(Preprint) AAS 15-733

# PROBABILITY DENSITY TRANSFORMATIONS ON ADMISSIBLE REGIONS FOR DYNAMICAL SYSTEMS

Johnny L. Worthy III*, Marcus J. Holzinger†

The admissible region as used for initial orbit determination is often expressed as a uniform multivariate probability density function (PDF). A multivariate PDF may be transformed and expressed in an alternate state space if the total probability is preserved over the transformation. This paper applies the general multivariate PDF transformation method to an admissible region to develop the conditions required for such a transformation. Because the probability must be preserved, it is shown that in general an admissible region PDF may not be transformed by a nonlinear transformation unless specific mapping conditions are met over all the state space volume. If this condition is not met then the transformation of an admissible region PDF yields incorrect probabilities over the state space. Further, it is also shown that if each state in an admissible region is locally observable then the constant gradient condition is lifted.

## INTRODUCTION

Observation systems often act in information deprived environments where a single observation cannot fully determine the state of an observed object. Characterizing the orbit of a space object from optical or radar measurements is an example of such an observation system since a single optical or radar measurement does not provide enough information to uniquely determine the state of the space object. However, with the increasing number of objects in orbit around the Earth, characterizing these space objects is an active area of research. Currently, over 20,000 objects larger than 10 cm are tracked and it is estimated that catastrophic collisions are likely to occur every 5 to 9 years [1]. These objects are currently tracked in the United States by a group of optical and radar sensors in the Space Surveillance Network (SSN) [2]. Both types of observing sensors operate in data deprived environments as optical measurements cannot determine the range to the target and radar measurements cannot determine the angular position of the space object. However, over long observation periods it is possible to use approaches such as Gauss's method or Lambert's method to fully estimate the state of a space object [3]. The difficulty with these types of observations is that over short observation periods (relative to the time scale of the dynamics) there is an unobservable subspace causing traditional methods, such as Gauss's, to fail. Because of this, the development of initial orbit determination approaches based on short-arc optical and radar measurements is a very active area of research.

Several nonlinear initial orbit determination approaches are based on the admissible region method. When an observation is too short to provide enough geometric data for an initial orbit estimate, a continuum of possible solution exist. The admissible region method uses hypothesized constraints

*Graduate Researcher, School of Aerospace Engineering, Georgia Institute of Technology
†Assistant Professor, School of Aerospace Engineering, Georgia Institute of Technology.

to bound the feasible solutions to a closed and computationally tractable state space volume [4]. First proposed by Milani et. al., the admissible region method constrains the possible solutions for a given too-short arc (TSA) observation using the dynamics of the orbiting bodies and hypothesized constraints [4] [5]. Many have extended the admissible region's applicability to SSA since Milani et. al. For example, methods have been presented that discretize the admissible region and consider solutions at discrete points [6] [7]. Multiple hypothesis filter or particle filter methods can also be initialized from discretized admissible regions [8]. Optimization methods to identify a best fitting orbit solution are presented by Siminski et. al. [9]. A boundary value problem approach is applied to the admissible region by Fujimoto and Alfriend which uses the angle-rate information to eliminate hypotheses [10].

In general, Bayesian estimation techniques utilize an a priori probability distribution of the initial state. Thus, admissible regions must be expressed probabilistically when used as a-priori distributions to initiate Bayesian estimation schemes. Fujimoto et. al. showed that the admissible region possesses a uniform probability density over the constrained unobservable state space volume; every state satisfying the constraint is equally likely to be true [11]. Without the inclusion of measurement, observer, and parameter uncertainty, however, a uniform probability distribution results in a probability discontinuity at the admissible region boundary. It is important to include uncertainty effects to remove this discontinuity when generating an admissible region in order to account for states that would otherwise have been assigned 0 probability. DeMars and Jah accomplish this by using a Gaussian mixture model (GMM) to approximate the the admissible region [12]. The GMM accounts for uncertainty in the Gaussian mixture approximation component size. An approximate analytical expression for the exact probability distribution of an admissible region is presented by Worthy and Holzinger by directly accounting for uncertainty and errors in the measurements and observer state [13]. Hussein et. al. generate a probabilistic admissible region by uniformly sampling from an admissible region constructed in a given state space [14]. The uniformly sampled points are then mapped into the desired state space and a Gaussian mixture is used to describe the new admissible region. This work motivates the question as to whether it is possible to mathematically transform an admissible region from a given state space into a different state space.

An admissible region is formulated in a specific state space based on the hypothesized constraints, however operational requirements may necessitate estimation schemes which operate in a different state space implying a probability mapping is needed. In general, probability mappings are only restricted by the requirement that the transformation be left invertible [15]. This condition preserves probability across the transformation and it is possible that additional conditions exist for admissible region probabilities. Due to this, care must be taken when transforming the admissible region probabilities between state spaces. The approach presented in [13] can be extended to assess the conditions under which an approximate admissible region may be initialized in a state space different from the original formulation. To address probability transformations of admissible regions, this paper comprises 1) formal analytical results regarding valid admissible region state space transformations, 2) a formal expression of the constant gradient condition which enables an admissible region probability distribution to be arbitrarily transformed into a different state space and 3) a novel extension of the approximate analytical probability distribution function for the transformation of an admissible region.

This paper is organized as follows. The general requirements and approach to transform probability distributions are described in §II.A with applicability to admissible regions shown in §II.B. In §III.A the theory is applied to assess the validity of the transformation of probability from topocentric spherical to cartesian coordinates and show the conditions for observability of the system.

## PROBABILITY TRANSFORMATIONS FOR ADMISSIBLE REGIONS

The general theory of probability transformations is an exhaustively studied topic in statistics and probability [15] [16] [17] with a wide range of applications. The purpose of this section is to introduce fundamental results regarding general probability mappings and apply it to the admissible region.

### General Probability Transformations

Given the probability density function (PDF) $f_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}_+$ of a random variable $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{x} \sim f_x(\mathbf{x})$, the cumulative distribution function (CDF) can be written as

$$F_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}[\mathbf{X} \leq \mathbf{x}] = \int_A f_X(\mathbf{x})\mathrm{d}\mathbf{x} \tag{1}$$

where the volume of integration is given by $A = (-\infty, x_1] \times \cdots \times (-\infty, x_n]$. Define a transformation $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ where $n \geq m$. Applying the transformation $\widetilde{\mathbf{x}} = \mathbf{g}(\mathbf{x})$, the CDF for the transformed variable is obtained using integration by substitution and is given by

$$F_{\widetilde{\mathbf{X}}}(\tilde{\mathbf{x}}) = \mathbb{P}[\widetilde{\mathbf{X}} \leq \tilde{\mathbf{x}}] = \int_{\tilde{A}} f_{\mathbf{X}}(\mathbf{g}^{-1}(\tilde{\mathbf{x}})) \cdot \mathrm{abs}\left(\left|\frac{\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})}{\partial \tilde{\mathbf{x}}}\right|\right) \tag{2}$$

where $\tilde{A} = (-\infty, \tilde{x}_1] \times \cdots \times (-\infty, \tilde{x}_n]$ and $\mathrm{abs}\left(\left|\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})/\partial \tilde{\mathbf{x}}\right|\right)$ is the determinant of the Jacobian matrix and the absolute value ensures $f_{\widetilde{\mathbf{X}}}(\tilde{\mathbf{x}})$ is non-negative for all values of $\tilde{\mathbf{x}}$ [15]. The integrand of Eqn. (2) is by definition the PDF of $\widetilde{\mathbf{X}} = \mathbf{g}(\mathbf{X})$ and thus the following foundational theorem in multivariate statistics gives the PDF of the transformed variable.

**Theorem 1** (Transformation theorem for continuous random variables [15])**.** *Given a PDF $f_{\boldsymbol{X}}(\boldsymbol{x})$ and a left-invertible transformation $\tilde{\boldsymbol{x}} = \boldsymbol{g}(\boldsymbol{x})$ then the PDF of the transformed variable $f_{\tilde{\boldsymbol{x}}}(\tilde{\boldsymbol{x}}_u)$ is given by*

$$f_{\widetilde{\boldsymbol{X}}}(\tilde{\boldsymbol{x}}) = \begin{cases} f_{\boldsymbol{X}}(\boldsymbol{g}^{-1}(\tilde{\boldsymbol{x}})) \cdot \mathrm{abs}(\left|\frac{\partial}{\partial \tilde{\boldsymbol{x}}}\boldsymbol{g}^{-1}(\tilde{\boldsymbol{x}})\right|) & \textit{for } \tilde{\boldsymbol{x}} \in \mathcal{R}(\boldsymbol{g}(\tilde{\boldsymbol{x}})) \\ 0 & \textit{otherwise} \end{cases} \tag{3}$$

*Proof.* The proof of Theorem 1 is given in [15]. □

This transformation of $\mathbf{X}$ into $\widetilde{\mathbf{X}}$ must also satisfy [18]

$$F_{\widetilde{\mathbf{X}}}(\tilde{\mathbf{x}}) = F_{\mathbf{X}}(\mathbf{g}^{-1}(\tilde{\mathbf{x}})) \tag{4}$$

where $F_{(\cdot)}$ denotes the CDF over $\tilde{\mathbf{x}}$ or $\mathbf{x}$. This implies that for a given transformation $\tilde{\mathbf{x}} = \mathbf{g}(\mathbf{x})$, the CDF must not be changed. In other words, if the CDF is known for $\mathbf{X}$ then the CDF is known for $\widetilde{\mathbf{X}}$.

**Corollary 1** (Equivalence of CDFs)**.** *Given a known CDF $F_X(\boldsymbol{x})$ for $\boldsymbol{x}$ and a once differentiable and right-invertible transformation $\tilde{\boldsymbol{x}} = \boldsymbol{g}(\boldsymbol{x})$, the CDF $F_{\tilde{X}}(\tilde{\boldsymbol{x}})$ for $\tilde{\boldsymbol{x}}$ must satisfy $F_{\widetilde{\boldsymbol{X}}}(\tilde{\boldsymbol{x}}) = F_{\boldsymbol{X}}(\tilde{\boldsymbol{x}})$.*

*Proof.* The proof of Corollary 1 follows directly from the derivation and analysis of Eqn. (1). By definition Eqn. (1) is equal to Eqn. (2) and thus Eqn. (4) must hold. □

**Admissible Region Transformations**

The purpose of the following subsections is to outline why, in general, and admissible region PDF cannot be transformed. The first subsection shows the application of the derivation of Eqn. (3) to the admissible region problem. Then the necessary conditions for an admissible region PDF to be transformed based on the definition of an admissible region are defined, followed by a discussion of the limitation of practical transformations satisfying these necessary conditions. The second subsection considers the case when an admissible region PDF is not considered to be uniform. The third subsection discusses the observability condition in the admissible region problem and discusses when Eqn. (3) may be applied to a PDF based on an admissible region without any additional conditions. As this section will prove, a uniform PDF representation is in contradiction to Theorem 1.

*Defining the Admissible Region*

The previous probability transformation techniques operate specifically on a PDF, thus it is important to distinguish the admissible region from a true PDF. The admissible region is generally taken to be a uniform PDF [11], however mathematically it is not a true PDF. Since every state within an admissible region satisfies the hypothesized constraints, there is no information available to indicate whether one admissible state is more probable than any other admissible state. Admissible regions are then often represented as uniform PDFs over the admissible region.

Defining the admissible region requires knowledge of a measurement model for the system being observed. Consider the a general nonlinear measurement model given by

$$\mathbf{y} = \mathbf{h}(\mathbf{x}; \mathbf{k}, t) \tag{5}$$

where the measurement function is defined as $\mathbf{h} : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R} \rightarrow \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^m$ is the measurement vector, $\mathbf{x} \in \mathbb{R}^n$ is the state, $\mathbf{k} \in \mathbb{R}^l$ is the parameter vector, which may include the observer state and any other necessary parameters, and $t$ is the time. As done in all admissible region approaches, the state vector can be partitioned in to determined states $\mathbf{x}_d \in \mathbb{R}^d$ and undetermined states $\mathbf{x}_u \in \mathbb{R}^u$ where $u + d = n$ [13]. This means that

$$\mathbf{y} = \mathbf{h}(\mathbf{x}_d; \mathbf{k}, t) \tag{6}$$

Admissible region approaches constrain this continuum of solutions using hypothesized constraints in the form $\kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0$ where $\kappa_i : \mathbb{R}^u \times \mathbb{R}^m \times \mathbb{R}^l \times \mathbb{R} \rightarrow \mathbb{R}$. The admissible region for the i$^{th}$ hypothesized constraint $\boldsymbol{\kappa}_i(\cdot)$ is then defined as

$$\mathcal{AR}_i := \{\mathbf{x}_u \in \mathbb{R}^u \mid \kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0\} \tag{7}$$

where $\mathcal{AR}_i \subseteq \mathbb{R}^u$. Furthermore, if there are $c$ such hypotheses then the total combined admissible region is given by

$$\mathcal{AR} = \bigcap_{i=1}^{c} \mathcal{AR}_i \tag{8}$$

where $\mathcal{AR}$ must be a compact set [19]. The requirement that $\mathcal{AR}$ be compact ensures the assumed uniform distribution has non-zero probability. Thus, each state $\mathbf{x} \in \mathcal{AR}$ can be assigned a non-zero uniform probability.

*Observability of Admissible States*

The admissible region approach allows the continuum of solutions possible for an underdetermined system to be bounded based on hypothesized constraints. The continuum of solutions for an underdetermined system indicates that the system is unobservable. The undetermined states are the unobservable states, and the admissible region bounds this unobservable subspace. Consider the general nonlinear dynamical and measurement model

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{9}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}; \mathbf{k}, t) \tag{10}$$

The conditions for observability of this system, given a single measurement, can be assessed by the observability gramian $\mathbf{P} \in \mathbb{S}_{+}^{n\times n}$ [20] which is given as

$$\mathbf{P}(t_f, t_0, \mathbf{x}(t)) = \int_{t_0}^{t} \mathbf{\Phi}^T(\tau, t_0) \frac{\partial \mathbf{h}(\mathbf{x}(\tau); \mathbf{k}, \tau)}{\partial \mathbf{x}(\tau)}^T \frac{\partial \mathbf{h}(\mathbf{x}(\tau); \mathbf{k}, \tau)}{\partial \mathbf{x}(\tau)} \mathbf{\Phi}(\tau, t_0) \mathrm{d}\tau \tag{11}$$

where $\mathbf{\Phi} : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ is the state transition matrix (STM). Consider the solution to the dynamical system given by Eqn. (9)

$$\mathbf{x}(t) = \phi(t; \mathbf{x}(t_0), t_0) \tag{12}$$

where $\phi$ is the flow function. The rank of the above observability gramian gives the dimension of the unobservable subspace of the system along $\mathbf{x}(t), t \in [t_0, t_f]$. A point in state space $\mathbf{x}(t)$ is observable if and only if $\mathrm{rank}\mathbf{P}(\mathrm{t_f}, \mathrm{t_0}, \mathbf{x}(\mathrm{t})) = \mathrm{n}$. If $\mathrm{rank}\mathbf{P}(\mathrm{t_f}, \mathrm{t_0}, \mathbf{x}(\mathrm{t})) < \mathrm{n}$ then there is an unobservable subspace which is realized as $\mathcal{N}(\mathbf{P}(t_f, t_0, \mathbf{x}(t)))$, the nullspace of the local observability gramian about $\mathbf{x}(t)$ over the time interval $t \in [t_0, t_f]$, and a state estimate admits a continuum of solutions that generate the same measurement sequence.

**Lemma 1** (Admissible Regions and System Observability)**.** *For all $\boldsymbol{x}(t)$ with $\boldsymbol{x}_u \in Æ\!R$ the observability gramian $\boldsymbol{P}(t_f, t_0, \boldsymbol{x}(t))$ has rank $\boldsymbol{P}(t_f, t_0, \boldsymbol{x}(t)) = d < n$ for the observation of a space object following Keplerian dynamics. Every point $\boldsymbol{x}_u$ subspace is therefore unobservable.*

*Proof.* Consider the Taylor series approximation of the state transition matrix.

$$\mathbf{\Phi}(\tau, t_0) = \mathbb{I}_6 + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\tau - t_0) + \text{H.O.T} \tag{13}$$

Taking the first order terms and assuming two body dynamics,

$$\mathbf{f}(\mathbf{x}, t) = \begin{bmatrix} \dot{\mathbf{r}} \\ -\frac{\mu}{||\mathbf{r}||^3}\mathbf{r} \end{bmatrix} \tag{14}$$

where $\mathbf{x} = \begin{bmatrix} \mathbf{r}^T & \dot{\mathbf{r}}^T \end{bmatrix}^T$ which allows,

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \mathbf{0} & \mathbb{I}_3 \\ \mathbf{M} & \mathbf{0} \end{bmatrix} \tag{15}$$

where $\mathbb{I}_v$ is the $v \times v$ identity matrix and

$$\mathbf{M} = \begin{bmatrix} \frac{3\mu r_x^2}{\|\mathbf{r}\|^5} - \frac{\mu}{\|\mathbf{r}\|^3} & \frac{3\mu r_x r_y}{\|\mathbf{r}\|^5} & \frac{3\mu r_x r_z}{\|\mathbf{r}\|^5} \\ \frac{3\mu r_y r_x}{\|\mathbf{r}\|^5} & \frac{3\mu r_y^2}{\|\mathbf{r}\|^5} - \frac{\mu}{\|\mathbf{r}\|^3} & \frac{3\mu r_y r_z}{\|\mathbf{r}\|^5} \\ \frac{3\mu r_z r_x}{\|\mathbf{r}\|^5} & \frac{3\mu r_z r_y}{\|\mathbf{r}\|^5} & \frac{3\mu r_z^2}{\|\mathbf{r}\|^5} - \frac{\mu}{\|\mathbf{r}\|^3} \end{bmatrix} \tag{16}$$

If the time interval is sufficiently small, then $\mathbf{M}(\tau - t_0)$ can be approximated to have a negligible contribution. To quantify sufficiently small, if

$$(\tau - t_0) \ll \frac{\|\mathbf{r}\|^3}{3\mu} \tag{17}$$

then each of the terms in $\mathbf{M}(\tau - t_0)$ are very small and considered negligible. Then,

$$\mathbf{\Phi}(\tau, t_0) \approx \mathbb{I}_6 + \begin{bmatrix} \mathbf{0} & (\tau - t_0)\mathbb{I}_3 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{18}$$

Over this time period satisfying Eqn. (17), the matrix $\partial \mathbf{h}(\mathbf{x}(\tau); \mathbf{k}, \tau)/\partial \mathbf{x}(\tau)$ may be considered a constant. With this, the observability gramian may be analytically integrated. Consider the partitioned state variables which allow,

$$\frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{h}}{\partial \mathbf{x}_d} & \frac{\partial \mathbf{h}}{\partial \mathbf{x}_u} \end{bmatrix} \tag{19}$$

By definition, $\partial \mathbf{h}/\partial \mathbf{x}_u = 0$ leaving

$$\frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{bmatrix} \mathbf{0} & \frac{\partial \mathbf{h}}{\partial \mathbf{x}_d} \end{bmatrix} \tag{20}$$

Given $n = 6$ and $d = 4$, the integration of Eqn. (11) using Eqn. (18) and assuming $t_0 = 0$ gives that

$$\mathbf{P}_1 = \begin{bmatrix} c_{1,1}\, t & c_{1,2}\, t & c_{1,3}\, t \\ c_{2,1}\, t & c_{2,2}\, t & c_{2,3}\, t \\ c_{3,1}\, t & c_{3,2}\, t & c_{3,3}\, t \\ t\ (c_{4,1_1}\, t + c_{4,1_2}) & t\ (c_{4,2_1}\, t + c_{4,2_2}) & t\ (c_{4,3_1}\, t + c_{4,3_2}) \\ c_{5,1}\, t^2 & c_{5,2}\, t^2 & c_{5,3}\, t^2 \\ c_{6,1}\, t^2 & c_{6,2}\, t^2 & c_{6,3}\, t^2 \end{bmatrix} \tag{21}$$

$$\mathbf{P}_2 = \begin{bmatrix} \frac{1}{2}t\ (c_{1,4_1}\, t + c_{1,4_2}) & c_{1,5}\, t^2 & c_{1,6}\, t^2 \\ t\ (c_{2,4_1}\, t + c_{2,4_2}) & c_{2,5}\, t^2 & c_{2,6}\, t^2 \\ t\ (c_{3,4_1}\, t + c_{3,4_2}) & c_{3,5}\, t^2 & c_{3,6}\, t^2 \\ t\ \left(c_{4,4_1}\, t^2 + c_{4,4_2}\, t + c_{4,4_3}\right) & t^2\ (c_{4,5_1}\, t + c_{4,5_2}) & c_{4,6_1}\, t^2\ (t + c_{4,6_2}) \\ t^2\ (c_{5,4_1}\, t + c_{5,4_2}) & c_{5,5}\, t^3 & c_{5,6}\, t^3 \\ t^2\ (c_{6,4_1} t + c_{6,4_2}) & c_{6,5}\, t^3 & c_{6,6}\, t^3 \end{bmatrix} \tag{22}$$

where

$$\mathbf{P}(t_f, t_0, \mathbf{x}(t)) = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 \end{bmatrix} \tag{23}$$

and each of the $c_{i,j}$ terms are constants which depend on $\partial\mathbf{h}/\partial\mathbf{x}_d$. Under the conditions given by Eqn. (17), it can be seen that for sufficiently short observations the rank of this matrix becomes 4. In particular, the higher order terms, $t^2$ and $t^3$, in the last two rows and last two columns cause these rows and columns to go to zero. Without loss of generality, the form of this matrix holds for any values of $n$ and $d$. The rank in this case is defined by the number of rows, or columns, containing terms to the 1$^{\text{st}}$ order of $t$. It can be shown that the number of rows, or columns, with $t$ is equal to the rank of $\partial\mathbf{h}/\partial\mathbf{x}_d$. This implies that the rank of $\mathbf{P}(t_f, t_0, \mathbf{x}(t))$ is $d$ as long as Eqn. (17) is satisfied. It can be stated that rank $\mathbf{P}(t_f, t_0, \mathbf{x}(t)) = n - u = d < n$. Because this is true for any point $\mathbf{x}_u$, all such points are unobservable.

□

*The Admissible Region PDF*

The probability that a given state $\mathbf{x}_u \in \mathbb{R}^u$ lies in the $i^{\text{th}}$ admissible region is then given by

$$\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] = \mathbb{P}[\kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0] \tag{24}$$

Without any additional information, the inequality defining $\mathcal{AR}_i$ in Equation (7) is a binary constraint and $\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] \in \{0, 1\}$ since each $\mathbf{x}_u$ has either 100% or 0% probability of satisfying the constraint. Thus the probability that $\mathbf{x}_u$ satisfies a given constraint $\kappa_i$ can be exactly expressed as a piecewise membership function defined as

$$m_i(\mathbf{x}_u) = \begin{cases} 1, & \kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0 \\ 0, & \kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) > 0 \end{cases} \tag{25}$$

Thus $\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] = m_i(\mathbf{x}_u)$, and the probability mass function for a particular constraint hypothesis can then be defined as [16]

$$f_{i,\mathbf{x}_u}(\mathbf{x}_u) = \frac{m_i(\mathbf{x}_u)}{\int_{\mathcal{AR}_i} \mathrm{d}\mathbf{x}_\mathrm{u}} \tag{26}$$

Eqn. (26) results in a uniform distribution, which is demonstrated in [11]. Applying the chain rule of probabilities, the general joint probability function can be written as [17]

$$\begin{aligned} f_{\mathbf{x}_u}(\mathbf{x}_u) &= \frac{\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}]}{\int_{\mathcal{AR}} d\mathbf{x}_u} \\ &= \frac{1}{\int_{\mathcal{AR}} d\mathbf{x}_u} \prod_{k=1}^{c} \mathbb{P}\left[\mathbf{x}_u \in \mathcal{AR}_k \mid \mathbf{x}_u \in \bigcap_{j=1}^{k-1} \mathcal{AR}_j\right] \end{aligned} \tag{27}$$

where the bracketed term gives the probability that $k^{\text{th}}$ constraint is satisfied given that each of the $k-1$ previous constraints are satisfied. If the constraints $\kappa_i$ are assumed to be independent, then by Bayes' rules the conditional probability terms evaluate to 1 and Eqn. (27) simplifies to

$$f_{\mathbf{x}_u}(\mathbf{x}_u) = \frac{\prod_{k=1}^{c} \mathbb{P}\left[\mathbf{x}_u \in \mathcal{AR}_k\right]}{\int_{\mathcal{AR}} d\mathbf{x}_u} \tag{28}$$

$$= \frac{\prod_{k=1}^{c} m_k(\mathbf{x}_u)}{\int_{\mathcal{AR}} d\mathbf{x}_u} \tag{29}$$

By this formulation, every $\mathbf{x}_u \in \mathcal{AR}$ is a candidate solution that satisfies the $c$ constraints and without additional information; no one state can be considered more likely than another. Thus $f_{\mathbf{x}_u}(\mathbf{x}_u)$ is a constant over $\mathcal{AR}$ and as such the admissible region must be considered a uniform distribution. This fact exactly agrees with work presented by Fujimoto and Scheeres stating that without any a priori information regarding the observation, an admissible region is expressed as a uniform PDF [21].

*Transformation of the Admissible Region PDF*

Suppose a user wishes to use the admissible region method to initiate an estimation procedure in a state space other than the state spate in which the admissible region constraints are formed. Following the general probability transformation approach a transformation $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ can be defined. This transformation can also be partitioned into $\mathbf{g}_u : \mathbb{R}^u \to \mathbb{R}^u$ and $\mathbf{g}_d : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\tilde{\mathbf{x}}_u = \mathbf{g}_u(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \tag{30}$$

$$\tilde{\mathbf{x}}_d = \mathbf{g}_d(\mathbf{x}_d, \mathbf{y}, \mathbf{k}, t) \tag{31}$$

For simplicity, this transformation will be expressed as $\mathbf{g}_u(\mathbf{x}_u; \cdot)$ for the remainder of this paper. In general, the transformation $\mathbf{g}_u(\mathbf{x}_u; \cdot)$ must be left invertible for Eqn (3) to be applicable. Additionally, the transformation must satisfy the condition that the underdetermined and determined states in the transformed space are still capable of being partitioned, leading to the following Lemma.

**Lemma 2** (Partitioned Transformed State)**.** *A PDF of an admissible region expressed in state space* $\boldsymbol{x}_u$ *may be transformed to state space* $\tilde{\boldsymbol{x}}_u = \boldsymbol{g}_u(\boldsymbol{x}_u; \cdot)$ *only if there exist some* $\tilde{\boldsymbol{x}}_d = \bar{\boldsymbol{g}}_d(\boldsymbol{y}; \cdot)$, $\bar{\boldsymbol{g}}_d : \mathbb{R}^m \to \mathbb{R}^d$ *such that* $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}_d, \boldsymbol{k}, t) = \tilde{\boldsymbol{h}}(\tilde{\boldsymbol{x}}_d, \boldsymbol{k}, t)$, $\tilde{\boldsymbol{h}} : \mathbb{R}^d \times \mathbb{R}^l \times \mathbb{R} \to \mathbb{R}^m$.

*Proof.* The undetermined states $\mathbf{x}_u$ are independent of the determined states $\mathbf{x}_d$ as defined in [13]. This enables the partitioning of the state space such that the measurement $\mathbf{y}$ is only a function of the determined states, the parameters $\mathbf{k}$, and time and can be expressed by

$$\mathbf{y} = \mathbf{h}(\mathbf{x}_u, \mathbf{x}_d, \mathbf{k}, t) = \mathbf{h}(\mathbf{x}_d, \mathbf{k}, t)$$

which by definition means $\mathbf{x}_d = \mathbf{h}^{-1}(\mathbf{y}, \mathbf{k}, t)$. If there is a transformation of $\mathbf{x}_u$, then the transformation can be given by

$$\begin{aligned}\tilde{\mathbf{x}}_d &= \mathbf{g}_d(\mathbf{x}_d) \\ &= \mathbf{g}_d(\mathbf{h}^{-1}(\mathbf{y}, \mathbf{k}, t))\end{aligned}$$

which can be defined as $\bar{\mathbf{g}}_d = \mathbf{g}_d \circ \mathbf{h}^{-1} : \mathbb{R}^m \to \mathbb{R}^d$ giving,

$$\tilde{\mathbf{x}}_d = \bar{\mathbf{g}}_d(\mathbf{y}; \cdot)$$

Thus, the measurement function is now expressed by

$$\mathbf{y} = \tilde{\mathbf{h}}(\mathbf{g}_u(\mathbf{x}_u; \cdot), \bar{\mathbf{g}}_d(\mathbf{y}; \cdot), \mathbf{k}, t)$$

For the admissible region problem, it is required that $\tilde{\mathbf{x}}$ can be partitioned into $\tilde{\mathbf{x}}_u$ and $\tilde{\mathbf{x}}_d$ such that $\mathbf{y}$ is independent of $\tilde{\mathbf{x}}_u$. In general $\tilde{\mathbf{h}}(\mathbf{g}_u(\mathbf{x}_u; \cdot), \bar{\mathbf{g}}_d(\mathbf{y}; \cdot), \mathbf{k}, t) \neq \tilde{\mathbf{h}}(\bar{\mathbf{g}}_d(\mathbf{y}; \cdot), \mathbf{k}, t)$ since the transformation is not necessarily a function solely of $\mathbf{x}_u$. Thus, the function $\bar{\mathbf{g}}_d(\mathbf{y}; \cdot)$ must be defined to ensure that the determined variables are transformed such that the transformed undetermined states remain independent of the measurements. If a transformation $\bar{\mathbf{g}}_u(\mathbf{y};\cdot)$ cannot be defined such that this is true then

$$\mathbf{y} = \tilde{\mathbf{h}}(\mathbf{g}_u(\mathbf{x}_u;\cdot), \bar{\mathbf{g}}_d(\mathbf{y};\cdot), \mathbf{k}, t) \neq \tilde{\mathbf{h}}(\bar{\mathbf{g}}_d(\mathbf{y};\cdot), \mathbf{k}, t)$$

and the admissible region formulation is invalid. $\square$

The result of Lemma 2 essentially requires that if the undetermined states can be transformed then they must remain independent of the observations. Because this is a requirement for the formation of an admissible region, any transformation that does not satisfy Lemma 2 necessarily generates a region that can no longer be defined an $\mathit{Æ}$.

Assuming a transformation satisfying Lemma 2 exists, the admissible region in the transformed space can be defined. For the admissible region problem, since the constraint hypothesis is a function of a unique state $\mathbf{x}_u$,

$$\kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) = \tilde{\kappa}_i(\mathbf{g}_u(\mathbf{x}_u;\cdot), \mathbf{y}, \mathbf{k}, t) \tag{32}$$

$$\tilde{\kappa}_i(\tilde{\mathbf{x}}_u, \mathbf{y}, \mathbf{k}, t) = \kappa_i(\mathbf{g}_u^{-1}(\tilde{\mathbf{x}}_u;\cdot), \mathbf{y}, \mathbf{k}, t) \tag{33}$$

Eqns. (32) and (33) then imply that $\mathbb{P}[\tilde{\mathbf{x}}_u \in \widetilde{\mathit{Æ}}_i] = \mathbb{P}[\mathbf{x}_u \in \mathit{Æ}_i]$ and $m_i(\mathbf{x}_u) = m_i(\tilde{\mathbf{x}}_u)$ where,

$$\widetilde{\mathit{Æ}}_i := \{\tilde{\mathbf{x}}_u \in \mathbb{R}^u \mid \tilde{\boldsymbol{\kappa}}_i(\tilde{\mathbf{x}}_u, \mathbf{y}, \mathbf{k}, t) \leq 0\} \tag{34}$$

and

$$\tilde{m}_i(\tilde{\mathbf{x}}_u) = \begin{cases} 1, & \tilde{\boldsymbol{\kappa}}_i(\tilde{\mathbf{x}}_u, \mathbf{y}, \mathbf{k}, t) \leq 0 \\ 0, & \tilde{\boldsymbol{\kappa}}_i(\tilde{\mathbf{x}}_u, \mathbf{y}, \mathbf{k}, t) > 0 \end{cases} \tag{35}$$

The general PDF in the transformed space is given by

$$f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u) = \frac{1}{\int_{\widetilde{\mathit{Æ}}} \mathrm{d}\tilde{\mathbf{x}}_u} \prod_{k=1}^{c} \mathbb{P}\left[\tilde{\mathbf{x}}_u \in \widetilde{\mathit{Æ}}_k \mid \tilde{\mathbf{x}}_u \in \bigcap_{j=1}^{k-1} \widetilde{\mathit{Æ}}_j\right] \tag{36}$$

Assuming again that the constraint hypotheses are independent, the PDF expressed in $\tilde{\mathbf{x}}_u$ is given by,

$$f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u) = \frac{\prod_{k=1}^{c} \tilde{m}_k(\tilde{\mathbf{x}}_u)}{\int_{\widetilde{\mathit{Æ}}} \mathrm{d}\tilde{\mathbf{x}}_\mathrm{u}} \tag{37}$$

A general nonlinear transformation of a uniform PDF will yield a non-uniform PDF according to Eqn. (3). The uniform PDF of an admissible region is a probabilistic representation of the fact that each state $\mathbf{x}_u \in \mathit{Æ}$ is consistent with the measurement $\mathbf{y}$. Without any additional information, each state necessarily has equal probability which must also be true if $\mathbf{x}_u$ is expressed in any other state space. Given this fact, the necessary relationship between $f_{\mathbf{X}_u}(\mathbf{x}_u)$ and $f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u)$ is given by Theorem 2.

**Theorem 2** (Equivalence of Admissible Regions)**.** *Given $\boldsymbol{x}_u \in \mathit{Æ}$ and an invertible transformation $\tilde{x}_u = \boldsymbol{g}_u(\boldsymbol{x}_u;\cdot)$, a PDF transformation for the admissible region problem is only valid if the transformation $\boldsymbol{g}$ satisfies $|\partial \boldsymbol{x}_u / \partial \tilde{\boldsymbol{x}}_u| = \zeta \; \forall \, \boldsymbol{x}_u \in \mathit{Æ}$ where $\zeta$ is the ratio of the volume of the admissible region as expressed in both state spaces and $f_{\tilde{\mathbf{X}}_u}(\tilde{\boldsymbol{x}}_u) = \zeta f_{\boldsymbol{x}_u}(\boldsymbol{x}_u)$.*

*Proof.* The proof of Theorem 2 is given by way of contradiction. Assume there exists an invertible transformation $\mathbf{g}_u(\mathbf{x}_u;\cdot)$ for which $|\partial\mathbf{x}_u/\partial\tilde{\mathbf{x}}_u| \neq \zeta$ for some $\mathbf{x} \in \mathcal{AR}$. The relationship between $f_{\mathbf{x}_u}(\mathbf{x}_u)$ and $f_{\tilde{\mathbf{x}}_u}(\tilde{\mathbf{x}}_u)$ may be determined by applying Eqn. (3) as follows

$$f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u) = \frac{\prod_{k=1}^{c} m_k\left(\mathbf{g}^{-1}(\tilde{\mathbf{x}}_u)\right)}{\int_{\mathcal{AR}} \mathrm{d}\mathbf{x}_u} \mathrm{abs}\left(\left|\frac{\partial\mathbf{g}_{\mathrm{u}}^{-1}(\tilde{\mathbf{x}}_{\mathrm{u}})}{\partial\tilde{\mathbf{x}}_{\mathrm{u}}}\right|\right) \tag{38}$$

Each of the terms in Eqn. (38) have been defined thus far except for the Jacobian term $\left|\partial\mathbf{g}^{-1}(\tilde{\mathbf{x}}_u)/\partial\tilde{\mathbf{x}}_u\right|$. Rearranging Eqn. (38), by substituting the $\tilde{\mathbf{x}}_u$ PDF on the left hand side and multiplying by the denominator of the right hand side,

$$\frac{\prod_{k=1}^{c} \tilde{m}_k\left(\tilde{\mathbf{x}}_u\right)}{\int_{\widetilde{\mathcal{AR}}} \mathrm{d}\tilde{\mathbf{x}}_u} \int_{\mathcal{AR}} \mathrm{d}\mathbf{x}_u = \prod_{k=1}^{c} m_k\left(\mathbf{g}_u^{-1}(\tilde{\mathbf{x}}_u)\right) \mathrm{abs}\left(\left|\frac{\partial\mathbf{g}_{\mathrm{u}}^{-1}(\tilde{\mathbf{x}}_{\mathrm{u}})}{\partial\tilde{\mathbf{x}}_{\mathrm{u}}}\right|\right) \tag{39}$$

Note that for the admissible region approach $m_k(\mathbf{x}_u) = \tilde{m}_k(\tilde{\mathbf{x}}_u)$ since it is necessary that $\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] = \mathbb{P}[\tilde{\mathbf{x}}_u \in \widetilde{\mathcal{AR}}_i]$. Thus, dividing each side by $\prod_{k=1}^{c} \tilde{m}_k\left(\tilde{\mathbf{x}}_u\right)$ results in,

$$\frac{\int_{\mathcal{AR}} \mathrm{d}\mathbf{x}_u}{\int_{\widetilde{\mathcal{AR}}} \mathrm{d}\tilde{\mathbf{x}}_u} = \zeta = \mathrm{abs}\left(\left|\frac{\partial\mathbf{g}_{\mathrm{u}}^{-1}(\tilde{\mathbf{x}}_{\mathrm{u}})}{\partial\tilde{\mathbf{x}}_{\mathrm{u}}}\right|\right) \tag{40}$$

If $|\partial\mathbf{x}_u/\partial\tilde{\mathbf{x}}_u| \neq \zeta$ then,

$$\frac{\int_{\mathcal{AR}} \mathrm{d}\mathbf{x}_u}{\int_{\widetilde{\mathcal{AR}}} \mathrm{d}\tilde{\mathbf{x}}_u} \neq \zeta \tag{41}$$

which then implies $\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] \neq \mathbb{P}[\tilde{\mathbf{x}}_u \in \widetilde{\mathcal{AR}}_i]$ for Eqn. (39) to hold. But this is a contradiction since the admissible region requires $\mathbb{P}[\mathbf{x}_u \in \mathcal{AR}_i] = \mathbb{P}[\tilde{\mathbf{x}}_u \in \widetilde{\mathcal{AR}}_i]$ regardless of the transformation. □

Theorem 2 imposes a geometric constraint on the transformation $\mathbf{g}$ through the determinant of the Jacobian. The constraint requires the determinant to be constant which implies the distortion of the $\tilde{\mathbf{x}}_u$ state space relative to the $\mathbf{x}_u$ state space is the same at every point. This is necessary to ensure that any one point inside the admissible region in $\mathbf{x}_u$ remains inside the equivalent admissible region expressed in $\tilde{\mathbf{x}}_u$. Furthermore, this constraint limits the practical applicability of probability transformations to admissible region because useful state space transformation are often complex, nonlinear functions.

Given that a transformation $\mathbf{g}$ exists which satisfies Theorem 2, it is possible to define the transformed PDF. The final expression for the transformed PDF is then given by

$$f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u) = \zeta \frac{\prod_{k=1}^{c} m_k\left(\mathbf{x}_u\right)}{\int_{\mathcal{AR}} \mathrm{d}\mathbf{x}_u} \tag{42}$$

Eqn. (42) signifies that for the admissible region problem with no additional information, the admissible region of $\mathbf{x}_u$ expressed in any transformed state space $\tilde{\mathbf{x}}_u$ such that $\mathbf{g}^{-1}(\tilde{\mathbf{x}}_u)$ exists is necessarily uniform and simply scaled by a factor $\zeta$. Thus, the admissible region can be expressed in any state space, given that the transformation satisfies Theorem 2, as a uniform PDF which agrees with the work shown in [11]. It should be noted that useful transformations are often highly non-linear and as such will not typically satisfy the conditions presented by Theorem 2. It is likely that, in general, an admissible region PDF cannot be transformed since no practical transformations exists satisfying Theorem 2. If an admissible region PDF is transformed by a transformation not satisfying Theorem 2, then the PDF in the transformed space is no longer a uniform representation of the state space, and this non-uniform representation is not based on statistical information but based only on the transformation function. Because of this, any transformation not satisfying Theorem 2 generates an admissible region PDF misrepresenting the true distribution.

*Non-Uniform Æ PDFs*

While Eqn. (42) applies for transformations of uniform PDFs, it may also be applied to non-uniform admissible region PDFs. An approach for generating a non-uniform PDF of $\mathbf{x}_u$ is shown in [13]. The approximate analytical probability for a given admissible region is given by

$$\mathbb{P}[(\mathbf{x}_u \in \mathcal{AE}_i)] = m_i(\mathbf{x}_u) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\|\mathbf{x}_\text{u} - \mathbf{x}_{\text{u},\mathcal{B}_\perp,\text{i}}\|}{\sqrt{2\text{tr}\mathbf{P}_{\mathbf{x}_{\text{u},\mathcal{B}_\perp,\text{i}}}}}\right)\right] \tag{43}$$

which updates the piecewise membership function given by Eqn. (25) to a continuous membership function by including uncertainty effects. These uncertainties are quantified as the covariance matrix $\mathbf{P}_\mathbf{z}$ where $\mathbf{z}$ is the combined matrix of the measurements, parameters, and time. The quantity $\mathbf{x}_{u,\mathcal{B}_\perp,i}$ is the point on the boundary of $\mathcal{AE}_i$ orthogonal to $\mathbf{x}_u$ and $\mathbf{P}_{\mathbf{x}_{u,\mathcal{B}_\perp,i}}$ is the covariance calculated at this boundary point. Substituting Eqn. (43) into Eqn. (29) then gives the non-uniform PDF.

**Corollary 2** (Systematic Uncertainty in Admissible Regions)**.** *If the combined measurements, parameters, and time covariance matrix $\boldsymbol{P}_\mathbf{z}$ is known then transformation of the non-uniform admissible region probability is given by*

$$\mathbb{P}[(\tilde{\boldsymbol{x}}_u \in \widetilde{\mathcal{AE}}_i)] = \tilde{m}_i(\tilde{\boldsymbol{x}}_u) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\|\boldsymbol{g}_\text{u}(\boldsymbol{x}_\text{u};\cdot) - \boldsymbol{g}_\text{u}(\boldsymbol{x}_{\text{u},\mathcal{B}_\perp,\text{i}};\cdot)\|}{\sqrt{2\text{tr}\boldsymbol{P}_{\tilde{\boldsymbol{x}}_{\text{u},\mathcal{B}_\perp,\text{i}}}}}\right)\right] \tag{44}$$

*where $\boldsymbol{P}_{\tilde{\boldsymbol{x}}_{u,\mathcal{B}_\perp,i}}$ is the modified covariance matrix.*

*Proof.* Given the previous transformation $\tilde{\mathbf{x}}_u = \mathbf{g}_u(\mathbf{x}_u;\cdot)$, Eqn. (43) can be derived for $\tilde{\mathbf{x}}_u$. The simplified Taylor series expansion from Eqn. (17) in [13] now becomes

$$-\frac{\partial\kappa_i}{\partial\tilde{\mathbf{x}}_u}\frac{\partial\tilde{\mathbf{x}}_u}{\partial\mathbf{x}_u}\delta\mathbf{X}_u = \frac{\partial\kappa_i}{\partial\mathbf{z}}\delta\mathbf{Z} \tag{45}$$

Carrying the notation defined in [13], a new perpendicular vector $\tilde{\mathbf{p}}$ is defined as

$$\tilde{\mathbf{p}} = \frac{\partial\kappa_i}{\partial\tilde{\mathbf{x}}_u}\frac{\partial\mathbf{g}_u^{-1}(\mathbf{x}_u;\cdot)}{\partial\mathbf{x}_u}\bigg|_{\mathbf{x}_u} \tag{46}$$

The rest of the derivation can be carried out as specified in [13] by replacing $\mathbf{p}$ with $\tilde{\mathbf{p}}$ resulting in

$$\widetilde{\mathbf{M}} = \begin{bmatrix}\tilde{\mathbf{p}}^T \\ \mathbf{T}\end{bmatrix}^{-1}\begin{bmatrix}-\frac{\partial\kappa_i}{\partial\mathbf{z}} \\ \mathbf{0}\end{bmatrix}$$

where $\mathbf{T} \in \mathbb{R}^{u-1\times u}$ is a matrix of tangential unit vectors which gives

$$\mathbf{P}_{\tilde{\mathbf{x}}_u} = \widetilde{\mathbf{M}}\mathbf{P}_\mathbf{z}\widetilde{\mathbf{M}}^T \tag{47}$$

Eqn. (44) is obtained by substituting $\mathbf{P}_{\tilde{\mathbf{x}}_u}$ and $\mathbf{g}(\mathbf{x}_u)$ into Eqn. (43). □

Eqn. (44) defines the approximate analytical probability distribution function for an admissible region $\mathcal{AR}_i$ in the $\tilde{\mathbf{x}}_u$ space. Alternatively, from Eqns. (32) and (33)

$$\begin{aligned}\mathbb{P}[(\mathbf{x}_u \in \mathcal{AR}_i)] &= \mathbb{P}[(\tilde{\mathbf{x}}_u \in \widetilde{\mathcal{AR}}_i)] & (48)\\ &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\|\mathbf{g}_{\mathrm{u}}(\mathbf{x}_{\mathrm{u}};\cdot) - \mathbf{g}_{\mathrm{u}}(\mathbf{x}_{\mathrm{u}};\cdot)_{\mathrm{u},\mathcal{B}_\perp}\|}{\sqrt{2\mathrm{tr}\mathbf{P}_{\tilde{\mathbf{x}}_{\mathrm{u},\mathcal{B}_\perp}}}}\right)\right] & (49)\\ &\approx \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\|\mathbf{x}_{\mathrm{u}} - \mathbf{x}_{\mathrm{u},\mathcal{B}_\perp}\|}{\sqrt{2\mathrm{tr}\mathbf{P}_{\mathbf{x}_{\mathrm{u},\mathcal{B}_\perp}}}}\right)\right] & (50)\end{aligned}$$

Because of this, it is equivalent to directly map each $\mathbf{x}_u$ to $\tilde{\mathbf{x}}_u$ and assign each $\tilde{\mathbf{x}}_u = \mathbf{g}(\mathbf{x}_u)$ the probability of set membership $\mathbb{P}[\mathbf{x} \in \mathcal{AR}_i]$ or vice-versa.

*The Observability Condition*

Lemma 1 shows that the existence of the admissible region implies that there is a non-trivial unobservable subspace of the system given a short enough observation. However, it is possible for the system to become fully observable given enough observations or a long enough observation of the system. Thus, it is of interest to understand how the observability of a system affects the transformation of the PDF associated with an admissible region. By definition, after the initial measurement any state in the admissible region is locally unobservable and cannot be transformed by Eqn. (3) unless Theorem 2 is satisfied. However, if an additional measurement can be taken at a time $t$ such that each state $\mathbf{x}_u \in \mathcal{AR}$ is locally observable, then an a posteriori PDF can be constructed. This a posteriori PDF represents a true PDF over the state space and can be used directly with Eqn. (3) to transform probabilities between state spaces. As such, it is of interest to determine when the states $\mathbf{x}_u \in \mathcal{AR}$ become locally observable.

**Corollary 3** (Observability in Admissible Region Problems)**.** *If the observability gramian for the admissible region system satisfies* $\mathrm{rank}\boldsymbol{P}(\mathrm{t_f}, \mathrm{t_0}, \boldsymbol{x}(\mathrm{t})) = \mathrm{n}$ *where* $\boldsymbol{x}(t) = [\boldsymbol{x}_d(t)\ \ \boldsymbol{x}_u(t)]\,\forall\, \boldsymbol{x}_u \in \mathcal{AR}$ *then the PDF associated with the admissible region estimate may be transformed without the condition* $|\partial \boldsymbol{x}_u / \partial \tilde{\boldsymbol{x}}_u| = \zeta \;\forall\, \boldsymbol{x}_u \in \mathcal{AR}$.

*Proof.* The admissible region $\mathcal{AR}$ is, as defined, a subset of the unobservable state space where each state $\mathbf{x}_u \in \mathcal{AR}$ has no effect on the measurements. Since the mapping $\mathbf{h}$ from $\mathbf{x}$ to $\mathbf{y}$ cannot be a one-to-one and onto, each $\mathbf{x}_u \in \mathcal{AR}$ must necessarily have a uniform probability. Because this is also true in any transformed state space $\tilde{\mathbf{x}}$, the admissible region must necessarily be uniform in any state space. If a system is locally observable at $\mathbf{x}_k \in \mathbb{R}^n$, where $k$ is an arbitrary index, then there exists a measurement function $\mathbf{h}_o : \mathbb{R}^n \to \mathbb{R}^m$ where $\mathbf{h}_o$ is a one-to-one and onto function. Thus, $\mathbf{x}_j \neq \mathbf{x}_k \implies \mathbf{h}_o(\mathbf{x}_j) \neq \mathbf{h}_o(\mathbf{x}_k)$ and each unique observation corresponds to a unique state $\mathbf{x}$. If the transformation $\mathbf{g}(\mathbf{x})$ is also one-to-one and onto then there must also exist a measurement function $\tilde{\mathbf{h}}_o : \mathbb{R}^n \to \mathbb{R}^m$ such that $\tilde{\mathbf{x}}_j \neq \tilde{\mathbf{x}}_k \implies \tilde{\mathbf{h}}_o(\tilde{\mathbf{x}}_j) \neq \tilde{\mathbf{h}}_o(\tilde{\mathbf{x}}_k)$ and $\mathbf{h}_o(\mathbf{x}_j) = \tilde{\mathbf{h}}_o(\tilde{\mathbf{x}}_j) = \mathbf{y}$. A unique solution exists for a given observation, or set of observations, and a PDF can then be defined about that solution. Because this unique PDF cannot be identical in both state spaces, the condition given by $|\partial \mathbf{x}_u / \partial \tilde{\mathbf{x}}_u| = \zeta$ can no longer hold, and for an observable system the PDF can simply be transformed by Eqn. (3). □

The main result of Corollary 3 is that the PDF associated with a given $\mathcal{AR}$ may generally not be transformed until it is observable. Since there are likely no practical transformations that satisfy Theorem 2, the significance of Corollary 3 is in the fact that general admissible region PDF transformations are possible, but only once each state in $\mathcal{AR}$ becomes locally observable. Furthermore, by Lemma 1, if every $\mathbf{x}_u \in \mathcal{AR}$ is locally observable, then the region is necessarily not an admissible region.

## Additional Transformations

This section discusses additional transformations that apply to the probability transformation theorems, corollaries, and lemmas presented in this work.

*Linear Transformations*

The only set of functions that will always satisfy Theorem 2 are linear transformations leading to Remark 1.

**Remark 1:** Any linear transformation $\tilde{\mathbf{x}}_u = \mathbf{g}_u(\mathbf{x}_u) = \mathbf{T}_u\mathbf{x}_u$ such that $\mathbf{T}_u \in \mathbb{R}^{n\times n}, \mathrm{rank}\mathbf{T}_\mathrm{u} = \mathrm{n}$ that can be defined $\forall \mathbf{x}_u \in \mathcal{AR}$ will satisfy the requirements given by Lemma 1 and Theorem 2. Thus, for any linear transformation of an admissible region, $\zeta$ can be defined such that $f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u) = \zeta f_{\mathbf{X}_u}(\mathbf{x}_u)$.

Any invertible linear transformation of covariance in extended Kalman filters satisfies Theorem 2 as long as the covariance is sufficiently small []. Since linear transformations are trivial for an admissible region problem, it can be stated that in general $\mathcal{AR}$ cannot be transformed due to Theorem 2.

*Sigma Point Transformations*

An additional application of the general probability transformation comes from sigma point transformations and filters [25]. Sigma point filters use transformations of the sigma points of a Gaussian PDF to map the PDF over nonlinear transformations, used largely in the Unscented Kalman Filter. The sigma point transformation as originally defined relies on the fact that the transformation preserves the mean and covariance [26]. Alternatively, the sigma point transformation must preserve the PDF. Assume a PDF $f_{\mathbf{x}}(\mathbf{x})$ is known for a given $\mathbf{x}$, then the first order Taylor Series expansion of the inverse of the transformation $\tilde{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ is given by

$$\mathbf{x} + \delta\mathbf{x} = \mathbf{g}^{-1}(\tilde{\mathbf{x}}) + \frac{\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})}{\partial \tilde{\mathbf{x}}}\delta\tilde{\mathbf{x}} \tag{51}$$

$$\mathbf{x} + \delta\mathbf{x} = \mathbf{x} + \frac{\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})}{\partial \tilde{\mathbf{x}}}\delta\tilde{\mathbf{x}} \tag{52}$$

$$\delta\mathbf{x} = \frac{\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})}{\partial \tilde{\mathbf{x}}}\delta\tilde{\mathbf{x}} \tag{53}$$

Since a sigma point transformation aims to preserve the mean and covariance a transformation given by $|\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})/\partial\tilde{\mathbf{x}}| = 1$ is a valid sigma point transformation since the PDFs of $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are the same. However, if $|\partial \mathbf{g}^{-1}(\tilde{\mathbf{x}})/\partial\tilde{\mathbf{x}}| = c$ where $c$ is a constant for all $\mathbf{x}$ in the vicinity of the Gaussian PDF parameterized by the sigma points then the PDF is also preserved by the scaling factor $c$ and the PDFs can be written as $f_{\mathbf{x}}(\mathbf{x}) = f_{\tilde{\mathbf{x}}}(\tilde{\mathbf{x}})/c$. This result is analogous to Theorem 2, since the PDF must be preserved for admissible regions and the PDF must be preserved for sigma point transformations, the scaling factor $c$ is similar to $\zeta$ for admissible regions.

*Transformations Over Time*

General probability transformations also apply to transformations through time as shown by Park and Scheeres [22]. Here it is shown that the framework presented in this paper is consistent with these existing methods. Given an initial PDF for a system, it is often useful to know how that PDF changes over time. Consider the following system dynamics

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{54}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The solution is expressed as

$$\mathbf{x}(t) = \boldsymbol{\phi}(t; \mathbf{x}_0, t_0) \tag{55}$$

where the subscript '0' denotes the initial state, $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\boldsymbol{\phi}$ is the flow function satisfying

$$\frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{dt}} = \mathbf{f}(\boldsymbol{\phi}(t; \mathbf{x}_0, t_0), t) \tag{56}$$

$$\boldsymbol{\phi}(t_0; \mathbf{x}_0, t_0) = \mathbf{x}_0 \tag{57}$$

In the case of time transformations, the function $\boldsymbol{\phi}$ is the transformation function $\mathbf{g}(\cdot)$. The PDF transformation of a dynamical system over time comes from analysis of the Fokker-Planck equation. If the system introduced above satisfies the Itô stochastic differential equation, then the time evolution of the PDF stochastic variable $\mathbf{X}$ over time is given by the Fokker-Planck equation [23]

$$\frac{\partial f_{\mathbf{x}}(\mathbf{x}, t)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial \mathbf{x}_i} \left( f_{\mathbf{x}}(\mathbf{x}, t) \mathbf{f}_i(\mathbf{x}, t) \right) \tag{58}$$

assuming no diffusion terms. Park and Scheeres show the integral invariance of a PDF through the solution to this simplified Fokker-Planck equation for a system with no diffusion resulting in [22] [24].

$$f(t, \boldsymbol{\phi}(\mathbf{x}_0, t_0), t) = f(x_0, t_0) \left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right|^{-1} \tag{59}$$

which is the exact form given in Eqn. (3). Under Hamiltonian dynamics, Liouville's theorem proves that $|\partial \mathbf{x} / \partial \mathbf{x}_0| = 1$ for all time $t$ since the transformation over time is a Canonical transformation [22]. Thus, for a Hamiltonian system Eqn. (59) simplifies further since the Jacobian term evaluates to unity and if the PDF is known at any time, it is known for all time. This exactly matches with Theorem 2 since $|\partial \mathbf{x} / \partial \mathbf{x}_0| = \zeta = 1$ and at any time $t$ the PDF is given by $\zeta f(\mathbf{x}_0, t_0) = f(\mathbf{x}_0, t_0)$.

## Discussion

The results presented in this paper show that in general, transformations of admissible region probabilities are only possible under strict conditions outlined by Theorem 2. Notable acceptable transformations include linear transformations and transformations with constant Jacobians over the admissible region. If a nonlinear transformation is applied to an admissible region PDF that does not satisfy Theorem 2, then the resulting PDF is necessarily a mis-representation of the true PDF. Furthermore, if a filter is instantiated from this improperly transformed PDF then it may cause unnecessary inefficiency in filter convergence. However, once every state in the admissible region becomes observable then Theorem 1 can be applied to transform the PDF with appropriate $\tilde{\mathbf{x}} = \mathbf{g}(\mathbf{x})$as desired. As such, for any filter to be properly instantiated, it must remain in the state space of the original admissible region PDF formulation unless either Theorem 2 or Corollary 3 is satisfied.

## SIMULATION AND RESULTS

To demonstrate probability transformations as applied to admissible regions, consider the observation of an object in LEO from an observer in Socorro, NM. Following the approach described in [13], the measurement vector is given by,

$$\mathbf{y} = \begin{bmatrix} \alpha & \delta & \dot{\alpha} & \dot{\delta} \end{bmatrix}^T \tag{60}$$

with the object state vector,

$$\mathbf{x} = \begin{bmatrix} \mathbf{r} & \mathbf{v} \end{bmatrix} \tag{61}$$

where **r** and **v** are position and velocity of the space object. The state matrix may also be represented by the topocentric spherical coordinates,

$$\tilde{\mathbf{x}} = \begin{bmatrix} \alpha & \delta & \dot{\alpha} & \dot{\delta} & \rho & \dot{\rho} \end{bmatrix}^T \tag{62}$$

For this observation model the undetermined states are given by $\tilde{\mathbf{x}}_u = [\rho\ \dot{\rho}]$, where $\rho$ is the range and $\dot{\rho}$ is the range-rate. The true state of the object at time $t_0$ is given in canonical units as

$$\mathbf{r} = \begin{bmatrix} -0.9281 \\ -0.0489 \\ 0.6167 \end{bmatrix} \mathrm{DU} \qquad \mathbf{v} = \begin{bmatrix} -0.5171 \\ 0.1292 \\ -0.7662 \end{bmatrix} \mathrm{DU/TU} \tag{63}$$

where 1 $DU = 6378$ km and 1 $DU/TU = 7.90538$ km/s. An initial series of 2 measurements of the inertial bearings are gathered at 20 second intervals producing the following determined states, or observation, vector

$$\mathbf{x}_d = \begin{bmatrix} -3.0337\,\mathrm{rad} & -0.0538\,\mathrm{rad} & -0.1003\,\mathrm{rad/TU} & -0.4482\,\mathrm{rad/TU} \end{bmatrix} \tag{64}$$

From this information an admissible region can be constructed. The admissible region is then constructed such that the constraint hypotheses give a region where 10000 km $\leq a \leq 50000$ km and $e < 0.4$. A set of 4000 points are uniformly sampled from the admissible region to demonstrate the requirements on admissible region transformations and are shown in Figure 1. The upper bound on semi major axis is given by the solid line and the upper bound on eccentricity is given by the dotted line in Figure 1.

Initial orbit determination methods can then use these sampled points to initiate particle filters or multiple hypothesis filters to process new observations. For these particle filter methods, the state vector can be converted to cartesian coordinates for propagation. However, this involves a transformation of the state space which implies either Theorem 2 or Eqn. (3) must be applied. The transformation from $\tilde{\mathbf{x}}$ to $\mathbf{x}$ is given by,

$$\mathbf{r} = \mathbf{o} + \rho\hat{\mathbf{l}} \tag{65}$$

$$\mathbf{v} = \dot{\mathbf{o}} + \dot{\rho}\hat{\mathbf{l}} + \rho\dot{\alpha}\hat{\mathbf{l}}_\alpha + \rho\dot{\delta}\hat{\mathbf{l}}_\delta \tag{66}$$

where,

$$\begin{aligned} \hat{\mathbf{l}}^T &= \begin{bmatrix} \cos\alpha\cos\delta & \sin\alpha\cos\delta & \sin\delta \end{bmatrix} \\ \hat{\mathbf{l}}_\alpha^T &= \begin{bmatrix} -\sin\alpha\cos\delta & \cos\alpha\cos\delta & 0 \end{bmatrix} \\ \hat{\mathbf{l}}_\delta^T &= \begin{bmatrix} \cos\alpha\sin\delta & -\sin\alpha\sin\delta & \cos\delta \end{bmatrix} \end{aligned}$$

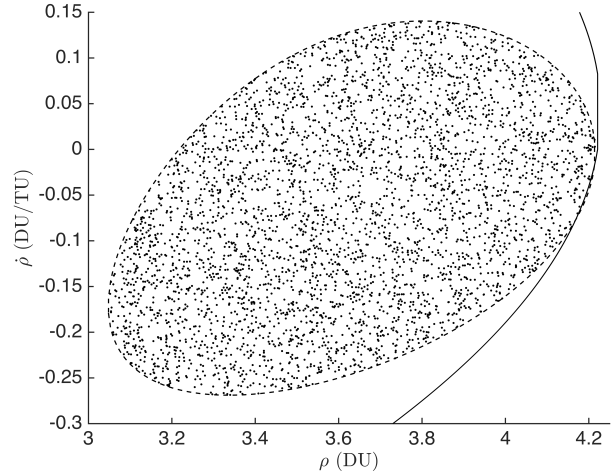


**Figure 1. A set of 4000 points sampled uniformly from the admissible region.**

and $\mathbf{o} \in \mathbb{R}^3$ is the observer position and $\dot{\mathbf{o}} \in \mathbb{R}^3$ is the observer velocity. This transformation is both one-to-one and onto as there is only one cartesian state corresponding to a given $\rho$, $\dot{\rho}$, and observation vector. The Jacobian of this transformation is clearly a function of $\rho$ and $\dot{\rho}$ and thus cannot be constant over the admissible region. After a single observation, the admissible region must still be expressed as a uniform distribution and transforming the sampled points into cartesian coordinates and expressing the PDF of the admissible region in cartesian coordinates violates Theorem 2. To demonstrate this, Figure 2 shows the values of the determinant of the Jacobian over the admissible region. Since the probability transformation of an admissible region requires this value to be constant, it is clear that the transformation to cartesian coordinates violates Theorem 2.

With a single observation and no consideration of uncertainty, each of the points sampled from the admissible region necessarily has a uniform spatial distribution. New measurements should allow the admissible region to become observable by taking into account the new information provided by the measurements. Once the system is observable, by Corollary 3, the admissible region PDF becomes a true PDF and the transformation is given directly by Eqn. (3). To test for observability, the condition number, $K(\mathbf{P}(t, t_0, \mathbf{x}(t)))$, for the observability gramian is computed for each value of $\rho$ and $\dot{\rho}$ shown in Figure 1. The inverse of the machine epsilon value $\delta_m^{-1}$ is also plotted, which indicates that any $K(\mathbf{P}(t, t_0, \mathbf{x}(t)) > \delta_m^{-1}$ is essentially infinity due to the precision of the computer. Then an additional observation is made 120 seconds after the initial set of observations. The additional observations are ingested by the particle filter and the updated observability gramian is computed. Figure 4 shows how the condition numbers for the observability gramian for each particle changes after the second observation is made. This change in condition number implies that the observability gramian becomes full rank after a second observation is made. At this point it is possible to transform the PDF expressed in terms of $\rho$ and $\dot{\rho}$ into cartesian coordinates by direct application of Eqn. (3). Figure 5 shows the updated PDF after the second observation is made and can equivalently be expressed in cartesian coordinates by Eqn. (3).

To demonstrate the importance of Theorem 2 and Corollary 3, consider the process shown in Figure 3 by which the cartesian PDF for these observations can be determined. The original admissible region in $\rho$ and $\dot{\rho}$ is represented by $\widetilde{\mathcal{AR}}_{t_0}$ and after the second observation is made the PDF over

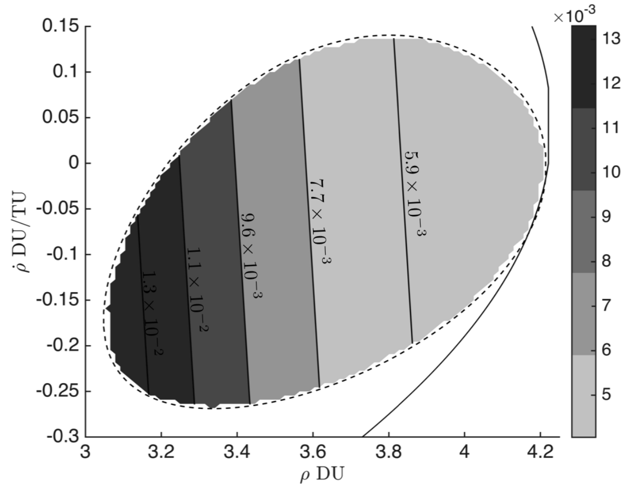


**Figure 2. Values of $|\partial \mathbf{x}/\partial \tilde{\mathbf{x}}|$ evaluated for each particle $\mathbf{x}(t)$**

the particles is given by $f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u)$. The admissible region given by $\mathcal{AR}$ represents the transformation of $\widetilde{\mathcal{AR}}_{t_0}$ while the system is still unobservable. It has already been shown that this particular transformation does not satisfy Theorem 2, thus it is expected that the resulting PDF in cartesian space given by $f_{\mathbf{X}_u}(\mathbf{x}_u)$ will not be equal to the transformation of $f_{\tilde{\mathbf{X}}_u}(\tilde{\mathbf{x}}_u)$ into cartesian coordinates once the system is observable. This subtle difference in approach will generate two different PDFs for the particles resulting from the second observation and mathematically the PDF generated from the unobservable transformation is incorrect. Figure 6 shows the resulting PDF for the unobservable and observable transformations outlined in Figure 3. The PDFs shown are represented as histograms of the particles for each cartesian state after the resampling step in the particle filter. As can be seen the PDFs are fairly different between the approaches. In general, once a particle filter is instantiated in a given state space using an admissible region, the PDF must remain expressed in that state space until the system is observable. The general exception to this are linear transformations which always satisfy the requirements of Theorem 2.

## CONCLUSIONS

The general theory of probability transformations is presented and then applied directly to the admissible region problem. It is found that general probability transformations are invalid for admissible regions, thus a constraint on transformations for admissible region problems is defined. The constraint is shown to ensure the admissible region remains a uniform distribution regardless of the state space it is expressed in. Furthermore it is shown that this requirement is only necessary while the system is unobservable, as once the system becomes observable the admissible region becomes a true PDF. It is also shown that probability transformations of admissible regions can also take into account measurement, parameter, and observer uncertainties.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. The authors would also like to acknowledge

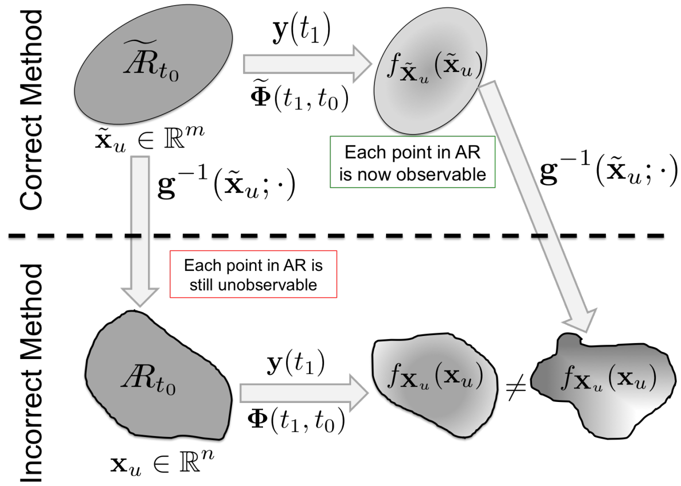


**Figure 3. Outline of the two approaches for generating the PDF in cartesian coordinates**

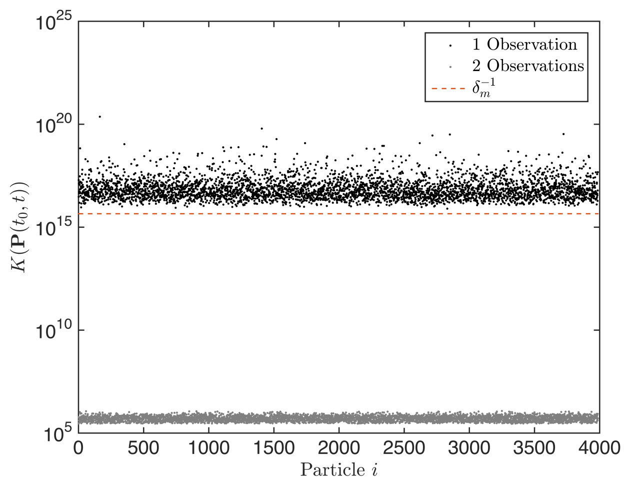


**Figure 4. Condition number of $\mathbf{P}(t, t_0, \mathbf{x}(t))$ computed for each particle $\mathbf{x}_u(t)$**

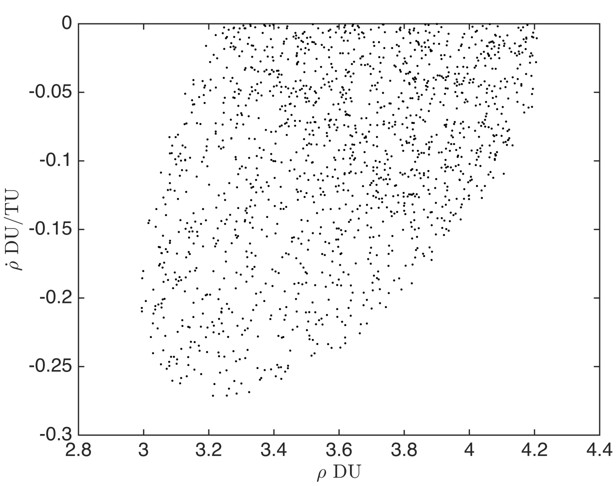


**Figure 5. PDF expressed in $\rho$ and $\dot{\rho}$ after the second observation**

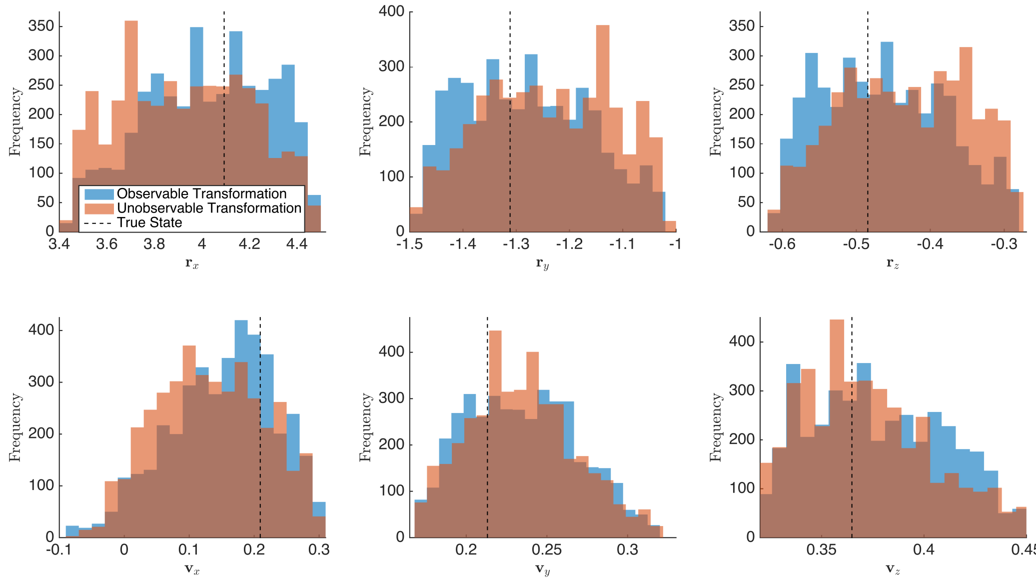


**Figure 6. Difference between the cartesian PDFs if the transformation from $\rho$ and $\dot{\rho}$ is applied before the system is observable.**

Islam Hussein, Christopher Roscoe, Matthew Wilkins, and Paul Schumacher for their discussions regarding this work.

## REFERENCES

[1] J.-C. Liou, "Modeling the Large and Small Orbital Debris Populations for Environment Remediation," tech. rep., NASA Orbital Debris Program Office, June 2014.

[2] S. A. Hildreth and A. Arnold, "Threats to US National Security Interests in Space: Orbital Debris Mitigation and Removal," DTIC Document, 2014.

[3] L. G. Taff, "On initial orbit determination," *Astronomical Journal*, Vol. 89, Sept. 1984, pp. 1426–1428, doi:10.1086/113644.

[4] A. Milani, G. F. Gronchi, M. d. Vitturi, and Z. Knežević, "Orbit determination with very short arcs. I admissible regions," *Celestial Mechanics and Dynamical Astronomy*, Vol. 90, No. 1-2, 2004, pp. 57–85, doi:10.1007/s10569-004-6593-5.

[5] A. Milani, G. F. Gronchi, Z. Knežević, M. E. Sansaturio, and O. Arratia, "Orbit determination with very short arcs: II. Identifications," *Icarus*, Vol. 179, No. 2, 2005, pp. 350 – 374, doi:10.1016/j.icarus.2005.07.004.

[6] J. M. Maruskin, D. J. Scheeres, and K. T. Alfriend, "Correlation of Optical Observations of Objects in Earth Orbit," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 1, 2009, pp. 194–209, doi:10.2514/1.36398.

[7] K. Fujimoto, J. Maruskin, and D. Scheeres, "Circular and zero-inclination solutions for optical observations of Earth-orbiting objects," *Celestial Mechanics and Dynamical Astronomy*, Vol. 106, No. 2, 2010, pp. 157–182, doi:10.1007/s10569-009-9245-y.

[8] K. DeMars, M. Jah, and P. Schumacher, "Initial Orbit Determination using Short-Arc Angle and Angle Rate Data," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 48, No. 3, 2012, pp. 2628–2637, doi:10.1109/TAES.2012.6237613.

[9] J. Siminski, O. Montenbruck, H. Fiedler, and T. Schildknecht, "Short-arc tracklet association for geostationary objects," *Advances in Space Research*, Vol. 53, No. 8, 2014, pp. 1184 – 1194, doi:10.1016/j.asr.2014.01.017.

[10] K. Fujimoto and K. T. Alfriend, "Optical Short-Arc Association Hypothesis Gating via Angle-Rate Information," *Journal of Guidance, Control, and Dynamics*, 2015/06/08 2015, pp. 1–12, 10.2514/1.G000927.

[11] K. Fujimoto, D. J. Scheeres, J. Herzog, and T. Schildknecht, "Association of optical tracklets from a geosynchronous belt survey via the direct Bayesian admissible region approach," *Advances in Space Research*, Vol. 53, No. 2, 2014, pp. 295 – 308, doi:10.1016/j.asr.2013.11.021.

[12] K. DeMars and M. K. Jah, "Probabilistic Initial Orbit Determination Using Gaussian Mixture Models," *Journal of Guidance Control, and Dynamics*, Vol. 36, No. 5, 2013, pp. 1324–1335, doi:10.2514/1.59844.

[13] J. L. Worthy and M. J. Holzinger, *Incorporating Uncertainty in Admissible Regions for Uncorrelated Detections*. American Institute of Aeronautics and Astronautics, 2015/02/17 2014, doi:10.2514/6.2014-4307.

[14] I. Hussein, C. Roscoe, M. Wilkins, and P. Schumacher, "Probabilistic Admissible Region for Short-Arc Angles-Only Observations," *Advanced Maui Optical and Space Surveillance Technologies Conference*, Vol. 1, 2014, p. 76.

[15] B. Flury, *A First Course in Multivariate Statistics*. Springer Texts in Statistics, Springer, 1997.

[16] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*. John Wiley & Sons, 5th ed., 2010.

[17] A. Chorin and O. Hald, *Stochastic Tools in Mathematics And Science*. Stochastic Tools in Mathematics and Science, Springer, 2006.

[18] R. Hogg, J. McKean, and A. Craig, *Introduction to Mathematical Statistics*. Pearson, 2013.

[19] G. Tommei, A. Milani, and A. Rossi, "Orbit determination of space debris: admissible regions," *Celestial Mechanics and Dynamical Astronomy*, Vol. 97, No. 4, 2007, pp. 289–304, doi:10.1007/s10569-007-9065-x.

[20] W. Brogan, "Modern Control Theory," 1991, p. 382.

[21] K. Fujimoto and D. J. Scheeres, "Correlation of Optical Observations of Earth-Orbiting Objects and Initial Orbit Determination," *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 1, 2012, pp. 208–221, doi: 10.2514/1.53126.

[22] R. S. Park and D. J. Scheeres, "Nonlinear Mapping of Gaussian Statistics: Theory and Applications to Spacecraft Trajectory Design," *Journal of Guidance, Control, and Dynamics*, Vol. 29, 2015/03/31 2006, pp. 1367–1375, 10.2514/1.20177.

[23] P. Maybeck, *Stochastic Models, Estimation, and Control*. Mathematics in Science and Engineering, Elsevier Science, 1982.

[24] K. Fujimoto and D. J. Scheeres, "Tractable Expressions for Nonlinearly Propagated Uncertainties," *Journal of Guidance, Control, and Dynamics*, Vol. 38, 2015/05/28 2015, pp. 1146–1151, 10.2514/1.G000795.

[25] D.-J. Lee and K. T. Alfriend, "Sigma Point Filtering for Sequential Orbit Estimation and Prediction," *Journal of Spacecraft and Rockets*, Vol. 44, 2015/05/11 2007, pp. 388–398, 10.2514/1.20702.

[26] S. Julier, "The scaled unscented transformation," *American Control Conference, 2002. Proceedings of the 2002*, Vol. 6, 2002, pp. 4555–4559 vol.6, 10.1109/ACC.2002.1025369.