A General Version of Price's Theorem

A Tool for Bounding the Expectation of Nonlinear Functions of Gaussian Random Vectors

Felix Voigtlaender[1,2] (ID)

Abstract

Assume that $X_\Sigma \in \mathbb{R}^n$ is a centered random vector following a multivariate normal distribution with positive definite covariance matrix Σ. Let $g : \mathbb{R}^n \to \mathbb{C}$ be measurable and of moderate growth, say $|g(x)| \lesssim (1 + |x|)^N$. We show that the map $\Sigma \mapsto \mathbb{E}[g(X_\Sigma)]$ is smooth, and we derive convenient expressions for its partial derivatives, in terms of certain expectations $\mathbb{E}[(\partial^\alpha g)(X_\Sigma)]$ of partial (distributional) derivatives of g. As we discuss, this result can be used to derive bounds for the expectation $\mathbb{E}[g(X_\Sigma)]$ of a nonlinear function $g(X_\Sigma)$ of a Gaussian random vector X_Σ with possibly correlated entries. For the case when $g(x) = g_1(x_1) \cdots g_n(x_n)$ has tensor-product structure, the above result is known in the engineering literature as *Price's theorem*, originally published in 1958. For dimension $n = 2$, it was generalized in 1964 by McMahon to the general case $g : \mathbb{R}^2 \to \mathbb{C}$. Our contribution is to unify these results, and to give a mathematically fully rigorous proof. Precisely, we consider a normally distributed random vector $X_\Sigma \in \mathbb{R}^n$ of arbitrary dimension $n \in \mathbb{N}$, and we allow the nonlinearity g to be a general tempered distribution. To this end, we replace the expectation $\mathbb{E}[g(X_\Sigma)]$ by the dual pairing $\langle g, \phi_\Sigma \rangle_{\mathcal{S}', \mathcal{S}}$, where ϕ_Σ denotes the probability density function of X_Σ.

✉ Felix Voigtlaender
felix@voigtlaender.xyz

1   Katholische Universität Eichstätt-Ingolstadt, Lehrstuhl Wissenschaftliches Rechnen, Ostenstraße 26, 85072 Eichstätt, Germany

2   Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, 10623 Berlin, Germany

1 Introduction

In this introduction, we first present a precise formulation of our version of Price's theorem, the proof of which we defer to Sect. 4. We then briefly discuss the relevance of this theorem: In a nutshell, it is a useful tool for estimating the expectation of a nonlinear function $g(X_\Sigma)$ of a Gaussian random vector $X_\Sigma \in \mathbb{R}^n$ with possibly correlated entries. In Sect. 3, we consider a specific example application which illustrates this. The relation of our result to the classical versions [6,8] of Price's theorem is discussed in Sect. 2.

1.1 Our Version of Price's Theorem

Let us denote by $\mathrm{Sym}_n := \left\{ A \in \mathbb{R}^{n \times n} \ : \ A^T = A \right\}$ the set of symmetric matrices, and by

$$\mathrm{Sym}_n^+ := \left\{ A \in \mathrm{Sym}_n \ : \ \forall x \in \mathbb{R}^n \setminus \{0\} : \langle x, Ax \rangle > 0 \right\}$$

the set of (symmetric) positive definite matrices, where we write $\langle x, y \rangle := x^T y$ for the standard scalar product of $x, y \in \mathbb{R}^n$ and $|x| := \sqrt{\langle x, x \rangle}$ for the usual Euclidean norm. For $\Sigma \in \mathrm{Sym}_n^+$, let

$$\phi_\Sigma : \mathbb{R}^n \to (0, \infty), x \mapsto \left[(2\pi)^n \cdot \det \Sigma \right]^{-\frac{1}{2}} \cdot e^{-\frac{1}{2} \langle x, \Sigma^{-1} x \rangle}, \qquad (1.1)$$

and note that ϕ_Σ is the density function of a centered random vector $X_\Sigma \in \mathbb{R}^n$ which follows a joint normal distribution with covariance matrix Σ—that is, $X_\Sigma \sim N(0, \Sigma)$; see for instance [5, Chapter 5, Theorem 5.1].

Let us briefly recall the notion of Schwartz functions and tempered distributions, which will play an important role in what follows. First, with $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$, any $\alpha \in \mathbb{N}_0^n$ will be called a *multiindex*, and we write $|\alpha| = \alpha_1 + \cdots + \alpha_n$ as well as $\partial^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$, and $z^\alpha = z_1^{\alpha_1} \cdots z_n^{\alpha_n}$ for $z \in \mathbb{C}^n$. Finally, given $\alpha, \beta \in \mathbb{N}_0^n$, we write $\beta \leq \alpha$ if $\beta_j \leq \alpha_j$ for all $j \in \{1, \dots, n\}$. With this notation, it is not hard to see that the density function ϕ_Σ from above belongs to the *Schwartz class*

$$\mathcal{S}(\mathbb{R}^n) = \Big\{ g \in C^\infty(\mathbb{R}^n; \mathbb{C}) \ : \ \forall \alpha \in \mathbb{N}_0^n \ \forall N \in \mathbb{N} \ \exists C > 0 \ \forall x \in \mathbb{R}^n \ :$$

$$|\partial^\alpha g(x)| \leq C \cdot (1 + |x|)^{-N} \Big\}$$

of smooth, rapidly decaying functions; see for instance [3, Chapter 8] for more details on this space. In fact, $\phi_\Sigma(x) = c_\Sigma \cdot e^{-\frac{1}{2} \langle \Sigma^{-1/2} x, \Sigma^{-1/2} x \rangle} = c_\Sigma \cdot \Phi(\Sigma^{-1/2} x)$, where Φ is the usual Gaussian function $\Phi(x) = e^{-\frac{1}{2} |x|^2}$, which is well-known to belong to $\mathcal{S}(\mathbb{R}^n)$.

The space $\mathcal{S}'(\mathbb{R}^n)$ of *tempered distributions* consists of all linear functionals $g : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ which are continuous with respect to the usual topology on $\mathcal{S}(\mathbb{R}^n)$; see [3, Sections 8.1 and 9.2] for the details. Since $\phi_\Sigma \in \mathcal{S}(\mathbb{R}^n)$, given any tempered

distribution $g \in \mathcal{S}'(\mathbb{R}^n)$, the function

$$\Phi_g : \mathrm{Sym}_n^+ \to \mathbb{C}, \Sigma \mapsto \langle g, \phi_\Sigma \rangle_{\mathcal{S}',\mathcal{S}} \tag{1.2}$$

is well-defined, where $\langle \cdot, \cdot \rangle_{\mathcal{S}',\mathcal{S}}$ denotes the (bilinear) dual pairing between $\mathcal{S}'(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$. As an important special case, note that if $g : \mathbb{R}^n \to \mathbb{C}$ is measurable and *of moderate growth*, in the sense that $x \mapsto (1 + |x|)^{-N} \cdot g(x) \in L^1(\mathbb{R}^n)$ for some $N \in \mathbb{N}$, then

$$\Phi_g(\Sigma) = \mathbb{E}\big[g(X_\Sigma)\big] \tag{1.3}$$

is just the expectation of $g(X_\Sigma)$, where $X_\Sigma \sim N(0, \Sigma)$. Here, we identify as usual the function g with the tempered distribution $\mathcal{S}(\mathbb{R}^n) \to \mathbb{C}, \varphi \mapsto \int g(x)\varphi(x)\,dx$.

The main goal of this note is to show for each $g \in \mathcal{S}'(\mathbb{R}^n)$ that the function Φ_g is smooth, and to derive an explicit formula for its partial derivatives. Thus, at least in the case of Equation (1.3), our goal is to *calculate the partial derivatives of the expectation of a nonlinear function g of a Gaussian random vector $X_\Sigma \sim N(0, \Sigma)$, as a function of the covariance matrix Σ of the vector X_Σ.*

In order to achieve a convenient statement of this result, we first introduce a bit more notation: Write $\underline{n} := \{1, \ldots, n\}$, and let

$$\begin{aligned}
I &:= \big\{(i, j) \in \underline{n} \times \underline{n} : i \le j\big\}, \quad I_\shortparallel := \big\{(i, i) : i \in \underline{n}\big\}, \\
I_< &:= \big\{(i, j) \in \underline{n} \times \underline{n} : i < j\big\},
\end{aligned} \tag{1.4}$$

so that $I = I_\shortparallel \uplus I_<$. Since for $n > 1$, the sets Sym_n and Sym_n^+ have empty interior in $\mathbb{R}^{n \times n}$ (because they only consist of symmetric matrices), it does not make sense to talk about partial derivatives of a function $\Phi : \mathrm{Sym}_n^+ \to \mathbb{C}$, unless one interprets Sym_n^+ as an open subset of the vector space Sym_n, rather than of $\mathbb{R}^{n \times n}$. As a means of fixing a coordinate system on Sym_n, we therefore parameterize the set of symmetric matrices by their "upper half"; precisely, we consider the following isomorphism between $\mathbb{R}^I$ and Sym_n:

$$\Omega : \mathbb{R}^I \to \mathrm{Sym}_n, \ \big(A_{i,j}\big)_{1 \le i \le j \le n} \mapsto \sum_{i \le j} A_{i,j} E_{i,j} + \sum_{i > j} A_{j,i} E_{i,j}. \tag{1.5}$$

Here, we denote by $(E_{i,j})_{i,j \in \underline{n}}$ the standard basis of $\mathbb{R}^{n \times n}$, meaning that $(E_{i,j})_{k,\ell} = \delta_{i,k} \cdot \delta_{j,\ell}$ with the usual Dirac delta $\delta_{i,k}$. Below, instead of calculating the partial derivatives of Φ_g, we will consider the function $\Phi_g \circ \Omega|_U$, where $U := \Omega^{-1}\big(\mathrm{Sym}_n^+\big) \subset \mathbb{R}^I$ is open.

In order to achieve a concise formulation of our version of Price's theorem, we need two non-standard notions regarding multiindices $\beta = \big(\beta(i, j)\big)_{(i,j) \in I} \in \mathbb{N}_0^I$. Namely, we define the **flattened version** of β as

$$\beta_\flat := \sum_{(i,j) \in I} \beta(i, j)\,(e_i + e_j) \in \mathbb{N}_0^n \quad \text{with the standard basis } (e_1, \ldots, e_n) \text{ of } \mathbb{R}^n,$$

$$\tag{1.6}$$

and in addition to $|\beta| = \sum_{(i,j) \in I} \beta(i,j)$, we will also use

$$|\beta|_{\shortparallel} := \sum_{(i,j) \in I_{\shortparallel}} \beta(i,j) = \sum_{i \in \underline{n}} \beta(i,i). \qquad (1.7)$$

With this notation, our main result reads as follows:

Theorem 1 (Generalized version of Price's theorem) *Let $g \in \mathcal{S}'(\mathbb{R}^n)$ be arbitrary. Then the function $\Phi_g \circ \Omega|_U : U \to \mathbb{C}$ is smooth and its partial derivatives are given by*

$$\partial^\beta \left(\Phi_g \circ \Omega \right)(A) = (1/2)^{|\beta|_{\shortparallel}} \cdot \langle \partial^{\beta_\flat} g \,, \phi_{\Omega(A)} \rangle_{\mathcal{S}',\mathcal{S}} \qquad \forall A \in U = \Omega^{-1}(\mathrm{Sym}_n^+) \forall \beta \in \mathbb{N}_0^I. \qquad (1.8)$$

Here $\partial^{\beta_\flat} g$ denotes the usual distributional derivative of g.

Remark 1 Note that even if one is in the setting of Equation (1.3) where $g : \mathbb{R}^n \to \mathbb{C}$ is of moderate growth, so that $\Phi_g(\Sigma) = \mathbb{E}[g(X_\Sigma)]$ is a "classical" expectation, it need *not* be the case that the derivative $\partial^{\beta_\flat} g$ is given by a *function*, let alone one of moderate growth. Therefore, it really is useful to consider the formalism of (tempered) distributions.

1.2 Relevance of Price's Theorem

An important application of Price's theorem is as follows: For certain values of the covariance matrix Σ, it is usually easy to precisely calculate the expectation $\mathbb{E}[g(X_\Sigma)]$—for example if Σ is a diagonal matrix, in which case the entries of X_Σ are independent. As a complement to such special cases where explicit calculations are possible, Price's theorem can be used to obtain (bounds for) the partial derivatives of the map $\Sigma \mapsto \mathbb{E}[g(X_\Sigma)]$. In combination with standard results from multivariable calculus, one can then obtain bounds for $\mathbb{E}[g(X_\Sigma)]$ for general covariance matrices Σ. Thus, *Price's theorem is a tool for estimating the expectation of a nonlinear function $g(X_\Sigma)$ of a Gaussian random vector X_Σ, even if the entries of X_Σ are correlated.*

An example for this type of reasoning will be given in Sect. 3. There, we apply our version of Price's theorem to show that if $f(x) = f_\tau(x)$ "clips" $x \in \mathbb{R}$ to the interval $[-\tau, \tau]$ and if $(X_\alpha, Y_\alpha) \sim N(0, \Sigma_\alpha)$ for $\Sigma_\alpha = \left(\begin{smallmatrix} 1 & \alpha \\ \alpha & 1 \end{smallmatrix} \right)$, then the map $F_\tau : [0,1] \to \mathbb{R}, \alpha \mapsto \mathbb{E}[f_\tau(X_\alpha) f_\tau(Y_\alpha)]$ is *convex* and satisfies $F_\tau(0) = 0$. Thus, $F_\tau(\alpha) \leq \alpha \cdot F_\tau(1)$, where $F_\tau(1)$ is easy to bound since $X_1 = Y_1$ almost surely. These facts constitute important ingredients in [4]; see Theorem A.4 and the proof of Lemma A.3 in that paper.

2 Comparison with the Classical Results

The original form of Price's theorem as stated in [8] only concerns the case when the nonlinearity $g(x) = g_1(x_1) \cdots g_n(x_n)$ has a tensor-product structure. In this special

case, the formula derived in [8] is identical to the one given by Theorem 1, up to notational differences.

This tensor-product structure assumption concerning g was removed by McMahon [6] and Papoulis [7] in the case of Gaussian random vectors of dimension $n = 2$ with covariance matrix of the form $\Sigma = \Sigma_\alpha = \left(\begin{smallmatrix} 1 & \alpha \\ \alpha & 1 \end{smallmatrix} \right)$ with $\alpha \in (-1, 1)$. Precisely, if $X_\alpha \sim N(0, \Sigma_\alpha)$, then [6] states for $g : \mathbb{R}^2 \to \mathbb{C}$ that

$$\Theta_g : (-1, 1) \to \mathbb{C}, \alpha \mapsto \mathbb{E}\left[g(X_\alpha)\right] \quad \text{is smooth with} \quad \Theta_g^{(n)}(\alpha) = \mathbb{E}\left[\frac{\partial^{2n} g}{\partial x_1^n \partial x_2^n}(X_\alpha)\right].$$
(2.1)

Based on the work by Papoulis, Brown [1] showed that Price's theorem holds for Gaussian random vectors X of general dimensionality and unit variance $\Sigma_{i,i} = \mathbb{E}[(X_\Sigma)_i^2] = 1$, if one takes derivatives with respect to the covariances $\Sigma_{i,j} = \mathbb{E}[(X_\Sigma)_i (X_\Sigma)_j]$ where $i \neq j$. In this setting, Brown also showed that Price's theorem *characterizes the normal distribution*; more precisely, if $(X_\Sigma)_\Sigma$ is a (sufficiently nice) family of random vectors with $\mathrm{Cov}(X_\Sigma) = \Sigma$ which satisfies the conclusion of Price's theorem, then $X_\Sigma \sim N(0, \Sigma)$ is necessarily normally distributed. This extends and corrects the original work of Price [8], where a similar claim was made.

Finally, we mention the article [9] in which a quantum-mechanical version of Price's theorem is established. In Sect. 2 of that paper, the author reviews the "classical" case of Price's theorem, and essentially derives the same formulas as in Theorem 1.

Despite their great utility, the existing versions of Price's theorem have some shortcomings—at least from a mathematical perspective:

- In [1,6,8], the assumptions regarding the functions $g_1, \ldots, g_n$ or g are never made explicit. In particular, it is assumed in [6,8] without justification that $g_1, \ldots, g_n$ or g can be represented as the sum of certain Laplace transforms. Likewise, Papoulis [7] assumes that g satisfies the decay condition $|g(x, y)| \lesssim e^{|(x,y)|^\beta}$ for some $\beta < 2$, but does not impose *any* restrictions on the regularity of g. Finally, [1] is mainly concerned with showing that Price's theorem *only* holds for normally distributed random vectors, and simply refers to [7] for the proof that Price's theorem does indeed hold for normal random vectors.

  None of the papers [1,6–8], explains the nature of the derivative of g (classical, distributional, etc.) which appears in the derived formula.
- In contrast, for calculating the k-th order derivatives of $\Sigma \mapsto \mathbb{E}[g(X_\Sigma)]$, it is assumed in [9] that the nonlinearity g is C^{2k}, with a certain decay condition concerning the derivatives. This *classical* smoothness of g, however, does not hold in many applications; see Sect. 3.

Differently from [1,6–9], our version of Price's theorem imposes precise, rather mild assumptions concerning the nonlinearity g (namely $g \in \mathcal{S}'(\mathbb{R}^n)$) and precisely explains the nature of the derivative $\partial^{\beta_\flat} g$ that appears in the theorem statement: this is just a distributional derivative.

Furthermore, maybe as a consequence of the preceding points, it seems that Price's theorem is not as well-known in the mathematical community as it deserves to be. *It is my hope that the present paper may promote this result.*

Before closing this section, we prove that—assuming g to be a tempered distribution—the result of [6,7] is indeed a special case of Theorem 1. With similar arguments, one can show that the forms of Price's theorem considered in [1,8,9] are covered by Theorem 1 as well.

Corollary 1 *Let* $g \in \mathcal{S}'(\mathbb{R}^2)$. *For* $\alpha \in (-1, 1)$, *let* $\Sigma_\alpha := \left(\begin{smallmatrix} 1 & \alpha \\ \alpha & 1 \end{smallmatrix} \right)$. *Let*

$$\Theta_g : (-1, 1) \to \mathbb{C}, \quad \alpha \mapsto \langle g, \phi_{\Sigma_\alpha} \rangle_{\mathcal{S}',\mathcal{S}},$$

where $\phi_{\Sigma_\alpha} : \mathbb{R}^2 \to (0, \infty)$ *denotes the probability density function of* $X_\alpha \sim N(0, \Sigma_\alpha)$.

Then Θ_g *is smooth with n-th derivative* $\Theta_g^{(n)}(\alpha) = \left\langle \frac{\partial^{2n} g}{\partial x_1^n \partial x_2^n}, \phi_{\Sigma_\alpha} \right\rangle_{\mathcal{S}',\mathcal{S}}$ *for* $\alpha \in (-1, 1)$.

Remark 2 In particular, if both g and the (distributional) derivative $\frac{\partial^{2n} g}{\partial x_1^n \partial x_2^n}$ are given by functions of moderate growth, then Equation (2.1) holds, i.e.,

$$\frac{d^n}{d\alpha^n} \mathbb{E}\big[g(X_\alpha) \big] = \mathbb{E}\left[\frac{\partial^{2n} g}{\partial x_1^n \partial x_2^n} (X_\alpha) \right].$$

Proof of Corollary 1 In the notation of Theorem 1, we have

$$\Theta_g(\alpha) = \big(\Phi_g \circ \Omega \big) \big(A^{(\alpha)} \big) \quad \text{with} \quad A^{(\alpha)}_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ \alpha, & \text{if } i \neq j \end{cases}$$

$$\text{for} \quad (i, j) \in I = \{(1, 1), (1, 2), (2, 2)\}.$$

Since $\Omega\big(A^{(\alpha)}\big) = \Sigma_\alpha$ is easily seen to be positive definite, we have $A^{(\alpha)} \in U$. Now, setting $\beta := n \cdot e_{(1,2)} \in \mathbb{N}_0^I$ (with the standard basis $e_{(1,1)}, e_{(1,2)}, e_{(2,2)}$ of $\mathbb{R}^I$), the flattened version $\beta_\flat$ of β satisfies $\beta_\flat = ne_1 + ne_2 = (n, n)$. Thus, Theorem 1 and the chain-rule show that Θ_g is smooth, with

$$\Theta_g^{(n)}(\alpha) = \frac{d^n}{d\alpha^n} \left[(\Phi_g \circ \Omega)\big(A^{(\alpha)}\big) \right] = \big[\partial^\beta (\Phi_g \circ \Omega) \big] \big(A^{(\alpha)} \big)$$

$$= \langle \partial^{\beta_\flat} g, \phi_{\Omega(A^{(\alpha)})} \rangle_{\mathcal{S}',\mathcal{S}} = \left\langle \frac{\partial^{2n} g}{\partial x_1^n \partial x_2^n}, \phi_{\Sigma_\alpha} \right\rangle_{\mathcal{S}',\mathcal{S}}.$$

$\square$

3 An Example of an Application of Price's Theorem

In this section, we derive bounds for the expectation $\mathbb{E}[f_\tau(X_\alpha) f_\tau(Y_\alpha)]$, where X_α, Y_α follow a joint normal distribution with covariance matrix $\left(\begin{smallmatrix} 1 & \alpha \\ \alpha & 1 \end{smallmatrix} \right)$, and where the non-linearity f_τ is just a truncation (or *clipping*) to the interval $[-\tau, \tau]$. We remark that this example has already been considered by Price [8] himself, but that his arguments are

not completely mathematically rigorous, as explained in Sect. 2. Precisely, we obtain the following result:

Lemma 1 *Let $\tau > 0$ be arbitrary, and define*

$$f_\tau : \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} \tau, & \text{if } x \geq \tau, \\ x, & \text{if } x \in [-\tau, \tau], \\ -\tau, & \text{if } x \leq -\tau. \end{cases}$$

For $\alpha \in [-1, 1]$, set $\Sigma_\alpha := \left(\begin{smallmatrix} 1 & \alpha \\ \alpha & 1 \end{smallmatrix} \right)$, and let $(X_\alpha, Y_\alpha) \sim N(0, \Sigma_\alpha)$. Finally, define

$$F_\tau : [-1, 1] \to \mathbb{R}^2, \qquad \alpha \mapsto \mathbb{E}\big[f_\tau(X_\alpha) \cdot f_\tau(Y_\alpha) \big].$$

Then F_τ is continuous and $F_\tau|_{[0,1]}$ is convex with $F_\tau(0) = 0$. In particular, $F_\tau(\alpha) \leq \alpha \cdot F_\tau(1)$ for all $\alpha \in [0, 1]$.

Proof It is easy to see that f_τ is bounded and Lipschitz continuous, so that $f_\tau \in W^{1,\infty}(\mathbb{R})$ with weak derivative $f_\tau' = \mathbb{1}_{(-\tau,\tau)}$. Therefore, using the notation $(g \otimes h)(x, y) = g(x) h(y)$, we see that $g_\tau := f_\tau \otimes f_\tau \in W^{1,\infty}(\mathbb{R}^2) \subset \mathcal{S}'(\mathbb{R}^2)$, with weak derivative $\frac{\partial^2 g_\tau}{\partial x_1 \partial x_2} = \mathbb{1}_{(-\tau,\tau)} \otimes \mathbb{1}_{(-\tau,\tau)} = \mathbb{1}_{(-\tau,\tau)^2}$. Directly from the definition of the weak derivative, in combination with Fubini's theorem and the fundamental theorem of calculus, we thus see for each $\phi \in \mathcal{S}(\mathbb{R}^2)$ that

$$\left\langle \frac{\partial^4 g_\tau}{\partial x_1^2 \partial x_2^2}, \phi \right\rangle_{\mathcal{S}',\mathcal{S}} = \left\langle \frac{\partial^2 g_\tau}{\partial x_1 \partial x_2}, \frac{\partial^2 \phi}{\partial x_1 \partial x_2} \right\rangle_{\mathcal{S}',\mathcal{S}} = \int_{-\tau}^{\tau} \int_{-\tau}^{\tau} \left(\frac{\partial^2 \phi}{\partial x_1 \partial x_2} \right)(t_1, t_2) \, dt_1 \, dt_2$$

$$= \int_{-\tau}^{\tau} \left(\frac{\partial \phi}{\partial x_2} \right)(\tau, t_2) - \left(\frac{\partial \phi}{\partial x_2} \right)(-\tau, t_2) \, dt_2$$

$$= \phi(\tau, \tau) - \phi(-\tau, \tau) - \phi(\tau, -\tau) + \phi(-\tau, -\tau).$$

Now, Corollary 1 shows that $F_\tau|_{(-1,1)} = \Theta_{g_\tau}$ is smooth with

$$F_\tau''(\alpha) = \left\langle \frac{\partial^4 g_\tau}{\partial x_1^2 \partial x_2^2}, \phi_{\Sigma_\alpha} \right\rangle_{\mathcal{S}',\mathcal{S}}$$
$$= \phi_{\Sigma_\alpha}(\tau, \tau) - \phi_{\Sigma_\alpha}(-\tau, \tau) - \phi_{\Sigma_\alpha}(\tau, -\tau) + \phi_{\Sigma_\alpha}(-\tau, -\tau)$$

for $\alpha \in (-1, 1)$. We want to show $F_\tau''(\alpha) \geq 0$ for $\alpha \in [0, 1)$. Since ϕ_{Σ_α} is symmetric, it suffices to show $\phi_{\Sigma_\alpha}(\tau, \tau) - \phi_{\Sigma_\alpha}(-\tau, \tau) \geq 0$, which is easily seen to be equivalent to

$$\exp\left(-\frac{1}{2(1 - \alpha^2)}(2\tau^2 - 2\alpha\tau^2) \right) \overset{!}{\geq} \exp\left(-\frac{1}{2(1 - \alpha^2)}(2\tau^2 + 2\alpha\tau^2) \right)$$

$$\iff 2\tau^2 + 2\alpha\tau^2 \overset{!}{\geq} 2\tau^2 - 2\alpha\tau^2 \iff 4\alpha\tau^2 \overset{!}{\geq} 0,$$

which clearly holds for $\alpha \in [0, 1)$.

To finish the proof, we only need to show that F_τ is continuous with $F_\tau(0) = 0$. To see this, let $(X, Z) \sim N(0, I_2)$, with the 2-dimensional identity matrix I_2. For $\alpha \in [-1, 1]$, it is then not hard to see that $Y_\alpha := \alpha X + \sqrt{1 - \alpha^2} Z$ satisfies $(X, Y_\alpha) \sim N(0, \Sigma_\alpha)$. Therefore, we see for $\alpha, \beta \in [-1, 1]$ that

$$
\begin{aligned}
|F_\tau(\alpha) - F_\tau(\beta)| &= \left| \mathbb{E}\left[g_\tau(X, Y_\alpha) \right] - \mathbb{E}\left[g_\tau(X, Y_\beta) \right] \right| \\
&= \left| \mathbb{E}\left[f_\tau(X) \cdot \left(f_\tau(Y_\alpha) - f_\tau(Y_\beta) \right) \right] \right| \\
(\text{since } |f_\tau(X)| \leq \tau) \quad &\leq \tau \cdot \mathbb{E}\left| f_\tau(Y_\alpha) - f_\tau(Y_\beta) \right| \\
(\text{since } f_\tau \text{ is 1-Lipschitz}) \quad &\leq \tau \cdot \mathbb{E}\left| Y_\alpha - Y_\beta \right| \\
&\leq \tau \cdot |\alpha - \beta| \cdot \mathbb{E}|X| + \tau \cdot \left| \sqrt{1 - \alpha^2} - \sqrt{1 - \beta^2} \right| \cdot \mathbb{E}|Z| \xrightarrow[\beta \to \alpha]{} 0,
\end{aligned}
$$

which shows that F_τ is indeed continuous. Furthermore, we see by independence of X, Z that

$$
F_\tau(0) = \mathbb{E}[f_\tau(X) \cdot f_\tau(Z)] = \mathbb{E}[f_\tau(X)] \cdot \mathbb{E}[f_\tau(Z)] = 0,
$$

since $\mathbb{E}[f_\tau(X)] = -\mathbb{E}[f_\tau(-X)] = -\mathbb{E}[f_\tau(X)]$, because of $X \sim -X$ and $f_\tau(x) = -f_\tau(-x)$ for $x \in \mathbb{R}$. □

4 The Proof of Theorem 1

The main idea of the proof is to use Fourier analysis, since the Fourier transform $\mathcal{F}\phi_\Sigma$ of the density function ϕ_Σ will turn out to be much easier to handle than ϕ_Σ itself. This is similar to the approach in [1,7] but slightly different from the approach in [6,8], where the Laplace transform is used instead.

For the Fourier transform, we will use the normalization

$$
\mathcal{F}\varphi(\xi) := \widehat{\varphi}(\xi) := \int_{\mathbb{R}^n} \varphi(x) \cdot e^{-i\langle x, \xi \rangle} \, dx \quad \text{for } \xi \in \mathbb{R}^n \text{ and } \varphi \in L^1(\mathbb{R}^n).
$$

It is well-known that the restriction $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ of $\mathcal{F}$ is a well-defined homeomorphism, with inverse $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, where $\mathcal{F}^{-1}\varphi(x) = (2\pi)^{-n} \cdot \mathcal{F}\varphi(-x)$. By duality, the Fourier transform also extends to a bijection $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ defined [1] by $\langle \mathcal{F}g, \varphi \rangle_{\mathcal{S}',\mathcal{S}} := \langle g, \mathcal{F}\varphi \rangle_{\mathcal{S}',\mathcal{S}}$ for $g \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Further, it is well-known for the distributional derivatives $\partial^\alpha g$ of $g \in \mathcal{S}'(\mathbb{R}^n)$ defined by $\langle \partial^\alpha g, \varphi \rangle_{\mathcal{S}',\mathcal{S}} = (-1)^{|\alpha|} \cdot \langle g, \partial^\alpha \varphi \rangle_{\mathcal{S}',\mathcal{S}}$ that if we set

$$
X^\alpha \cdot \varphi : \mathbb{R}^n \to \mathbb{C}, x \mapsto x^\alpha \cdot \varphi(x) \quad \text{and} \quad \langle X^\alpha \cdot g, \varphi \rangle_{\mathcal{S}',\mathcal{S}} = \langle g, X^\alpha \cdot \varphi \rangle_{\mathcal{S}',\mathcal{S}}
$$

---

[1] This definition is motivated by the identity $\int \widehat{f}(x) \cdot g(x)dx = \int \int f(\xi) e^{-i\langle x, \xi \rangle} g(x) d\xi dx = \int f(\xi) \widehat{g}(\xi) d\xi$ which is valid for $f, g \in L^1(\mathbb{R}^n)$ thanks to Fubini's theorem.

for $g \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then we have

$$\mathcal{F}\left[\partial^\alpha g\right] = i^{|\alpha|} \cdot X^\alpha \cdot \mathcal{F}g \quad \forall g \in \mathcal{S}'(\mathbb{R}^n), \ \alpha \in \mathbb{N}_0^n. \tag{4.1}$$

These results can be found e.g., in [2, Chapter 14], or (with a slightly different normalization of the Fourier transform) in [3, Sections 8.3 and 9.2].

Finally, we will use the formula

$$(2\pi)^n \cdot \mathcal{F}^{-1}\phi_\Sigma(\xi) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \phi_\Sigma(x)\,dx = \mathbb{E}\left[e^{i\langle \xi, X_\Sigma\rangle}\right]$$
$$= e^{-\frac{1}{2}\langle \xi, \Sigma\xi\rangle} =: \psi_\Sigma(\xi) \ \text{ for } \xi \in \mathbb{R}^n, \tag{4.2}$$

which is proved in [5, Chapter 5, Theorem 4.1]; in probabilistic terms, this is a statement about the characteristic function of the random vector $X_\Sigma \sim N(0, \Sigma)$.

Next, by the assumption of Theorem 1, we have $g \in \mathcal{S}'(\mathbb{R}^n)$ and hence $\mathcal{F}g \in \mathcal{S}'(\mathbb{R}^n)$. Thus, by the structure theorem for tempered distributions (see for instance [2, Theorem 17.10]), there are $L \in \mathbb{N}$, certain $\alpha_1, \ldots, \alpha_L \in \mathbb{N}_0^n$ and certain polynomially bounded, continuous functions $f_1, \ldots, f_L : \mathbb{R}^n \to \mathbb{C}$ satisfying $\mathcal{F}g = \sum_{\ell=1}^L \partial^{\alpha_\ell} f_\ell$, i.e., $g = \sum_{\ell=1}^L \mathcal{F}^{-1}(\partial^{\alpha_\ell} f_\ell)$. Since both sides of the target identity (1.8) are linear with respect to g, we can thus assume without loss of generality that $g = \mathcal{F}^{-1}(\partial^\alpha f)$ for some $\alpha \in \mathbb{N}_0^n$ and some continuous $f : \mathbb{R}^n \to \mathbb{C}$ which is polynomially bounded, say $|f(\xi)| \leq C \cdot (1 + |\xi|)^N$ for all $\xi \in \mathbb{R}^n$ and certain $C > 0$, $N \in \mathbb{N}_0$. We thus have

$$\Phi_g(\Sigma) = \langle g, \phi_\Sigma\rangle_{\mathcal{S}',\mathcal{S}} = \left\langle g, \mathcal{F}\mathcal{F}^{-1}\phi_\Sigma\right\rangle_{\mathcal{S}',\mathcal{S}} = \left\langle \mathcal{F}g, \mathcal{F}^{-1}\phi_\Sigma\right\rangle_{\mathcal{S}',\mathcal{S}}$$
$$\text{(Eq. (4.2))} = (2\pi)^{-n} \cdot \left\langle \partial^\alpha f, \psi_\Sigma\right\rangle_{\mathcal{S}',\mathcal{S}}$$
$$= (-1)^{|\alpha|} \cdot (2\pi)^{-n} \cdot \left\langle f, \partial^\alpha \psi_\Sigma\right\rangle_{\mathcal{S}',\mathcal{S}}$$
$$= (-1)^{|\alpha|} \cdot (2\pi)^{-n} \cdot \int_{\mathbb{R}^n} f(\xi) \cdot \left(\partial^\alpha \psi_\Sigma\right)(\xi)\,d\xi \quad \text{for all } \Sigma \in \text{Sym}_n^+. \tag{4.3}$$

Our first goal in the remainder of the proof is to show that one can justify "differentiation under the integral" with respect to $A_{i,j}$ with $\Sigma = \Omega(A)$ in the last integral in Equation (4.3).

It is easy to see that $A \mapsto \psi_{\Omega(A)}(\xi)$ is smooth, with partial derivative

$$\partial_{A_{i,j}}\psi_{\Omega(A)}(\xi) = e^{-\frac{1}{2}\langle \xi, \Omega(A)\xi\rangle} \cdot \partial_{A_{i,j}}\left(-\frac{1}{2} \cdot \sum_{k,\ell=1}^n \left[\Omega(A)\right]_{k,\ell} \cdot \xi_k \xi_\ell\right)$$
$$= \begin{cases} -\frac{1}{2} \cdot \xi_i\,\xi_j \cdot \psi_{\Omega(A)}(\xi), & \text{if } i = j, \\ -\xi_i\,\xi_j \cdot \psi_{\Omega(A)}(\xi), & \text{if } i < j \end{cases}$$

for all $\xi \in \mathbb{R}^n$ and arbitrary $(i, j) \in I$ and $A \in U$. Given $\beta \in \mathbb{N}_0^I$, let us write ∂_A^β for the partial derivative of order β with respect to $A \in \mathbb{R}^I$. Then, a straightforward

induction using the preceding identity shows (with $|\beta|_{11}$ and β_b as in (1.7) and (1.6)) that

$$\partial_A^\beta \psi_{\Omega(A)}(\xi) = (-1)^{|\beta|} \cdot \left(\frac{1}{2}\right)^{|\beta|_{11}} \cdot \xi^{\beta_b} \cdot \psi_{\Omega(A)}(\xi) \qquad \forall \, \beta \in \mathbb{N}_0^I, \, \xi \in \mathbb{R}^n, \, A \in U.$$
(4.4)

Next, we show for arbitrary $\gamma \in \mathbb{N}_0^n$ that there is a polynomial $p_{\alpha,\gamma} = p_{\alpha,\gamma}(\Xi, B)$ in the variables $\Xi \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n \times n}$ that satisfies

$$\partial_\xi^\alpha \big[\xi^\gamma \cdot \psi_\Sigma(\xi)\big] = p_{\alpha,\gamma}(\xi, \Sigma) \cdot \psi_\Sigma(\xi) \qquad \forall \, \Sigma \in \mathrm{Sym}_n^+, \, \xi \in \mathbb{R}^n.$$
(4.5)

To see this, we first note that a direct computation using the identity $\partial_{\xi_i}(\xi_k \xi_\ell) = \delta_{i,k} \xi_\ell + \delta_{i,\ell} \xi_k$ and the symmetry of Σ shows that $\partial_{\xi_i} \psi_\Sigma(\xi) = -(\Sigma \xi)_i \cdot \psi_\Sigma(\xi)$. By induction, and since $(\Sigma \xi)_i$ is a polynomial in ξ, Σ, we therefore see that for each $\beta \in \mathbb{N}_0^n$ there is a polynomial $p_\beta = p_\beta(\Xi, B)$ in the variables $\Xi \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n \times n}$ satisfying $\partial_\xi^\beta \psi_\Sigma(\xi) = \psi_\Sigma(\xi) \cdot p_\beta(\xi, \Sigma)$. Therefore, the Leibniz rule shows

$$\partial_\xi^\alpha \big[\xi^\gamma \psi_\Sigma(\xi)\big] = \sum_{\beta \in \mathbb{N}_0^n \text{ with } \beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta \psi_\Sigma(\xi) \cdot \partial^{\alpha-\beta} \xi^\gamma$$

$$= \psi_\Sigma(\xi) \sum_{\beta \in \mathbb{N}_0^n \text{ with } \beta \le \alpha} \binom{\alpha}{\beta} p_\beta(\xi, \Sigma) \, \partial^{\alpha-\beta} \xi^\gamma,$$

which proves Equation (4.5).

Now we are ready to justify differentiation under the integral (as in [3, Theorem 2.27]) for the last integral appearing in Equation (4.3), with $\Sigma = \Omega(A)$, that is, for the function

$$U \to \mathbb{C}, \qquad A \mapsto \int_{\mathbb{R}^n} f(\xi) \cdot \big(\partial^\alpha \psi_{\Omega(A)}\big)(\xi) \, d\xi.$$

Indeed, let $A_0 \in U$ be arbitrary. Since U is open, there is some $\varepsilon > 0$ satisfying $\overline{B_\varepsilon}(A_0) \subset U$, for the *closed* ball $\overline{B_\varepsilon}(A_0) = \big\{A \in \mathbb{R}^I : |A - A_0| \le \varepsilon\big\}$, with the Euclidean norm $|\cdot|$ on $\mathbb{R}^I$. The *open* ball $B_\varepsilon(A_0)$ is defined similarly.

Now, with

$$\sigma_{\min}(A) := \inf_{\substack{x \in \mathbb{R}^n \\ |x|=1}} \langle x, Ax \rangle \qquad \text{for } A \in \mathbb{R}^{n \times n}$$

we have for $A, B \in \mathbb{R}^{n \times n}$ and arbitrary $x \in \mathbb{R}^n$ with $|x| = 1$ that

$$\sigma_{\min}(A) \le \langle x, Ax \rangle = \langle x, Bx \rangle + \langle x, (A - B)x \rangle \le \langle x, Bx \rangle + \|A - B\|.$$

Since this holds for all $|x| = 1$, we get $\sigma_{\min}(A) \le \sigma_{\min}(B) + \|A - B\|$, and by symmetry $|\sigma_{\min}(A) - \sigma_{\min}(B)| \le \|A - B\|$. Therefore, the continuous function

$A \mapsto \sigma_{\min}(\Omega(A))$ has a positive(!) minimum on the compact set $\overline{B_\varepsilon}(A_0)$, so that $\langle \xi, \Omega(A)\xi \rangle \geq c \cdot |\xi|^2$ for all $\xi \in \mathbb{R}^n$ and $A \in \overline{B_\varepsilon}(A_0)$, for a positive $c > 0$. Furthermore, there is some $K = K(A_0) > 0$ with $\|\Omega(A)\| \leq K$ for all $A \in \overline{B_\varepsilon}(A_0)$.

Now, since the map $U \times \mathbb{R}^n \ni (A, \xi) \mapsto \psi_{\Omega(A)}(\xi) \in \mathbb{C}$ is smooth, we have (in view of Equations (4.4) and (4.5)) for arbitrary $\beta \in \mathbb{N}_0^I$, $A \in U$ and $\xi \in \mathbb{R}^n$ that

$$
\begin{aligned}
\partial_A^\beta \left[f(\xi) \cdot \left(\partial^\alpha \psi_{\Omega(A)} \right)(\xi) \right] &= f(\xi) \cdot \partial_\xi^\alpha \left[\partial_A^\beta \psi_{\Omega(A)}(\xi) \right] \\
\text{(Eq. (4.4))} &= f(\xi) \cdot (-1)^{|\beta|} \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot \partial_\xi^\alpha \left[\xi^{\beta_\flat} \cdot \psi_{\Omega(A)}(\xi) \right] \\
\text{(Eq. (4.5))} &= f(\xi) \cdot (-1)^{|\beta|} \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot p_{\alpha, \beta_\flat}(\xi, \Omega(A)) \cdot \psi_{\Omega(A)}(\xi).
\end{aligned}
\tag{4.6}
$$

Using the polynomial growth restriction concerning f, we thus see that there is a constant $C_{\alpha,\beta} > 0$ and some $M_{\alpha,\beta} \in \mathbb{N}$ with

$$
\begin{aligned}
\left| \partial_A^\beta \left[f(\xi) \cdot \left(\partial^\alpha \psi_{\Omega(A)} \right)(\xi) \right] \right| &= \left| f(\xi) \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot p_{\alpha, \beta_\flat}(\xi, \Omega(A)) \cdot \psi_{\Omega(A)}(\xi) \right| \\
&\leq C \cdot (1 + |\xi|)^N \cdot C_{\alpha,\beta} \cdot (1 + |\xi| + \|\Omega(A)\|)^{M_{\alpha,\beta}} \cdot \\
&\quad e^{-\frac{1}{2}\langle \xi, \Omega(A)\xi \rangle} \\
&\leq C_{\alpha,\beta} C \cdot (1 + |\xi|)^N \cdot (1 + |\xi| + K)^{M_{\alpha,\beta}} \cdot e^{-\frac{c}{2}|\xi|^2} \\
&=: h_{\alpha,\beta,A_0,f}(\xi),
\end{aligned}
$$

for all $\xi \in \mathbb{R}^n$ and all $A \in B_\varepsilon(A_0)$. Since $h_{\alpha,\beta,A_0,f}$ is independent of $A \in B_\varepsilon(A_0)$ and since we clearly have $h_{\alpha,\beta,A_0,f} \in L^1(\mathbb{R}^n)$, [3, Theorem 2.27] and Equation (4.3) show that the function

$$
B_\varepsilon(A_0) \to \mathbb{C}, \quad A \mapsto (-1)^{|\alpha|} \cdot (2\pi)^n \cdot \Phi_g(\Omega(A)) = \int_{\mathbb{R}^n} f(\xi) \cdot \left(\partial^\alpha \psi_{\Omega(A)} \right)(\xi) \, d\xi
$$

is smooth, with partial derivative of order $\beta \in \mathbb{N}_0^I$ given by

$$
\begin{aligned}
\partial_A^\beta \left[(-1)^{|\alpha|} \cdot (2\pi)^n \cdot \Phi_g(\Omega(A)) \right] &= \int_{\mathbb{R}^n} \partial_A^\beta \left[f(\xi) \, (\partial_\xi^\alpha \psi_{\Omega(A)})(\xi) \right] d\xi \\
\text{(Eq. (4.6))} &= (-1)^{|\beta|} \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot \int_{\mathbb{R}^n} f(\xi) \cdot \partial_\xi^\alpha \left(\xi^{\beta_\flat} \cdot \psi_{\Omega(A)}(\xi) \right) d\xi \\
&= (-1)^{|\beta|} \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot \left\langle f, \partial^\alpha \left[X^{\beta_\flat} \cdot \psi_{\Omega(A)} \right] \right\rangle_{\mathcal{S}', \mathcal{S}} \\
\text{(Eq. (4.2))} &= (-1)^{|\beta| + |\alpha|} \cdot \left(\frac{1}{2} \right)^{|\beta|_{li}} \cdot (2\pi)^n \cdot \left\langle X^{\beta_\flat} \cdot \partial^\alpha f, \mathcal{F}^{-1} \phi_{\Omega(A)} \right\rangle_{\mathcal{S}', \mathcal{S}}
\end{aligned}
$$

$$\left(g = \mathcal{F}^{-1}(\partial^\alpha f), \text{ Eq. (4.1), and } (-1)^{|\beta|} = i^{|\beta_\flat|}\right)$$

$$= \left(\frac{1}{2}\right)^{|\beta|_{\shortparallel}} \cdot (2\pi)^n \cdot (-1)^{|\alpha|} \cdot \left\langle \mathcal{F}[\partial^{\beta_\flat} g], \ \mathcal{F}^{-1}\phi_{\Omega(A)} \right\rangle_{\mathcal{S}',\mathcal{S}} .$$

In combination, this shows that $\Phi_g \circ \Omega$ is smooth on $B_\varepsilon(A_0)$, with partial derivatives given by

$$\partial^\beta \left[\Phi_g \circ \Omega\right](A) = \left(\frac{1}{2}\right)^{|\beta|_{\shortparallel}} \cdot \left\langle \mathcal{F}[\partial^{\beta_\flat} g], \ \mathcal{F}^{-1}\phi_{\Omega(A)} \right\rangle_{\mathcal{S}',\mathcal{S}}$$

$$= \left(\frac{1}{2}\right)^{|\beta|_{\shortparallel}} \cdot \left\langle \partial^{\beta_\flat} g, \ \phi_{\Omega(A)} \right\rangle_{\mathcal{S}',\mathcal{S}} ,$$

as claimed. Since $A_0 \in U$ was arbitrary, the proof is complete. $\qquad\square$

References

1. Brown, J.: Generalized form of Price's theorem and its converse. IEEE Trans. Inform. Theory **13**(1), 27–30 (1967)
2. Duistermaat, J.J., Kolk, J.A.C.: Distributions. Birkhäuser Boston Inc, Boston, MA (2010)
3. Folland, G.B.: Real Analysis: Modern Techniques and Their Applications. Pure and Applied Mathematics, 2nd edn. Wiley, New York (1999)
4. Genzel, M., Kutyniok, G., März, M.: ℓ^1-analysis minimization and generalized (co-)sparsity: when does recovery succeed? Appl. Comput. Harmon. Anal. (2020). https://doi.org/10.1016/j.acha.2020.01.002
5. Gut, A.: An Intermediate Course in Probability. Springer Texts in Statistics, 2nd edn. Springer, New York (2009)
6. McMahon, E.: An extension of Price's theorem (corresp.). IEEE Trans. Inform. Theory **10**(2), 168–168 (1964)
7. Papoulis, A.: Comments on 'an extension of Price's theorem' by McMahon. E. L. IEEE Trans. Inform. Theory **11**(1), 154–154 (1965)
8. Price, R.: A useful theorem for nonlinear devices having Gaussian inputs. IRE Trans. **4**, 69–72 (1958)
9. Vladimirov, I.G.: A quantum mechanical version of Price's theorem for Gaussian states. In: Control Conference (AUCC), 2014 4th Australian, pp. 118–123. IEEE, (2014)