

$$\big(y_i^1+g_i^1\big)^2+\big(y_i^2+g_i^2\big)^2+\big(y_i^3+g_i^3\big)^2\stackrel{d}{=}\big(\sqrt{\lambda_i}+g_i^1\big)^2+(g_i^2)^2+(g_i^3)^2. \tag{22}$$

Indeed, if we take an orthogonal transformation Q such that $Qy_i=\|y_i\|e_1$ then we can write the left hand side as

$$\begin{aligned}\|y_i\|^2+2\langle y_i,g_i\rangle+\|g_i\|^2&=\|y_i\|^2+2\langle Qy_i,Qg_i\rangle+\|Qg_i\|^2\\&=\|y_i\|^2+2\|y_i\|\langle e_1,Qg_i\rangle+\|Qg_i\|^2\\&\stackrel{d}{=}\|y_i\|^2+2\|y_i\|\langle e_1,g_i\rangle+\|g_i\|^2\\&=\big(\sqrt{\lambda_i}+g_i^1\big)^2+(g_i^2)^2+(g_i^3)^2.\end{aligned}$$

This shows that the noncentral chi-squared distribution depends on the shift parameters in y_i only through $\lambda_i = \|y_i\|^2$, which is called the *noncentrality parameter*.

Let us compute the density of $(\sqrt{\lambda}+g)^2$ for $\lambda \geq 0$ and $g \sim \mathcal{N}(0,1)$. Since

$$\mathbb{P}\big((\sqrt{\lambda}+g)^2 \leq x\big) = \mathbb{P}\big(-\sqrt{x}-\sqrt{\lambda} \leq g \leq \sqrt{x}-\sqrt{\lambda}\big)$$

for $x > 0$, taking the derivative in x, the density on $(0,\infty)$ equals

$$\frac{1}{\sqrt{2\pi x}} e^{-(x+\lambda)/2} \cosh(\sqrt{\lambda x}) = \frac{1}{\sqrt{2\pi x}} e^{-(x+\lambda)/2} \sum_{k\geq 0} \frac{x^k \lambda^k}{(2k)!}.$$

It is a well-known property of the gamma function that

$$\Gamma(k+1/2) = \frac{2^{1-2k}\sqrt{\pi}\Gamma(2k)}{\Gamma(k)} = \frac{\sqrt{\pi}(2k)!}{2^{2k}k!},$$

so if we denote the density of $\Gamma(k+1/2,1/2)$ distribution by

$$\Gamma(k+1/2,1/2;x) = \frac{(1/2)^{k+1/2}}{\Gamma(k+1/2)} x^{k+1/2-1} e^{-x/2}$$

then the density of $(\sqrt{\lambda}+g)^2$ can be rewritten as

$$\sum_{k\geq 0} \frac{(\lambda/2)^k}{k!} e^{-\lambda/2} \Gamma(k+1/2,1/2;x). \tag{23}$$

Remark. It is interesting to observe that, since $\Gamma(k+1/2,1/2)$ is a χ^2_{1+2k}–distribution with $1+2k$ degrees of freedom, this representation for the density implies that

$$(\sqrt{\lambda}+g)^2 \stackrel{d}{=} g_1^2 + \ldots + g_{1+2M}^2,$$

where M has the Poisson distribution Poiss$(\lambda/2)$ with the mean $\lambda/2$. $\square$

From the above representation, since the distribution of $(g_i^2)^2 + (g_i^3)^2$ is $\Gamma(1,1/2)$, the density of the distribution of

$$\big(\sqrt{\lambda_i} + g_i^1\big)^2 + (g_i^2)^2 + (g_i^3)^2$$

conditionally on $\lambda_i = \|y_i\|^2$ is equal to

$$p(\lambda_i, x) := \sum_{k\geq 0} \frac{(\lambda_i/2)^k}{k!} e^{-\lambda_i/2} \Gamma(k+3/2,1/2;x). \tag{24}$$

This implies that the density of

$$Z^1 + Z^2 + Z^3 = \sum_{\ell=1}^{3} \frac{\delta}{2}\Big(\big(y_1^\ell + g_1^\ell\big)^2, \ldots, \big(y_n^\ell + g_n^\ell\big)^2\Big)$$

conditionally on (y^1, y^2, y^3) is equal to

$$\prod_{i\le n} \frac{2}{\delta} p\Big(\lambda_i, \frac{\delta x_i}{2}\Big)$$

and the unconditional density $h_\tau(x)$ equals

$$h_\tau(x) = \mathbb{E}\prod_{i\le n} \frac{2}{\delta} p\Big(\lambda_i, \frac{\delta x_i}{2}\Big), \tag{25}$$

where the expectation is with respect to y^1, y^2, y^3, which can be also written as

$$h_\tau(x) = \sum_{k_1,\ldots,k_n\ge 0} \mathbb{E}\prod_{i\le n} \frac{(\lambda_i/2)^{k_i}}{k_i!} e^{-\lambda_i/2}\Gamma(k_i + 3/2, 1/2; x_i). \tag{26}$$

If we introduce the notation

$$p_{k_1,\ldots,k_n} := \mathbb{E}\prod_{i\le n} \frac{(\lambda_i/2)^{k_i}}{k_i!} e^{-\lambda_i/2} \tag{27}$$

then, obviously, $\sum_{k_1,\ldots,k_n\ge 0} p_{k_1,\ldots,k_n} = 1$, so we can rewrite $h_\tau(x)$ as a mixture

$$h_\tau(x) = \sum_{k_1,\ldots,k_n\ge 0} p_{k_1,\ldots,k_n} \prod_{i\le n} \Gamma(k_i + 3/2, 1/2; x_i). \tag{28}$$

Using this representation, it is now easy to check the properties of $h_\tau(x)$ used in the proof above, because each factor

$$\Gamma(k_i + 3/2, 1/2; x_i) = \frac{(1/2)^{k_i+1/2}}{\Gamma(k_i+1/2)} x_i^{k_i+3/2-1} e^{-x_i/2}$$

is zero at $x_i = 0$ and

$$\frac{\partial}{\partial x_i}\Gamma(k_i + 3/2, 1/2; x_i) = \Gamma(k_i + 1/2, 1/2; x_i) - \frac{1}{2}\Gamma(k_i + 3/2, 1/2; x_i),$$

so the integrability and the fact that $\lim_{x_i\downarrow 0} \partial_J h_\tau(x) = 0$ for any $i \notin J$ should not be difficult to check.

References

[1] C. Borell, *A Gaussian correlation inequality for certain bodies in R^n*, Math. Ann. **256** (1981), 569–573.

[2] G. Hargé, *A particular case of correlation inequality for the Gaussian measure*, Ann. Probab. **27** (1999), 1939–1951.

[3] C. G. Khatri, *On certain inequalities for normal distributions and their applications to simultaneous confidence bounds*, Ann. Math. Statist. **38** (1967), 1853–1867.

[4] L. D. Pitt, *A Gaussian correlation inequality for symmetric convex sets*, Ann. Probab. **5** (1977), 470–474.

[5] T. Royen, *A simple proof of the Gaussian correlation conjecture extended to multivariate gamma distributions*, Far East J. Theor. Stat. **48** (2014), 139–145.

[6] G. Schechtman, T. Schlumprecht, J. Zinn, *On the Gaussian measure of the intersection*, Ann. Probab. **26** (1998), 346–357.

[7] Z. Šidák, *Rectangular confidence regions for the means of multivariate normal distributions*, J. Amer. Statist. Assoc. **62** (1967), 626–633.