Affine Diffusions with Non-Canonical State Space

Peter Spreij & Enno Veerman

Published online: 18 Jun 2012.

Submit your article to this journal ⬈

View related articles ⬈

Taylor & Francis
Taylor & Francis Group

Affine Diffusions with Non-Canonical State Space

PETER SPREIJ AND ENNO VEERMAN

Korteweg-de Vries Institute of Mathematics, University of Amsterdam, Amsterdam, The Netherlands

Multidimensional affine diffusions have been studied in detail for the case of a canonical state space. We present results for general state spaces and provide a complete characterization of all possible affine diffusions with polyhedral and quadratic state space. We give necessary and sufficient conditions on the behavior of drift and diffusion on the boundary of the state space in order to obtain invariance and to prove strong existence and uniqueness.

1. Introduction

Affine diffusions, introduced in the pioneering paper [10] by Duffie and Kan, are widely used in finance for modelling the term structure of interest rates. Their main attraction lies in the fact that they imply closed form expressions for bond prices. Affine diffusions are p-dimensional Markov processes that solve an *affine* stochastic differential equation (SDE) driven by a Brownian motion, i.e., an SDE with a drift $\mu(x)$ and diffusion matrix $\theta(x)$, both affine functions in the argument x. There are three important issues in the theory of affine diffusions, to wit

- *stochastic invariance* of a subset $\mathscr{X}$ of $\mathbb{R}^p$, the state space,
- the existence and uniqueness of strong solutions to the SDE with values in $\mathscr{X}$,
- the validity of the so-called *affine transform formula* for exponential moments.

Most of the theory that has recently been developed, concerns affine diffusions with a canonical state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^{p-m}$, henceforth referred to as the state space in *standard canonical form*, due to its tractable appearance which might ease the verification of possible technical conditions. The notion of a canonical state space has been introduced in [7]. Worth mentioning is the seminal paper [9] by Duffie, Filipović, and Schachermayer, who provide a complete characterization of (regular)

affine processes, allowing jumps as well, under the assumption of a standard canonical state space. Regarding the three issues mentioned above, for affine SDEs with canonical state space it is relatively easy to establish strong existence and uniqueness, as well as to derive conditions for invariance, see, for example, [7, 11]. Moreover, until recently, the affine transform formula has only been fully verified for affine diffusions with a standard canonical state space, for [11].

This article together with a companion paper [20] contribute to the theory of affine diffusions with a *non-canonical state space*. We will characterize all affine diffusions, focussing our attention on *polyhedrons* (of which the standard canonical state space is a special case) as well as *quadratic* state spaces (those of which the boundary is characterized by a quadratic function), though more is possible. For example, the matrix-valued affine processes and related Wishart processes treated in [5, 13] have the cone of positive semi-definite matrices as their state space. Worth mentioning is the Ph.D. thesis of Cuchiero [4], where affine processes are considered on general symmetric cones. Our results extend the classification of [12] for the two-dimensional case to higher dimensions. In [12] it is shown that besides an intersection of halfspaces, also a parabolic state space is possible. We will see that in higher dimensions the quadratic state spaces are not limited to the parabolic ones; there exist also affine diffusions whose state space is a Lorentz cone.

This article concerns the first two of three mentioned issues for affine diffusions that live on a non-canonical state space. Results on the third one are presented in the companion paper [20], where we extend the results in [11] on the validity of the affine transform formula for canonical to general state spaces. Returning to the first issue, in the current paper we derive conditions for the drift and diffusion matrix on the boundary of $\mathcal{X}$ to ensure stochastic invariance for both the polyhedral and the quadratic state space. For the standard canonical state space these conditions are often called *admissibility conditions*, see [7, 9]. The second issue, existence of a unique strong solution to an affine SDE, is in general not straightforward, as the square root of an affine matrix valued function θ is not locally Lipschitz continuous for singular θ. This article follows two approaches to solve this problem.

The first one is by invoking a result by Yamada and Watanabe [23, Theorem 1], as is done in [7, 11]. This result is essentially only applicable for the standard canonical state space. Under invariance conditions though, we prove that a general polyhedral state space can be transformed in some kind of canonical form, not necessarily the standard one, for which the result by Yamada and Watanabe does apply. For a parabolic state space unique strong solutions can be similarly obtained by application of an appropriate modification of this result.

The second way to obtain unique strong solutions is to impose conditions for invariance of $\{x \in \mathbb{R}^p : \theta(x) \text{ strictly positive definite}\}$ (also denoted by $\{\theta > 0\}$), an approach followed in [10] for affine diffusions with a diagonalizable diffusion matrix and in [18] for matrix-valued diffusions. Strong existence and uniqueness is guaranteed, as the unique positive definite square root of θ is locally Lipschitz continuous on $\{\theta > 0\}$. In this article we derive invariance conditions for general state spaces, following the arguments in [18]. This enables us to obtain existence and uniqueness of affine diffusions whose state space is a Lorentz cone.

As a side note, we mention that invariance of $\{\theta > 0\}$ is also important for applications. For example, in an affine term structure model one often desires an affine structure of the underlying SDE under both the risk-neutral and the physical measure. For this purpose, the invariance conditions for $\{\theta > 0\}$ are relevant in view

of [2, 8]. Therefore, we will provide these conditions not only for the Lorentz cone but also for the polyhedral and parabolic state space.

The remainder of this article is organized as follows. After introducing in Section 2 some notation and defining affine SDEs and diffusions more carefully, as well as presenting a more detailed description of our aims, we discuss in Section 3 stochastic invariance of the state space. Necessary conditions (admissibility conditions) on the behavior of the drift and the diffusion matrix on the boundary of a general closed convex state space are derived, whereas sufficient conditions are obtained for particular cases. In this section we also provide sufficient conditions for stochastic invariance of an open state space.

The invariance conditions derived in Section 3 are used in Section 4 and Section 5 to characterize all affine diffusions with polyhedral respectively quadratic state space. For the former we also give sufficient conditions, extending those known from the literature [3, 7, 11], under which the diffusion matrix can be diagonalized. In particular, we show that the classical model of [10] can be transformed into the canonical form of [7]. The results from convex analysis that we use in Section 4 are stated and proved in Appendix A. In Section 5, we show that for quadratic state spaces there are essentially only two types of state spaces possible, a (multidimensional) parabola and the Lorentz cone. For each of these types we are able to give a full characterization of the possible diffusion matrices.

2.   Definitions, Approach, and Notation

Let $p \in \mathbb{N}$. We are given a p-dimensional stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{2.1}$$

for continuous functions $\mu : \mathbb{R}^p \to \mathbb{R}^p$ and $\sigma : \mathbb{R}^p \to \mathbb{R}^{p \times p}$ that satisfy the linear growth condition

$$\|\mu(x)\| + \|\sigma(x)\| \leq M(1 + \|x\|), \quad \text{for all } x \in \mathbb{R}^p, \quad \text{some } M > 0. \tag{2.2}$$

By Theorems IV.2.3 and IV.2.4 in [15], for all initial conditions $x_0 \in \mathbb{R}^p$ there exists a *weak solution* (X, W) to (2.1), that is, there exists a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions, with a p-dimensional $\mathcal{F}_t$-Brownian motion W and an adapted p-dimensional stochastic process X, such that $X_0 = x_0$ a.s. and (2.1) holds. Let us also recall the result from Yamada and Watanabe [16, Theorem 21.14] that (2.1) has a *unique strong solution* if and only if weak existence and pathwise uniqueness holds.

We use the following definitions.

Definition 2.1.   We call a measurable set $\mathcal{X} \subset \mathbb{R}^p$ *stochastically invariant*, if for all $x_0 \in \mathcal{X}$, there exists a weak solution (X, W) to (2.1) with initial condition x_0 such that $X_t \in \mathcal{X}$ almost surely, for all $t \geq 0$.

Definition 2.2.   The SDE (2.1) is called an *affine SDE* with state space $\mathcal{X} \subset \mathbb{R}^p$ if

(1) it has a unique strong solution;
(2) $\mathcal{X}$ is stochastically invariant;

(3) the drift μ and diffusion matrix $\theta = \sigma\sigma^\top$ are affine in x on $\mathscr{X}$, i.e.

$$\mu(x) = ax + b, \quad \theta(x) = A^0 + \sum_{i=1}^{p} A^i x_i, \quad \text{for all } x \in \mathscr{X}, \tag{2.3}$$

for some $a \in \mathbb{R}^{p \times p}$, $b \in \mathbb{R}^p$, symmetric $A^i \in \mathbb{R}^{p \times p}$.

The unique strong solution to an affine SDE is called an affine diffusion, which corresponds to the definition in [11] in view of [20, Theorem 3.7].

Throughout the next sections we will address the following topics. Given a state space $\mathscr{X}$, we determine all affine functions $\mu : \mathbb{R}^p \to \mathbb{R}^p$ and $\theta : \mathbb{R}^{p \times p} \to \mathbb{R}^p$ such that there exists a continuous square root σ of θ on $\mathscr{X}$ for which (2.1) is an affine SDE. To that end, the following three aspects have to be taken into consideration.

- First, it is necessary that $\mathscr{X} \subset \{\theta \geq 0\}$, since $\theta(x) = \sigma(x)\sigma(x)^\top$ for $x \in \mathscr{X}$.
- Second, for stochastic invariance, one has to impose conditions on μ and θ on the boundary of $\mathscr{X}$, in order to prevent the solution X from leaving the state space $\mathscr{X}$.
- Third, one has to construct a square root σ such that (2.1) admits a unique strong solution that stays in $\mathscr{X}$. Remarkably, for the polyhedral and parabolic state space we consider, the conditions for stochastic invariance enable the construction of such a square root σ, see Theorems 4.3 and 5.7. For the Lorentz cone state space we obtain unique strong solutions by imposing conditions for invariance of $\{\theta > 0\}$, see Theorem 5.16.

Although obvious, it is worth noting that if X is an affine diffusion with drift $\mu(x)$, diffusion matrix $\theta(x)$ and state space $\mathscr{X}$, then $LX + \ell$ is an affine diffusion with drift $L\mu(L^{-1}(x - \ell))$, diffusion matrix $L\theta(L^{-1}(x - \ell))L^\top$ and state space $L\mathscr{X} + \ell$, for non singular $L \in \mathbb{R}^{p \times p}$, $\ell \in \mathbb{R}^p$. Therefore, for the tasks as outlined above, it suffices to characterize all affine diffusions where the state space is in a certain "canonical" form (not to be confused with the standard canonical form), thereby obtaining all remaining diffusions by affine transformations.

Remark 2.3. The distribution of X does not change with different choices of the square root σ (as long as X stays in $\mathscr{X}$ for these choices), since it is determined by the generator, which depends on σ only through $\sigma\sigma^\top$. Note, however, that if strong existence and uniqueness holds for one particular choice of σ, it does not automatically hold for other choices. For instance take $\theta(x) = 1$ in $\mathbb{R}$ and consider the one-dimensional SDE $dX_t = \sigma(X_t)dW_t$ with $\sigma\sigma^\top = \theta$. Existence and uniqueness of a strong solution holds when we take $\sigma(x) = 1$, while one only has a weak solution for the choice $\sigma(x) = \text{sgn}(x+)$, see [17, Example 5.3.5].

Matrix notation. The following notation regarding matrices and vectors is used throughout. Let $p, q \in \mathbb{N}$, $P = \{1, \ldots, p\}$, $Q = \{1, \ldots, q\}$, $A \in \mathbb{R}^{p \times q}$, $I \subset P$, $J \subset Q$. Write $I = \{i_1, \ldots, i_{\#I}\}$, $J = \{j_1, \ldots, j_{\#J}\}$, with $i_1 \leq i_2 \leq \cdots \leq i_{\#I}$ and $j_1 \leq j_2 \leq \cdots \leq j_{\#J}$. Then A_{IJ} denotes the ($\#I \times \#J$)-matrix with elements $(A_{IJ})_{kl} = A_{i_k j_l}$. If $\#I = 1$, say $I = \{i\}$, we write A_{iJ} instead. If $J = Q$ then we write A_I instead. In particular, A_i denotes the ith row of A. The jth column is denoted by A^j and the transpose of A is denoted by $A^\top$. The above notation is also used for matrix-valued functions ϕ, for example, $\phi^\top(x)$ stands for $(\phi(x))^\top$.

For $a_1, \ldots, a_p \in \mathbb{R}$ we write $\mathrm{diag}(a_1, \ldots, a_p)$ for the p-dimensional diagonal matrix D with diagonal elements $D_{ii} = a_i$, $i \in P$. We also write $\mathrm{diag}(a)$ instead, where a denotes the vector with elements a_i, sometimes explicitly denoted by $a = \mathrm{vec}(a_1, \ldots, a_p)$. For a vector v we write $|v|$ for the vector with elements $|v_i|$ and analogously $\sqrt{|v|}$ denotes the vector with elements $\sqrt{|v_i|}$. The identity matrix is written as I. We write $[A]$ for the linear span of the row vectors of A. If A and B are two matrices with the same column dimension, then $A \perp B$ stands for $[A] \perp [B]$, i.e. $A_i B_j^\top = 0$ for all i and j. The unique symmetric positive semi-definite square root of a positive semi-definite matrix A is denoted by $A^{1/2}$. For a square matrix A we write $|A|$ for $(AA^\top)^{1/2}$. We will often make use of the fact that the matrix-valued function $A \mapsto |A|^{1/2}$ is continuous, which is a consequence of [1, Theorem X.1.1].

3. Stochastic Invariance

In this section, we obtain necessary and in some cases sufficient boundary conditions for stochastic invariance, see Definition 2.1. We first consider stochastic invariance of a closed convex set $\mathscr{X} \subset \mathbb{R}^p$, for which we make use of the fact that it can be written as an intersection of halfspaces, that is,

$$\mathscr{X} = \bigcap_{i \in I} \{u_i \geq 0\}, \tag{3.1}$$

with I some index set and $u_i : \mathbb{R}^p \to \mathbb{R} : x \mapsto \gamma_i x + \delta_i$, for some $\gamma_i \in \mathbb{R}^{1 \times p}$, $\delta_i \in \mathbb{R}$. We denote the ith boundary segment $\mathscr{X} \cap \{u_i = 0\}$ of $\partial \mathscr{X}$ with $\partial \mathscr{X}_i$. The following proposition is partly proved in [11, Lemma B.1]. We give a more intuitive proof, involving an appropriate change of measure.

Proposition 3.1. *Let $\mathscr{X} \subset \mathbb{R}^p$ be a closed convex set given by (3.1) and assume $\mathscr{X}$ is stochastically invariant. Then necessarily it holds that*

$$\forall i \in I, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \sigma(x) = 0 \tag{3.2}$$

$$\forall i \in I, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \mu(x) \geq 0. \tag{3.3}$$

Proof. We give a proof by contradiction. Suppose there is an $i \in I$ and $x_0 \in \partial \mathscr{X}_i$ such that (3.2) or (3.3) does not hold. Then there exists $c \in \mathbb{R}$ such that η defined as $\eta(x) = \gamma_i \mu(x) + c \gamma_i \sigma(x) \sigma(x)^\top \gamma_i^\top$ is negative in x_0. Let (X, W) be a weak solution to (2.1) with initial condition x_0 on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$. There exists a stopping time $\tau_1 > 0$ such that $L = \mathscr{E}(c \gamma_i \sigma(X) 1_{[0,\tau_1]} \cdot W)$ is a martingale. Take $T > 0$ arbitrarily, then we can change $\mathbb{P}$ into an equivalent probability measure $\mathbb{Q}$ on $\mathscr{F}_T$ by $\mathrm{d}\mathbb{Q} = L_T \mathrm{d}\mathbb{P}$. By Girsanov's Theorem, $W^\mathbb{Q}$ defined by $\mathrm{d}W_t^\mathbb{Q} = \mathrm{d}W_t - c \sigma(X_t)^\top \gamma_i^\top 1_{[0,\tau_1]}(t) \mathrm{d}t$, is a Brownian motion under $\mathbb{Q}$ on $[0, T]$. Hence, X solves an SDE under $\mathbb{Q}$ for $t \in [0, T]$, namely

$$\mathrm{d}X_t = (\mu(X_t) + c \sigma(X_t) \sigma(X_t)^\top \gamma_i^\top 1_{[0,\tau_1]}(t)) \mathrm{d}t + \sigma(X_t) \mathrm{d}W_t^\mathbb{Q}.$$

Let $\tau_2 > 0$ be a stopping time such that $(\sigma(X) \cdot W^\mathbb{Q})_{t \wedge \tau_2}$ is a $\mathbb{Q}$-martingale and $\eta(X_t) < 0$ for $t \in [0, \tau_2]$, $\mathbb{Q}$-a.s. (this is possible since $\eta(X_0) = \eta(x_0) < 0$ and $t \mapsto X_t$

and η are continuous, $\mathbb{P}$-a.s., whence $\mathbb{Q}$-a.s.). Write $\tau = \tau_1 \wedge \tau_2$. Then it holds that $\tau > 0$, $\mathbb{Q}$-a.s. and, therefore,

$$E_{\mathbb{Q}} u_i(X_{T \wedge \tau}) = \delta_i + \gamma_i E_{\mathbb{Q}} X_{T \wedge \tau} = u_i(x_0) + E_{\mathbb{Q}} \int_0^{T \wedge \tau} \eta(X_t) \mathrm{d}t < 0.$$

This implies $\mathbb{Q}(\forall t \in [0, T] : u_i(X_t) \geq 0) < 1$ and by equivalence of $\mathbb{P}$ and $\mathbb{Q}$ also $\mathbb{P}(\forall t \in [0, T] : u_i(X_t) \geq 0) < 1$, which contradicts the stochastic invariance of $\mathscr{X}$. $\square$

If $\mathscr{X} \subset \mathbb{R}^p$ given by (3.1) is a convex polyhedron, then I is finite, say $I = \{1, \ldots, q\}$, for some $q \in \mathbb{N}$. We write $\gamma \in \mathbb{R}^{q \times p}$ for the matrix with row vectors γ_i, $\delta \in \mathbb{R}^q$ for the vector with elements δ_i and $u : \mathbb{R}^p \to \mathbb{R}^q : x \mapsto \gamma x + \delta$. The necessary conditions obtained in Proposition 3.1 are sufficient when $\mathscr{X}$ is a convex polyhedron congruent to the canonical space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^{p-m}$.

Proposition 3.2. *Let $\mathscr{X} \subset \mathbb{R}^p$ be a convex polyhedron given by (3.1) with $I = \{1, \ldots, q\}$ for some $q \in \mathbb{N}$ and suppose γ has full row-rank. Then $\mathscr{X}$ is stochastically invariant if and only if (3.2) and (3.3) hold.*

Proof. This follows by [21, Remark 3.6], since $\mathscr{X}$ is congruent to the standard canonical state space, as γ has full row-rank. $\square$

Proposition 3.2 has the following corollary, which will turn out to be the building block for proving stochastic invariance for affine diffusions with general polyhedral state space, as considered in Section 4.

Corollary 3.3. *Let $a \in \mathbb{R}^{p \times p}$, $b \in \mathbb{R}^p$, $c \in \mathbb{R}^p_{\geq 0}$ and suppose the drift of (2.1) is given by*

$$\mu(x) = ax + b, \tag{3.4}$$

and the diffusion coefficient by $\sigma(x) = diag(\sqrt{c_1|x_1|}, \ldots, \sqrt{c_p|x_p|})$, for all $x \in \mathbb{R}^p$. Then $\mathbb{R}^p_{\geq 0}$ is stochastically invariant if and only if

$$\forall i \in P, \quad \forall j \in P \setminus \{i\} : a_{ij} \geq 0 \quad and \quad b_i \geq 0. \tag{3.5}$$

To obtain conditions for stochastic invariance of a quadratic state space, we note that if $\mathscr{X} \subset \mathbb{R}^p$ is a closed convex set with $\partial \mathscr{X} \subset \{\Phi = 0\}$ for some C^1-function $\Phi : \mathbb{R}^p \to \mathbb{R}$, then we can take (3.1) with $I = \partial \mathscr{X}$, $\gamma_i = \nabla \Phi(i)$ and $\delta_i = -\gamma_i i$, for $i \in I$, where we write $\nabla \Phi$ for the gradient of Φ (written as a row vector). The necessary conditions for invariance (3.2) and (3.3) in this case read

$$\forall x \in \partial \mathscr{X} : \nabla \Phi(x) \sigma(x) = 0$$

$$\forall x \in \partial \mathscr{X} : \nabla \Phi(x) \mu(x) \geq 0.$$

The next proposition gives necessary and sufficient conditions for a particular case. We write $\nabla^2 \Phi = \nabla^\top \nabla \Phi$ for the Hessian of a C^2-function Φ.

Proposition 3.4. *Let* $\Phi : \mathbb{R}^p \to \mathbb{R} : x \mapsto x_1 - f(x_{P \setminus \{1\}})$ *for some* C^2-function $f :$ $\mathbb{R}^{p-1} \to \mathbb{R}$. *Then* $\mathscr{X} = \{\Phi \geq 0\}$ *is invariant for* (2.1) *if and only if*

$$\nabla\Phi(x)\sigma(x) = 0 \tag{3.6}$$

$$\nabla\Phi(x)\mu(x) \geq -\frac{1}{2}tr(\nabla^2\Phi(x)\sigma(x)\sigma(x)^\top), \tag{3.7}$$

for all $x \in \partial\mathscr{X} = \{\Phi = 0\}$.

Proof. Let (X, W) be a weak solution to (2.1) on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ with initial condition $x_0 \in \mathscr{X}$. It holds that $X_t \in \mathscr{X}$ if and only if $(\Phi(X_t), X_{P \setminus \{1\}, t}) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^{p-1}$. Itô's formula gives

$$d\Phi(X_t) = \nabla\Phi(X_t)dX_t + \frac{1}{2}tr(\nabla^2\Phi(X_t)d\langle X_t \rangle)$$

$$= \left(\nabla\Phi(X_t)\mu(X_t) + \frac{1}{2}tr(\nabla^2\Phi(X_t)\sigma(X_t)\sigma(X_t)^\top)\right)dt$$

$$+ \nabla\Phi(x)\sigma(X_t)dW_t.$$

Applying Proposition 3.2 gives the result. $\qquad\square$

Note that if $\mathscr{X}$ is convex in Proposition 3.4, equivalently f is a convex function, then $\nabla^2 f$ is positive semi-definite on its domain, when $\nabla^2\Phi$ is negative semi-definite and it follows that

$$tr(\nabla^2\Phi(X_t)\sigma(x)\sigma(x)^\top) = tr(\sigma(x)^\top\nabla^2\Phi(x)\sigma(x)) \leq 0, \quad \text{for all } x \in \mathbb{R}^p.$$

Thus, condition (3.7) is stronger than (3.2), whence (3.2) and (3.3) are in general not sufficient for stochastic invariance. For more results we refer to [6, 19].

We conclude this section by giving sufficient conditions for stochastic invariance of an open set. The idea behind the proof of the following result is taken from [18].

Proposition 3.5. *Let* $\Phi : \mathbb{R}^p \to \mathbb{R}$ *be a* C^2-function *and suppose* $\mathscr{X}^\circ$ *is a connected component* (*i.e. maximal connected subset*) *of* $\{\Phi > 0\}$. *Then* $\mathscr{X}^\circ$ *is invariant for* (2.1) *if there exists an open neighborhood* O *of* $\partial\mathscr{X}$ *such that*

$$\nabla\Phi(x)\mu(x) \geq -\frac{1}{2}tr(\nabla^2\Phi(x)\theta(x)) + \frac{1}{2}\Phi(x)^{-1}\nabla\Phi(x)\theta(x)(\nabla\Phi(x))^\top, \tag{3.8}$$

for all x *in* $O \cap \mathscr{X}^\circ$, *where we write* $\theta = \sigma\sigma^\top$.

Proof. Let (X, W) be a weak solution to (2.1) on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ with initial condition $x_0 \in \mathscr{X}^\circ$. By Itô's formula it holds for $t < \tau_0 :=$ $\inf\{s \geq 0 : \Phi(X_s) = 0\}$ that

$$d\log\Phi(X_t) = \Phi(X_t)^{-1}d\Phi(X_t) - \frac{1}{2}\Phi(X_t)^{-2}d\langle\Phi(X)\rangle_t$$

$$= \Phi(X_t)^{-1}(f(X_t)dt + \sigma(X_t)dW_t),$$

where

$$f(x) = \nabla\Phi(x)\mu(x) + \frac{1}{2}(\text{tr}(\nabla^2\Phi(x)\theta(x)) - \Phi(x)^{-1}\nabla\Phi(x)\theta(x)(\nabla\Phi(x))^\top).$$

Suppose $A := \{\tau_0 < \infty\}$ has positive probability. Since X does not explode, it holds for each $\omega \in A$ that there exists $\varepsilon > 0$ such that $X_t(\omega) \in O \cap \mathscr{X}^\circ$ for $\tau_0(\omega) - \varepsilon < t < \tau_0(\omega)$, whence $f(X_t(\omega)) \geq 0$. Therefore, $\int_0^t \Phi(X_s)^{-1}f(X_s)\mathrm{d}s$ does not tend to $-\infty$ on A for $t \uparrow \tau_0$. This yields

$$\lim_{t\uparrow\tau_0} \int_0^t \Phi(X_s)^{-1}\sigma(X_s)\mathrm{d}W_s = -\infty,$$

which is impossible, by the arguments of [18, Proposition 3.4].   □

Remark 3.6. Note that (3.8) implies (3.6) and is stronger than (3.7), as $\theta = \sigma\sigma^\top$ is positive semi-definite.

Proposition 3.5 yields tractable conditions for stochastic invariance of an open set in case of an affine SDE.

Proposition 3.7. *Consider the situation of Proposition* 3.5 *and suppose μ and $\theta := \sigma\sigma^\top$ are affine functions given by* (2.3). *Then $\mathscr{X}^\circ$ is invariant for* (2.1) *if*

$$\nabla\Phi(x)\theta(x) = \Phi(x)v^\top, \quad \text{for some constant } v \in \mathbb{R}^p, \tag{3.9}$$

$$\nabla\Phi(x)(\mu(x) - \frac{1}{2}\sum_{i=1}^p (A^i)^i) \geq 0, \quad \text{for all } x \in \mathscr{X}, \tag{3.10}$$

where $(A^i)^i$ denotes the ith column of the matrix A^i.

Proof. In view of Proposition 3.5 it suffices to show that

$$\text{tr}(\nabla^2\Phi(x)\theta(x)) - \Phi(x)^{-1}\nabla\Phi(x)\theta(x)(\nabla\Phi(x))^\top = -\nabla\Phi(x)\sum_{i=1}^p (A^i)^i,$$

under the assumption that $\nabla\Phi(x)\theta(x) = \Phi(x)v^\top$, for some constant vector $v \in \mathbb{R}^p$. Differentiating the right, respectively, the left-hand side of the latter and applying the product rule yields

$$\nabla^\top\Phi(x)v^\top = \nabla^\top(\nabla\Phi(x)\theta(x)) = \nabla^2\Phi(x)\theta(x) + \nabla\Phi(x)\nabla^\top\theta(x),$$

where $\nabla\Phi(x)\nabla^\top\theta(x)$ is short-hand notation for the matrix with row vectors $\nabla\Phi(x)\frac{\partial}{\partial x_i}\theta(x)$. Hence,

$$\text{tr}(\nabla^2\Phi(x)\theta(x)) = \nabla\Phi(x)v - \text{tr}(\nabla\Phi(x)\nabla^\top\theta(x)).$$

The result follows, since $\nabla\Phi(x)v = \Phi(x)^{-1}\nabla\Phi(x)\theta(x)(\nabla\Phi(x))^\top$ and

$$\text{tr}(\nabla\Phi(x)\nabla^\top\theta(x)) = \sum_{i=1}^p \nabla\Phi(x)\frac{\partial\theta^i(x)}{\partial x_i} = \nabla\Phi(x)\sum_{i=1}^p (A^i)^i.$$

                                                                            □

4. Polyhedral State Space

4.1. *General Diffusion Matrix*

Throughout this section, we assume $\mathscr{X}$ is a polyhedron given by

$$\mathscr{X} = \bigcap_{i=1}^{q} \{u_i \geq 0\}, \tag{4.1}$$

for some affine function $u : \mathbb{R}^q \to \mathbb{R}^p : x \mapsto \gamma x + \delta$, $\gamma \in \mathbb{R}^{q \times p}$, $\delta \in \mathbb{R}^q$, some $q \in \mathbb{N}$. We write $Q = \{1, \ldots, q\}$ and assume Q is *minimal* in the sense that $\bigcap_{i \in Q'} \{u_i \geq 0\} \neq \mathscr{X}$ for all $Q' \subset Q$ with $Q' \neq Q$. In addition we assume μ is of the form (3.4) and we are given an affine function θ by

$$\theta : \mathbb{R}^p \to \mathbb{R}^{p \times p} : x \mapsto A^0 + \sum_{i=1}^{p} A^i x_i, \tag{4.2}$$

for some symmetric $A^i \in \mathbb{R}^{p \times p}$ and we assume

$$\emptyset \neq \mathscr{X}^\circ \subset \mathscr{X} \subset \{x \in \mathbb{R}^p : \theta(x) \geq 0\} =: \mathscr{D}. \tag{4.3}$$

We will often make use of the fact that for a symmetric matrix S and vector v it holds that $Sv = 0$ is equivalent with $v^\top S v = 0$. In particular if $\theta = \sigma \sigma^\top$ on $\mathscr{X}$, then (3.2) is equivalent to

$$\forall i \in I, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \theta(x) = 0.$$

In Section 3, we have derived necessary boundary conditions on μ and θ to have $\mathscr{X}$ stochastically invariant. We show in this subsection that for a polyhedral state space, these conditions are also sufficient for the existence of a square root σ of θ on $\mathscr{X}$ such that (2.1) is an affine SDE. In the next proposition we prove that θ can be transformed in a block-diagonal form, which we use in the proof of Theorem 4.3 to construct the square root σ. The proposition improves upon the results given in the appendix of [10] and generalizes [11, Lemma 7.1].

Proposition 4.1. *Let $\mathscr{X} \subset \mathbb{R}^p$ be a convex polyhedron given by* (4.1) *and satisfying* (4.3) *and let θ be given by* (4.2). *Assume*

$$\forall i \in Q, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \theta(x) = 0. \tag{4.4}$$

Then there exists a non singular $L \in \mathbb{R}^{p \times p}$ and a vector $\ell \in \mathbb{R}^p$ such that

$$L\theta(L^{-1}(x - \ell))L^\top = \begin{pmatrix} diag(x_M, 0_N) & 0 \\ 0 & \Psi(x_{M \cup N}) \end{pmatrix}, \tag{4.5}$$

for some index sets $M = \{1, \ldots, m\}$, $N = \{m+1, \ldots, m+n\}$ and affine function Ψ. In addition, we have

$$L\mathscr{X} + \ell = \mathbb{R}_{\geq 0}^m \times \mathscr{C} \times \mathbb{R}^{p-m-n}, \tag{4.6}$$

for some convex polyhedron $\mathscr{C} = \bigcap_{i=1}^{q-m} \{\tilde{u}_i \geq 0\} \subset \mathbb{R}_{\geq 0}^n$ with $\tilde{u}_i(x) = x_i$ for $i \leq n$.

Proof. We divide the proof into a couple of steps.

Step 1. There exists $B \in \mathbb{R}^{q \times p}$ such that for $i \in Q$ it holds that

$$\gamma_i \theta(x) = B_i u_i(x), \tag{4.7}$$

$$B_i \gamma_i^\top > 0, \quad \text{if } B_i \neq 0 \tag{4.8}$$

$$B_i \gamma_j^\top = 0, \quad \text{for } j \in Q \backslash \{i\}. \tag{4.9}$$

This is shown as follows. Fix $i \in Q$. By (4.4) and Proposition A.4, there exists $B_i \in \mathbb{R}^{1 \times p}$ such that $\gamma_i \theta^j(x) = B_{ij} u_i(x)$ for $j \in P$. By assumption there exists $x_0 \in \mathscr{X}^\circ$. It holds that $u_i(x_0) > 0$, so we can write

$$B_i = u_i(x_0)^{-1} \gamma_i \theta(x_0). \tag{4.10}$$

By positive semi-definiteness of $\theta(x_0)$, it holds that $B_i \gamma_i^\top \geq 0$. We have $B_i \gamma_i^\top = 0$ if and only if $\gamma_i \theta(x_0) = 0$, that is, $B_i = 0$. This yields (4.8). Moreover, if $j \in Q \backslash \{i\}$, then by symmetry of θ it holds that

$$B_i \gamma_j^\top u_i(x) = \gamma_i \theta(x) \gamma_j^\top = \gamma_j \theta(x) \gamma_i^\top = B_j \gamma_i^\top u_j(x).$$

This implies $B_i \gamma_j^\top = 0$, since Q is minimal.

Step 2. It holds that

$$\operatorname{rank} \gamma_M = \operatorname{rank} B_M = m \tag{4.11}$$

$$\mathbb{R}^{1 \times p} = [\gamma_M] \oplus [B_M]^\perp, \tag{4.12}$$

where

$$M = \{i \in Q : B_i \neq 0\}. \tag{4.13}$$

Indeed, by (4.9) we have

$$B_M \gamma_M^\top = \operatorname{diag}(B_1 \gamma_1^\top, \ldots, B_m \gamma_m^\top), \tag{4.14}$$

which has full rank m by (4.8). This implies (4.11) as well as $[\gamma_M] \cap [B_M]^\perp = \{0\}$, which yields (4.12).

Step 3. It holds that

$$[B_M]^\perp = [\gamma_N] \oplus [\eta], \tag{4.15}$$

$$[\gamma_{Q \backslash M}] = [\gamma_N], \tag{4.16}$$

where $N \subset Q \backslash M$ is such that $\gamma_{M \cup N}$ has full rank, $\operatorname{rank} \gamma_{M \cup N} = \operatorname{rank} \gamma$, for some $\eta \in \mathbb{R}^{(p-m-n) \times p}$, with $n = \#N$. The first equality follows immediately from (4.11) and (4.12), while the second holds since $[\gamma_{Q \backslash M}] \subset [B_M]^\perp$ by (4.9) and $[\gamma_{M \cup N}] = [\gamma]$.

Step 4. Permute indices such that $M = \{1, \ldots, m\}$, $N = \{m + 1, \ldots, m + n\}$. We define $L \in \mathbb{R}^{p \times p}$ by

$$L_{M \cup N} = \gamma_{M \cup N} \tag{4.17}$$

$$L_{P \setminus (M \cup N)} = \eta. \tag{4.18}$$

Then L is non singular, by (4.12) and (4.15). Moreover, the previous steps yield

$$L\theta(x)L^{\top} = \begin{pmatrix} \text{diag}(c_1 u_1(x), \ldots, c_m u_m(x), 0_N) & 0 \\ 0 & \Phi(x) \end{pmatrix},$$

with $c_i = B_i \gamma_i^{\top}$ and Φ defined by $\Phi(x) = \eta \theta(x) \eta^{\top}$. It holds that $c_i > 0$ by (4.8). By rescaling u_i we may assume $c_i = 1$. Then we take $\ell \in \mathbb{R}^p$ such that $\ell_{M \cup N} = \delta_{M \cup N}$.

Step 5. It remains to show that we can write $\Phi(x)$ as an affine function of $u_{M \cup N}(x)$. This is an immediate consequence of the assumption that $\mathscr{X} \subset \{\theta \geq 0\}$, as this yields that $\theta(x)$ only depends on $u(x)$, which is a function of $u_{M \cup N}(x)$. $\square$

Remark 4.2. For the canonical state space as treated in [9, 11] one has (4.6) with $\mathscr{C} = \mathbb{R}^n_{\geq 0}$ and

$$\Psi(x_{M \cup N}) = B^0 + \sum_{i=1}^{m+n} B^i x_i,$$

with B^i positive semi-definite.

Theorem 4.3. *Let $\mathscr{X} \subset \mathbb{R}^p$ be given by (4.1) and satisfying (4.3). There exists an affine SDE with drift μ, diffusion matrix θ and polyhedral state space $\mathscr{X}$ if and only if*

$$\forall i \in Q, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \theta(x) = 0, \tag{4.19}$$

$$\forall i \in Q, \quad \forall x \in \partial \mathscr{X}_i : \gamma_i \mu(x) \geq 0. \tag{4.20}$$

Proof. The "only if" part is Proposition 3.1. We prove the "if"-part.

Assume (4.19) and (4.20). By Proposition 4.1 we may assume

$$\theta(x) = \begin{pmatrix} \text{diag}(x_M, 0_N) & 0 \\ 0 & \Psi(x_{M \cup N}) \end{pmatrix}, \text{ for } x \in \mathscr{X} \tag{4.21}$$

and

$$\mathscr{X} = \mathbb{R}^m_{\geq 0} \times \mathscr{C} \times \mathbb{R}^{p-m-n}, \tag{4.22}$$

for some index sets $M = \{1, \ldots, m\}$, $N = \{m + 1, \ldots, m + n\}$, an affine function Ψ and some convex polyhedron $\mathscr{C} \subset \mathbb{R}^n_{\geq 0}$, such that $u_{M \cup N}(x) = x_{M \cup N}$. As a square root of θ on $\mathscr{X}$ we take

$$\sigma(x) = |\theta(x)|^{1/2} = \begin{pmatrix} \text{diag}(\sqrt{|x_M|}, 0_N) & 0 \\ 0 & |\Psi(x_{M \cup N})|^{1/2} \end{pmatrix}. \tag{4.23}$$

It remains to show that there exists a unique strong solution X to (2.1) which stays in $\mathscr{X}$ for all $x_0 \in \mathscr{X}$.

Let $x_0 \in \mathbb{R}^p$ be arbitrary. By continuity of the coefficients μ and σ and satisfaction of the linear growth condition (2.2), there exists a weak solution (X, W) to (2.1) on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$. To show stochastic invariance of $\mathscr{X}$, note that by condition (4.20) and Proposition A.2 there exist $\tilde{a} \in \mathbb{R}^{q \times q}$ with $\tilde{a}_{ij} \geq 0$ for $i, j \in Q$ with $i \neq j$, and $\tilde{b} \in \mathbb{R}^q_{\geq 0}$ such that $\gamma \mu(x) = \tilde{a} u(x) + \tilde{b}$. Hence, $u(X_t)$ satisfies the q-dimensional SDE

$$\mathrm{d}u(X_t) = (\tilde{a} u(X_t) + \tilde{b})\mathrm{d}t + \mathrm{diag}\left(\sqrt{|X_{M,t}|}, 0_{Q \setminus M}\right)\mathrm{d}\widetilde{W}_t, \tag{4.24}$$

with initial condition $u(x_0)$, where $\widetilde{W}$ is a q-dimensional Brownian motion with $\widetilde{W}_M = W_M$ (possibly after extending Ω). For this SDE strong existence and uniqueness holds by [23, Theorem 1]. Together with Corollary 3.3 this yields $u(X_t) \in \mathbb{R}^q_{\geq 0}$, $\mathbb{P}$-a.s., for all $t \geq 0$ if $u(x_0) \in \mathbb{R}^q_{\geq 0}$. In other words, $X_t \in \mathscr{X}$, $\mathbb{P}$-a.s., for all $t \geq 0$ if $x_0 \in \mathscr{X}$. Thus $\mathscr{X}$ is stochastically invariant.

We now show pathwise uniqueness for (2.1). Therefore, let $(\widetilde{X}, W)$ be another solution on the same filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ with initial condition x_0. Pathwise uniqueness for (4.24) implies that $X_{M \cup N} = u_{M \cup N}(X) = u_{M \cup N}(\widetilde{X}) = \widetilde{X}_{M \cup N}$ a.s. Write $R = P \setminus (M \cup N)$. Since X_R does not appear in the diffusion part of the SDE, we have

$$\mathrm{d}(X_{R,t} - \widetilde{X}_{R,t}) = a_{RR}(X_{R,t} - \widetilde{X}_{R,t})\mathrm{d}t, \quad X_{R,0} - \widetilde{X}_{R,0} = 0.$$

So $X_R - \widetilde{X}_R$ solves a linear ODE which has 0 as its unique solution, whence $X_R = \widetilde{X}_R$ a.s. and the result follows. □

By an inspection of the proof of Theorem 4.3 we see that if X solves an affine SDE with polyhedral state space $\mathscr{X}$ given by (4.6), then $u(X)$ solves the affine SDE (4.24) with *admissible* parameters in the sense of [7] (that is, $\tilde{a}$ has non negative off-diagonal elements and $\tilde{b} \in \mathbb{R}^q_{\geq 0}$). Extending the dimension by considering $(u(X), X_R)$ instead of X, we get another affine SDE with state space $u(\mathscr{X}) \times \mathbb{R}^r$ (with $r = \#R$) and diffusion matrix $\tilde{\theta}$ (say). Now if θ can be written as an affine transformation of u with positive semi-definite matrices, then $\tilde{\theta}$ is positive semi-definite on the whole of $\mathbb{R}^q_{\geq 0} \times \mathbb{R}^r$. In that case, $u(\mathscr{X}) \times \mathbb{R}^r$ can be enlarged to the canonical state space $\mathbb{R}^q_{\geq 0} \times \mathbb{R}^r$ and the resulting SDE is of the canonical form as in [11]. We elaborate on this in the next subsection.

Remark 4.4. It is worth noting that in case an affine diffusion has a non-degenerate diffusion matrix, its polyhedral state space is necessarily (an affine transformation) of the canonical form $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^{p-m}$. Indeed, by (4.5) we have $N = \emptyset$, whence $n = 0$ and (4.6) reads $L\mathscr{X} + \ell = \mathbb{R}^m_{\geq} \times \mathbb{R}^{p-m}$.

Example 4.5. By considering degenerate diffusion matrices, one can construct affine diffusions with a polyhedral state space that is not (an affine transformation) of the canonical form. Here, we give an example of an affine diffusion with a degenerate but non-zero diffusion matrix, on a polyhedral *cone*, which is not a *symmetric* cone (and hence different from the cones in [14], so in particular of a non-canonical form).

Let $\mathscr{C} = \bigcap_{i=1}^{4}\{u_i \geq 0\} \subset \mathbb{R}^3_{\geq 0}$, with $u_1(x) = 2x_1 + 2x_2 + x_3$, $u_2(x) = -2x_1 - 2x_2 + x_3$, $u_3(x) = 2x_1 - 2x_2 + x_3$ and $u_4(x) = -2x_1 + 2x_2 + x_3$ and consider the 4-dimensional polyhedral cone $\mathscr{X} = \mathscr{C} \times \mathbb{R}$. Note that $\mathscr{X}$ is not a symmetric cone, as it is not self-dual. Let μ and θ be given as

$$\mu_N(x) = \begin{pmatrix} -x_1 \\ -x_2 \\ x_3 \end{pmatrix}, \quad \mu_{P \setminus N} \text{ arbitrary}, \quad \theta(x) = \begin{pmatrix} \operatorname{diag}(0_N) & 0 \\ 0 & x_3 \end{pmatrix},$$

with $N = \{1, 2, 3\}$. Then μ and θ fulfil (4.19) and (4.20), so by Theorem 4.3 there exists an affine diffusion with state space $\mathscr{X}$, drift μ and diffusion matrix θ.

4.2. *Diagonalizable Diffusion Matrix*

In [3], it is shown that the diffusion matrix $\theta(x)$ of an affine SDE with a standard canonical state space cannot be diagonalized in general. That is, there does not exist a non singular matrix L such that $L\theta(L^{-1}x)L^\top$ is diagonal. In this subsection we show that for a large class of polyhedral state spaces, including the standard canonical state space, diagonalization of the diffusion matrix is still possible in a different way, by extending the dimension. We also provide sufficient conditions for this as well as we give an example of an affine diffusion whose diffusion matrix is not diagonalizable.

Proposition 4.6. *Let X be an affine diffusion with drift μ, diffusion matrix θ and polyhedral state space $\mathscr{X}$. Then X is in distribution equal to an affine transformation of an affine diffusion with diagonal diffusion matrix and canonical state space of the form $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^{p-m}$ if and only if*

$$\theta(x) = B^0 + \sum_{i=1}^{q} B^i u_i(x), \quad \text{for some positive semi-definite } B^i. \tag{4.25}$$

Proof. We first prove the "only if" part. Suppose $X = L\widetilde{X} + \ell$ for some matrix L and vector ℓ, where $\widetilde{X}$ is an affine diffusion with diagonal diffusion matrix $\tilde{\theta}$ and polyhedral state space $\widetilde{\mathscr{X}}$. Then it holds that $L\tilde{\theta}(x)L^\top = \theta(Lx + \ell)$ and $\mathscr{X} = L\widetilde{\mathscr{X}} + \ell$. Since $\tilde{\theta}(x)$ is diagonal, we have

$$\theta(Lx + \ell) = L\tilde{\theta}(x)L^\top = \sum_i d_i(x)L^i(L^i)^\top, \tag{4.26}$$

for some affine real-valued functions d_i. Note that $L^i(L^i)^\top \geq 0$ and $d_i(x) \geq 0$ for $x \in \widetilde{\mathscr{X}}$. To show that θ is of the form (4.25), it suffices to write d_i as $d_i(x) = \sum_j \lambda_{ij} u_i(Lx + \ell) + c_i$ for non-negative λ_{ij} and c_i. Note that (4.26) yields $d_i(x + y) = d_i(x)$ for $y \in \ker L$ (otherwise, replace $d_i(x)$ by $d_i(\Pi(x))$, where Π denotes the projection onto $(\ker L)^\top$). Hence, the affine map

$$\phi_i : Lx + \ell \mapsto d_i(x)$$

is well defined and since $\phi_i(y) \geq 0$ for $y \in \mathscr{X} = L\widetilde{\mathscr{X}} + \ell$, Proposition A.1 yields the result.

Next we prove the "if" part. Suppose (4.25) holds. By Proposition 4.1 we may assume θ is of the form (4.21) with

$$\Psi(x_{M \cup N}) = \Lambda^0 + \sum_{i=1}^{q} \Lambda^i u_i(x),$$

for some positive semi-definite Λ^i. In this case $(u(X), X_{P \setminus (M \cup N)})$ is an affine diffusion where $u(X)$ satisfies (4.24). Therefore, we assume without loss of generality that $u(x) = x_{M \cup N}$, so that $Q = M \cup N$, $q = m + n$ and $\mathscr{X} = \mathbb{R}^q_{\geq 0} \times \mathbb{R}^{p-q}$. Since $\Lambda^i \geq 0$, its unique positive semi-definite square root $(\Lambda^i)^{1/2}$ exists. We write

$$\Lambda^{1/2} = \left((\Lambda^0)^{1/2} \quad (\Lambda^1)^{1/2} \quad \cdots \quad (\Lambda^q)^{1/2} \right)$$

$$w(x_Q) = \mathrm{vec}(1_Q, x_1 1_Q, \ldots, x_q 1_Q).$$

Now, note that

$$\Psi(x_{M \cup N}) = \Lambda^{1/2} \mathrm{diag}(w(x_Q))(\Lambda^{1/2})^\top,$$

so X is an affine diffusion with drift μ and diffusion matrix

$$\theta(x) = \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \Lambda^{1/2} \end{pmatrix} \begin{pmatrix} \mathrm{diag}(x_M, 0_N) & 0 \\ 0 & \mathrm{diag}(w(x_Q)) \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \Lambda^{1/2} \end{pmatrix}^\top.$$

We can diagonalize this by extending the dimension. Define a non singular square matrix T by

$$T = \begin{pmatrix} \Lambda^{1/2} & \mathrm{I} \\ \mathrm{I} & 0 \end{pmatrix},$$

and let $\widetilde{X}$ be an affine diffusion with drift $(\mu(x), 0)$ and diffusion matrix

$$\begin{pmatrix} \mathrm{I} & 0 \\ 0 & T \end{pmatrix} \begin{pmatrix} \mathrm{diag}(x_M, 0_N) & 0 \\ 0 & \mathrm{diag}(w(x_Q), 0_Q) \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ 0 & T \end{pmatrix}^\top,$$

and with the law of $\widetilde{X}_P$ equal to the law of X_0. Then

$$\begin{pmatrix} \mathrm{I} & 0 \\ 0 & T^{-1} \end{pmatrix} \widetilde{X}$$

solves an affine SDE with diagonal diffusion matrix $\mathrm{diag}(x_M, 0_N, w(x_Q), 0_Q)$ and it is easy to check that $\widetilde{X}_P$ satisfies an affine SDE with drift μ, diffusion matrix θ and initial condition the law of X_0. Hence, $\widetilde{X}_P$ is in distribution equal to X in view of Remark 2.3, which yields the result. $\qquad \square$

The relevance of diagonalizable diffusion matrices θ is elucidated in affine term structure models. In such models the short rate is an affine transformation of an affine diffusion X, the state factor. In view of Proposition 4.6, for an unobservable state factor X, we may assume without loss of generality that θ is diagonal when it is of the form (4.25). In particular this applies to affine diffusions in canonical

form. This observation reveals that affine diffusions with canonical state space and a nondiagonal diffusion matrix have the same potential as those with a diagonal diffusion matrix, which answers the implicit question in the concluding section of [3].

The following example shows that there exist affine diffusions with non-canonical polyhedral state space that do not meet the criteria of Proposition 4.6.

Example 4.7. Consider the polyhedron $\mathcal{X} = \mathcal{C} \times \mathbb{R}^2 \subset \mathbb{R}^4$, where we take $\mathcal{C} = \bigcap_{i=1}^3 \{u_i \geq 0\} \subset \mathbb{R}^2_{\geq 0}$, with $u_1(x) = x_1$, $u_2(x) = x_2$, $u_3(x) = x_1 + x_2 - \frac{3}{2}$. Let μ and θ be given as

$$\mu_N(x) = \begin{pmatrix} -x_1 + 1 \\ -x_2 + 1 \end{pmatrix}, \quad \mu_{P \setminus N} \text{ arbitrary}, \quad \theta(x) = \begin{pmatrix} \text{diag}(0_N) & 0 \\ 0 & \Phi(x) \end{pmatrix},$$

with $N = \{1, 2\}$ and

$$\Phi(x) = \begin{pmatrix} x_1 + \frac{1}{2} & 1 \\ 1 & x_2 + \frac{1}{2} \end{pmatrix}.$$

Then μ and θ fulfil (4.19) and (4.20), so by Theorem 4.3 there exists an affine diffusion with state space $\mathcal{X}$, drift μ and diffusion matrix θ. However, one can show that θ is not of the form (4.25).

We now give sufficient conditions for (4.25). In Proposition 4.8 below we prove that (4.25) does not only hold under full-row rankness of γ (equivalent to the canonical state space) but also under the weaker condition of full row-rankness of $\begin{pmatrix} \delta & \gamma \end{pmatrix}$ and an additional assumption.

Proposition 4.8. *Let $\mathcal{X} \subset \mathbb{R}^p$ be given by (4.1) and satisfying (4.3). Suppose that either*

 (i) *γ has full row-rank or that*
(ii) *$\begin{pmatrix} \delta & \gamma \end{pmatrix}$ has full row-rank and for all $i \in Q$, $x \in \mathbb{R}^p$ it holds that $u_j(x) = 0$ for all $j \in Q \setminus \{i\}$ implies $u_i(x) \geq 0$.*

Then (4.25) holds.

Proof. The case when γ has full row rank is easy, since the state space can be transformed in the canonical form $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^{p-m}$. We consider the case when γ has not full row rank. First we assume $\delta_i \geq 0$ for all $i \in Q$.

Let $\Gamma \in \mathbb{R}^{(p+1) \times (p+1)}$ be a non singular matrix such that $\Gamma_Q = \begin{pmatrix} \delta & \gamma \end{pmatrix}$. We extend the dimension of δ and γ by writing $\Gamma = \begin{pmatrix} \delta & \gamma \end{pmatrix}$. We also write $u_i(x) = \gamma_i x + \delta_i$, $x \in \mathbb{R}^p$, for all $i \leq p + 1$, in other words

$$u_i(x) = \Gamma_i \begin{pmatrix} 1 \\ x \end{pmatrix}.$$

Let $N = \Gamma^{-1}$. Then we have

$$\begin{pmatrix} 1 \\ x \end{pmatrix} = \Gamma^{-1} \Gamma \begin{pmatrix} 1 \\ x \end{pmatrix} = \sum_{i=1}^{p+1} N^i u_i(x), \quad \text{for } x \in \mathbb{R}^p.$$

Let $A = \begin{pmatrix} A^0 & A^1 & \dots & A^p \end{pmatrix}$ and abuse notation by writing

$$Ay = \sum_{i=1}^{p+1} A^{i-1} y_i, \quad \text{for } y \in \mathbb{R}^{p+1}.$$

Then we can write

$$\theta(x) = A \begin{pmatrix} 1 \\ x \end{pmatrix} = \sum_{i=1}^{p+1} (AN^i) u_i(x).$$

Define $\Lambda^i = AN^i$ for $i \le p+1$. It is sufficient to prove that $\Lambda^i \ge 0$ for $i \le q$ and $\Lambda^i = 0$ for $i > q$.

Let $i \le q$ and write $N^i = (N_1^i, N_2^i)$ with $N_1^i \in \mathbb{R}$, $N_2^i \in \mathbb{R}^p$. We consider two cases.

Case 1. Suppose $N_1^i \ne 0$. Then $N_2^i / N_1^i \in \{\theta \ge 0\} =: \mathscr{D}$. Indeed, for $j \in Q$, we have

$$u_j(N_2^i / N_1^i) = \frac{1}{N_1^i} \Gamma_j N^i = \frac{1}{N_1^i} 1_{\{i=j\}}.$$

So $u_j(N_2^i / N_1^i) = 0$ for $j \in Q \backslash \{i\}$, when by assumption $u_i(N_2^i / N_1^i) \ge 0$, that is, $N_1^i > 0$ and $N_2^i / N_1^i \in \mathscr{X} \subset \mathscr{D}$. It follows that

$$0 \le \theta(N_2^i / N_1^i) = \frac{1}{N_1^i} AN^i = \frac{1}{N_1^i} \Lambda^i,$$

and thus $\Lambda^i \ge 0$, as $N_1^i > 0$.

Case 2. Suppose $N_1^i = 0$. Then we have

$$\Lambda^i = AN^i = \lim_{\varepsilon \downarrow 0} \varepsilon A \begin{pmatrix} 1 \\ N_2^i / \varepsilon \end{pmatrix} = \lim_{\varepsilon \downarrow 0} \varepsilon \theta(N_2^i / \varepsilon).$$

So we have to prove $N_2^i / \varepsilon \in \mathscr{D}$ for ε small enough. For $j \in Q$ it holds that

$$1_{\{i=j\}} = \Gamma_j \begin{pmatrix} 0 \\ N_2^i \end{pmatrix} = \gamma_j N_2^i.$$

This gives

$$u_j(N_2^i / \varepsilon) = \delta_j + 1_{\{i=j\}} / \varepsilon \ge 0, \quad \text{for } j \in Q, \varepsilon > 0,$$

since $\delta_j \ge 0$ for $j \in Q$ by assumption. Hence, $N_2^i / \varepsilon \in \mathscr{X} \subset \mathscr{D}$ for $\varepsilon > 0$ and, thus, $\Lambda^i \ge 0$.

We have just shown that $\Lambda^i \ge 0$ for $i \le q$. We now show that $\Lambda^i = 0$ for $i > q$. Let $i > q$ and take $x_0 \in \mathscr{X}$. Write $R = P \cup \{p+1\}$. Since rank $\gamma = p$ and rank $\gamma_{R \backslash \{i\}} = p - 1$, there exists $\xi \in \mathbb{R}^p$ such that $\gamma_j \xi = 0$ for all $j \in R \backslash \{i\}$ and $\gamma_i \xi \ne 0$. Then for all

$k \in \mathbb{R}$ we have $u_j(x_0 + k\xi) = u_j(x_0)$ for $j \in R\backslash\{i\}$, in particular for all $j \in Q$. Hence, $x_0 + k\xi \in \mathscr{X}$ for all $k \in \mathbb{R}$. Therefore,

$$0 \le \theta(x_0 + k\xi) = \sum_{j=1}^{p+1} \Lambda^j u_j(x_0 + k\xi) = \theta(x_0) + k\Lambda^i \gamma_i \xi, \quad \text{for all } k \in \mathbb{R}.$$

Since $\gamma_i \xi \ne 0$, it follows that $\Lambda^i = 0$.

Now consider the general case without the restriction $\delta_i \ge 0$ for all $i \in Q$. Let $x_0 \in \mathscr{X}$ and for $i \in Q$ define

$$\tilde{u}_i : \mathbb{R}^p \to \mathbb{R} : x \mapsto \tilde{\gamma}_i x + \tilde{\delta}_i,$$

by $\tilde{u}_i(x) = u_i(x + x_0)$, $x \in \mathbb{R}^p$, i.e. $\tilde{\delta}_i = \delta_i + \gamma_i x_0$ and $\tilde{\gamma}_i = \gamma_i$. Moreover, write $\tilde{\theta}(x) = \theta(x + x_0)$, $x \in \mathbb{R}^p$. Then $\tilde{\delta}_i = \tilde{u}_i(0) = u_i(x_0) \ge 0$, $i \in Q$. Note that if $\tilde{u}_j(x) = 0$ for all $j \in Q\backslash\{i\}$, then $u_j(x + x_0) = 0$ for all $j \in Q\backslash\{i\}$ and hence $\tilde{u}_i(x) = u_i(x + x_0) \ge 0$ by assumption. Moreover,

$$\begin{pmatrix} \tilde{\delta} & \tilde{\gamma} \end{pmatrix} = \begin{pmatrix} \delta + \gamma x_0 & \gamma \end{pmatrix} = \begin{pmatrix} \delta & \gamma \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x_0 & \mathrm{I} \end{pmatrix},$$

which has full row rank. Therefore, we are in the previous situation, for which we have proved the existence of positive semi-definite $\Lambda^i \in \mathbb{R}^{p \times p}$ for $i \in Q \cup \{0\}$ such that

$$\theta(x + x_0) = \tilde{\theta}(x) = \Lambda^0 + \sum_{i=1}^q \Lambda^i \tilde{u}_i(x) = \Lambda^0 + \sum_{i=1}^q \Lambda^i u_i(x + x_0).$$

This gives the result. $\qquad\square$

Corollary 4.9. *Let $\mathscr{X} \subset \mathbb{R}^p$ be given by* (4.1). *Suppose $\mathscr{X}$ is contained in a polyhedron $\mathscr{C} \subset \{\theta \ge 0\}$ which meets the criteria of Proposition* 4.8. *Then* (4.25) *holds.*

Proof. Write $\mathscr{C} = \bigcap_{i=1}^r \{v_i \ge 0\}$, for some affine functions v_i, with $r \in \mathbb{N}$. By Proposition 4.8 it holds that

$$\theta(x) = B^0 + \sum_{i=1}^r B^i v_i(x), \tag{4.27}$$

for some positive semi-definite B^i. By Proposition A.1 there exist $\lambda_{ij} \ge 0$, $c_i \ge 0$ such that

$$v_i = \sum_{j=1}^r \lambda_{ij} u_j + c_i.$$

Plugging this in into (4.27) we get the result. $\qquad\square$

Example 4.10. In the two-dimensional case, the polyhedrons which satisfy the conditions of Proposition 4.8 are the "triangles" (including those with vertices and edges in "infinity," like $\{0 \le x_1 \le 1\} \cap \{x_2 \ge 0\}$ and $\{x_1 \ge 0\}$). Thus by Corollary 4.9,

if $\mathcal{X}$ is contained in a triangle that is a subset of $\{\theta \geq 0\}$, then (4.25) holds and $\mathcal{X}$ can be transformed in canonical form (see the proof of Proposition 4.6). However, this condition is sufficient, but not necessary. For example let

$$\theta(x) = \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} x_2,$$

and take $u_1(x) = 2x_1 - x_2$, $u_2(x) = -\frac{1}{2}x_1 + x_2$, $u_3(x) = -\frac{9}{4} + x_1 + x_2$. Then $\{\theta \geq 0\} = \{x \in \mathbb{R}^2_{\geq 0} : x_2 \geq \frac{1}{x_1}\}$ and $\mathcal{X} = \bigcap_{i=1}^{3}\{u_i \geq 0\} \subset \{\theta \geq 0\}$, but $\mathcal{X}$ is not contained in a triangle which is a subset of $\{\theta \geq 0\}$. Still we can write θ as an affine transformation of the u_i's with positive semi-definite coefficients, namely

$$\theta = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} u_1 + \frac{1}{9}\begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix} u_2 + \frac{1}{9}\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} u_3.$$

4.3. *Classical Model*

In this subsection, we revisit the classical model as introduced by Duffie and Kan in [10]. We assume θ is of the form

$$\theta = \Sigma \operatorname{diag}(v)\Sigma^{\top} = \sum_{i=1}^{p} \Sigma^i (\Sigma^i)^{\top} v_i, \tag{4.28}$$

with $\Sigma \in \mathbb{R}^{p \times p}$, $v : \mathbb{R}^p \to \mathbb{R}^p : x \mapsto \beta x + \alpha$ for some $\beta \in \mathbb{R}^{p \times p}$, $\alpha \in \mathbb{R}^p$. In addition we assume $\mathcal{X} \subset \bigcap_{i=1}^{p}\{v_i \geq 0\}$ and $\mathcal{X}^{\circ} \neq \emptyset$. Under conditions (4.19) and (4.20), the proof of Theorem 4.3 constructs a square root σ of θ on $\mathcal{X}$ such that (2.1) has a unique strong solution. We show that the natural choice

$$\sigma = \Sigma \operatorname{diag}(\sqrt{|v|}) = \begin{pmatrix} \Sigma^1 \sqrt{|v_1|} & \dots & \Sigma^p \sqrt{|v_p|} \end{pmatrix}$$

also gives strong existence and uniqueness for (2.1). This is not immediately clear in view of Remark 2.3.

Proposition 4.11. *Suppose θ is of the form* (4.28) *and assume* $\mathcal{X}^{\circ} \neq \emptyset$, $\mathcal{X} \subset \bigcap_{i=1}^{p}\{v_i \geq 0\}$ *and conditions* (4.19) *and* (4.20) *are met. Then* (2.1) *is an affine SDE for* $\sigma = \Sigma \operatorname{diag}(\sqrt{|v|})$.

Proof. Since conditions (4.19) and (4.20) are sufficient for invariance by Theorem 4.3, it suffices to prove existence and uniqueness of a strong solution.

Note that X solves an affine SDE with $\sigma(x) = \Sigma \operatorname{diag}(\sqrt{|v(x)|})$ if and only if $LX + \ell$ solves an affine SDE with $\sigma(x) = L\Sigma \operatorname{diag}(\sqrt{|\tilde{v}(x)|})$, where $\tilde{v}(x) = v(L^{-1}(x - \ell))$, for some non singular matrix L and vector ℓ. By Proposition 4.1 we can, therefore, assume without loss of generality that θ is of the form (4.21) and $\mathcal{X}$ of the form (4.22). Since all square roots of a positive semi-definite matrix are related by an orthogonal transformation, we have

$$\Sigma \operatorname{diag}(\sqrt{v(x)}) = \begin{pmatrix} \operatorname{diag}(\sqrt{x_M}, 0_N) & 0 \\ 0 & (\Psi(x_{M \cup N}))^{1/2} \end{pmatrix} O(x), \tag{4.29}$$

for $x \in \mathcal{X}$, with $O(x)$ an orthogonal matrix, possibly depending on x. We show that there exists a matrix $U \in \mathbb{R}^{(m+n) \times p}$ with orthonormal rows such that

$$\Sigma_{M \cup N} \operatorname{diag}(\sqrt{|v(x)|}) = \operatorname{diag}(\sqrt{|x_M|}, 0_N) U, \quad \text{for all } x \in \mathbb{R}^p.$$

Let $x_0 \in \mathcal{X}^\circ$ and define $U = O_{M \cup N}(x_0)$. We have to show that

$$\Sigma_{ij} \sqrt{|v_j(x)|} = \sqrt{|x_i|} U_{ij}, \quad \text{for all } i \in M, \quad \text{all } j. \tag{4.30}$$

For $i \in M$ we have

$$\Sigma_{ij} \sqrt{v_j(x)} = \sqrt{x_i} O_{ij}(x), \quad \text{for all } j \quad \text{and all } x \in \mathcal{X}.$$

If $\Sigma_{ij} \neq 0$, then

$$v_j(x) = \Sigma_{ij}^{-2} x_i O_{ij}^2(x), \quad \text{for all } x \in \mathcal{X},$$

which yields $v_j(x) = c x_i$ for some $c \geq 0$, by Proposition A.4. If $c > 0$, then $O_{ij}(x)$ is constant, when equal to U_{ij} and (4.30) follows. If $c = 0$ or $\Sigma_{ij} = 0$, then $O_{ij}(x) = 0$ for $x_i \neq 0$, whence $U_{ij} = 0$ and again (4.30) holds.

Let (X, W) be a weak solution to (2.1) on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Then UW is an $(m + n)$-dimensional Brownian motion and it follows that $X_{M \cup N}$ solves an SDE with diffusion part $\operatorname{diag}(\sqrt{|x_M|}, 0_N)$. The strong existence and uniqueness for (2.1) follows along the same lines as the proof of Theorem 4.3, as $\sigma(x)$ only depends on $x_{M \cup N}$ by Proposition A.1. $\qquad\square$

In the case that $\{\theta > 0\} \neq \emptyset$, we can strengthen the result of Proposition 4.6.

Proposition 4.12. *Consider the situation of Proposition 4.11. If $\{\theta > 0\} \neq \emptyset$, then the solution to (2.1) can be obtained by a bijective affine transformation of an affine diffusion with diagonal diffusion matrix and canonical state space of the form $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^{p-m}$.*

Proof. As in the proof of Proposition 4.11 we may assume θ is of the form (4.21) and $\mathcal{X}$ of the form (4.22). Since $\{\theta > 0\} \neq \emptyset$ we have $N = \emptyset$. Take $x_0 \in \mathcal{X}^\circ = \mathbb{R}_{>0}^m \times \mathbb{R}^{p-m}$ with $x_{0,i} = 1$ for $i \in M$. By strict positive definiteness of $\Psi(x_{0,M})$ we can apply the linear transformation

$$x \mapsto \begin{pmatrix} I_{MM} & 0 \\ 0 & (\Psi(x_{0,M}))^{-1/2} \end{pmatrix} x,$$

so that we may assume without loss of generality that $\Psi(x_{0,M}) = \mathrm{I}$, i.e. $\theta(x_0) = \mathrm{I}$. Since $v_i(x_0) > 0$ for all i, we can substitute $v_i/v_i(x_0)$ for v_i and $\sqrt{v_i(x_0)} \Sigma^i$ for Σ^i, which does not affect $\sigma(x)$ and yields $v_i(x_0) = 1$ for all i. Hence

$$\Sigma \Sigma^\top = \theta(x_0) = \mathrm{I},$$

in other words, Σ is orthogonal. Thus, we can write

$$\Sigma \operatorname{diag}(v(x)) = \begin{pmatrix} \operatorname{diag}(x_M) & 0 \\ 0 & \Psi(x_M) \end{pmatrix} \Sigma.$$

By symmetry, $\Psi(x_M)$ can be diagonalized by an orthogonal matrix $U(x_M)$ for all x_M. We show that $U(x_M)$ is constant. The above display yields

$$\Sigma_{P\backslash M}\,\mathrm{diag}(v(x)) = \Psi(x_M)\Sigma_{P\backslash M},$$

so for all x_M the eigenvectors of $\Psi(x_M)$ are in the span of the columns of $\Sigma_{P\backslash M}$, since rank $\Sigma_{P\backslash M} = p - m \geq$ rank $\Psi(x_M)$. Hence, the eigenvectors do not depend on x_M, which implies that $U(x_M)$ is constant, say equal to an orthogonal matrix O. Thus, $O^\top\Psi(x_M)O = \mathrm{diag}(w(x_M))$, for some affine vector-valued function w. Applying the orthogonal transformation

$$x \mapsto \begin{pmatrix} I_{MM} & 0 \\ 0 & O^\top \end{pmatrix} x,$$

we may assume $\Psi(x_M)$ is diagonal. Hence,

$$\Sigma\,\mathrm{diag}(v(x)) = \begin{pmatrix} \mathrm{diag}(x_M) & 0 \\ 0 & \mathrm{diag}(w(x_M)) \end{pmatrix}\Sigma,$$

with Σ orthogonal. It is easy to show that this yields

$$\Sigma\,\mathrm{diag}(\sqrt{|v(x)|}) = \begin{pmatrix} \mathrm{diag}(\sqrt{|x_M|}) & 0 \\ 0 & \mathrm{diag}(\sqrt{|w(x_M)|}) \end{pmatrix}\Sigma.$$

Since an orthogonal transformation of a Brownian motion is again a Brownian motion, the Σ on the right-hand side can be absorbed in the underlying Brownian motion. The result follows. $\qquad\square$

Using the above we are able to give an alternative proof of the existence and uniqueness results from [10], slightly strengthening the statements made there, see Remark 4.15 below. We first show how the well-known condition for invariance of the open state space $\mathbb{R}^p_{>0}$ follows from Proposition 3.7.

Proposition 4.13. *Let* $\theta(x) = diag(x)$ *and* $\mathscr{X} = \mathbb{R}^p_{\geq 0}$. *Then* $\mathscr{X}^\circ$ *is stochastically invariant if*

$$\forall i, \quad \forall j \neq i : a_{ij} \geq 0 \quad and \quad b_i \geq \frac{1}{2}.$$

Proof. We apply Proposition 3.7 with $\Phi(x) = \det\theta(x) = x_1 x_2 \ldots x_p$. It holds that

$$\nabla\Phi(x)\theta(x) = \Phi(x)\begin{pmatrix} 1 & \ldots & 1 \end{pmatrix},$$

and $\sum_{i=1}^p (A^i)^i) = \begin{pmatrix} 1 & \ldots & 1 \end{pmatrix}^\top$, when

$$\nabla\Phi(x)\left(\mu(x) - \frac{1}{2}\sum_{i=1}^p (A^i)^i\right) = \sum_i \prod_{j\neq i} x_j\left(a_i x + b_i - \frac{1}{2}\right).$$

This is non negative for all $x \in \mathscr{X} = \mathbb{R}^p_{\geq 0}$ if $a_{ij} \geq 0$ and $b_i \geq \frac{1}{2}$ for all i, j. By applying a measure transformation with density $\mathscr{E}(\lambda^\top\sigma(X) \cdot W)$ for some $\lambda \in \mathbb{R}^p$ (which yields

a probability measure by [20, Proposition 3.4]), we see that the sign of the diagonal elements a_{ii} is irrelevant for stochastic invariance of $\mathscr{X}^\circ$. □

Corollary 4.14. *Suppose θ is of the form* (4.28) *and assume $\mathscr{X}^\circ \neq \emptyset$, $\mathscr{X} \subset \bigcap_{i=1}^{p}\{v_i \geq 0\}$ and $\{\theta > 0\} \neq \emptyset$. Then* (2.1) *is an affine SDE for $\sigma = \Sigma\sqrt{|v|}$ if and only if, up to a reparametrization of Σ and v, $\mathscr{X} = \bigcap_{i=1}^{q}\{v_i \geq 0\} = \{\theta \geq 0\} =: \mathscr{D}$ for some $q \leq p$ and*

$$\forall i \leq q, \quad \forall j \leq p : \beta_i \Sigma^j = 0 \;\; or \;\; v_i = v_j, \tag{4.31}$$

$$\forall i \leq q, \quad \forall x \in \partial \mathscr{D}_i : \beta_i(ax + b) \geq 0. \tag{4.32}$$

Moreover, if we strengthen (4.32) *to*

$$\forall i \leq q, \forall x \in \partial \mathscr{D}_i : \beta_i(ax + b) \geq \frac{1}{2}\beta_i \Sigma \Sigma^\top \beta_i^\top, \tag{4.33}$$

then $\mathscr{D}^\circ = \bigcap_{i=1}^{q}\{v_i > 0\}$ is stochastically invariant.

Proof. Note that $\{\theta > 0\} \neq \emptyset$ implies Σ is non singular. Hence, $\theta \geq 0$ if and only if $\Sigma^{-1}\theta(\Sigma^{-1})^\top = \mathrm{diag}(v) \geq 0$, so

$$\mathscr{D} = \bigcap_{i=1}^{p}\{v_i \geq 0\}. \tag{4.34}$$

Therefore, if $\mathscr{X}$ is stochastically invariant, then $\mathscr{X} = \mathscr{D}$, since $\theta(x)$ is singular for $x \in \partial\mathscr{X}$ by condition (4.19), required for invariance. For the first part of the proposition, it remains to show that conditions (4.19) and (4.20) are equivalent to (4.31) and (4.32). Permuting and rescaling the elements in v and column vectors in Σ we may assume $v_i = u_i$ for $i \leq q$, that is,

$$\mathscr{X} = \bigcap_{i=1}^{q}\{v_i \geq 0\}.$$

Now, Condition (4.19) reads

$$\forall i \in Q, j \in P : \beta_i \Sigma^j = 0 \text{ or } v_j = cv_i \;\; \text{for some } c > 0,$$

by Proposition A.4. The constant c can be taken equal to 1 by rescaling v and Σ. This gives the equivalence between (4.19) and (4.20), while the equivalence between (4.31) and (4.32) is immediate.

For the second part of the proposition we follow the proof of Proposition 4.1 up to Step 4. If we take

$$u_i = (\beta_i \Sigma \Sigma^\top \beta_i^\top)^{-1}v_i, \text{ for } i \leq q,$$

then (4.10) gives

$$c_i = B_i \gamma_i^\top = v_i(x_0)^{-1}(\beta_i \Sigma \Sigma^\top \beta_i^\top)^{-1}\beta_i \Sigma \, \mathrm{diag}(v(x_0))\Sigma^\top \beta_i^\top = 1,$$

by choosing $x_0 \in \mathscr{D}^\circ$ such that $v_i(x_0) = 1$ for $i \leq q$. Hence the affine transformation $Lx + \ell$ from Proposition 4.1 satisfies

$$(Lx + \ell)_i = u_i(x) = (\beta_i \Sigma \Sigma^\top \beta_i^\top)^{-1} v_i(x), \quad \text{for } i \in Q.$$

Let X be an affine diffusion with SDE (2.1). We have that $Y := (LX + \ell)_Q$ satisfies an affine SDE of the form

$$\mathrm{d}Y_t = \tilde{\mu}(Y_t)\mathrm{d}t + \mathrm{diag}(\sqrt{|Y_t|})\mathrm{d}W_t,$$

with state space $\mathbb{R}^q_{\geq 0}$, where we write

$$\tilde{\mu}(x) = L\mu(L^{-1}(x - \ell)) = \tilde{a}x + \tilde{b},$$

for some matrix $\tilde{a}$ with non-negative off-diagonal elements and vector $\tilde{b}$ with non-negative components. Note that $\mathscr{D}^\circ$ is invariant for X if and only if $\mathbb{R}^q_{>0}$ is invariant for Y. By Proposition 4.13, the latter holds if $\tilde{b}_i \geq \frac{1}{2}$ for all i, or equivalently, if $\tilde{\mu}_i(y) \geq \frac{1}{2}$ for all $y \in \mathbb{R}^q_{\geq 0}$ with $y_i = 0$. Substituting $y = Lx + \ell$ gives

$$\tilde{\mu}_i(Lx + \ell) \geq \frac{1}{2} \quad \text{for all } x \in \partial \mathscr{D}_i.$$

Since $\tilde{\mu}(x) = L\mu(L^{-1}(x - \ell))$ and $L_i = (\beta_i \Sigma \Sigma^\top \beta_i^\top)^{-1}\beta_i$, the result follows. $\qquad \square$

Remark 4.15. In [10], it is assumed that $q = p$. Moreover, the strong existence and uniqueness is only proved under (4.31) and (4.33), whereas the case that the process X might hit the boundary $\partial \mathscr{X}$ is not treated. Note also that in [10] the inequality in (4.33) is strict.

5. Quadratic State Space

In this section, we consider affine diffusions where the boundary of the state space $\mathscr{X}$ is quadratic instead of linear. Let us be given a quadratic function

$$\Phi(x) = x^\top Ax + b^\top x + c, \tag{5.1}$$

for some symmetric non-zero $A \in \mathbb{R}^{p \times p}$, $b \in \mathbb{R}^p$, $c \in \mathbb{R}$. We take $\mathscr{X} = \{\theta \geq 0\}$ where θ is given by (4.2) (thus, $\mathscr{X}$ is convex) and we assume $\mathscr{X}^\circ$ is a non-empty connected component (maximal connected subset) of $\{\Phi > 0\}$ or $\{\Phi < 0\}$. Note that then automatically the boundary of the state space is quadratic, i.e. $\partial \mathscr{X} \subset \{\Phi = 0\}$. By the following proposition, there are only three types possible for Φ.

Proposition 5.1. *Let $\mathscr{X} \subset \mathbb{R}^p$ be convex and assume $\mathscr{X}^\circ$ is a non-empty connected component of $\{\Phi > 0\}$ or $\{\Phi < 0\}$, with Φ given by (5.1). Then there exists an affine transformation such that either $\Phi(x) = x_1 - \sum_{i=2}^q x_i^2$, $\Phi(x) = \sum_{i=1}^q x_i^2 + d$ or $\Phi(x) = x_1^2 - \sum_{i=2}^q x_i^2 + d$, for some $1 \leq q \leq p$, $d \in \mathbb{R}$.*

Proof. Since A is symmetric, it is diagonalizable by an orthogonal matrix. By further scaling one can take the diagonal elements equal to -1, 0 or 1. Using the equality

$$x^\top x + b^\top x = \left(x^\top + \frac{1}{2}b^\top\right)\left(x + \frac{1}{2}b\right) - \frac{1}{4}b^\top b,$$

we can apply an affine transformation such that for some disjoint $Q, Q' \subset P\backslash\{1\}$ the quadratic function Φ is of the form

$$\Phi(x) = x_1 - \sum_{i \in Q} x_i^2 + \sum_{i \in Q'} x_i^2, \tag{5.2}$$

or

$$\Phi(x) = x_1^2 - \sum_{i \in Q} x_i^2 + \sum_{i \in Q'} x_i^2 + d, \tag{5.3}$$

for some $d \in \mathbb{R}$. If Φ is of the form (5.2), then $\mathscr{X}$ is of the form

$$\mathscr{X} = \left\{x \in \mathbb{R}^p : x_1 \geq \sum_{i \in Q} x_i^2 - \sum_{i \in Q'} x_i^2\right\},$$

possibly after replacing x_1 by $-x_1$ and interchanging Q and Q'. Convexity of $\mathscr{X}$ yields that the Hessian of $\sum_{i \in Q} x_i^2 - \sum_{i \in Q'} x_i^2$ is positive semi-definite, which implies that $Q' = \emptyset$. Permuting coordinates gives $Q = \{2, \ldots, q\}$ with $q = \#Q + 1$.

Now assume Φ is of the form (5.3). We have to show that either $Q' = \emptyset$ or $\#Q \leq 1$. Define the function f by $f(x_{Q \cup Q'}) = \sum_{i \in Q} x_i^2 - \sum_{i \in Q'} x_i^2 - d$. There are two possible forms $\mathscr{X}$ can assume, namely

$$\mathscr{X} = \{x \in \mathbb{R}^p : x_1 \geq f(x_{Q \cup Q'})^{1/2}, x_{Q \cup Q'} \in K\}$$

$$\text{or } \mathscr{X} = \{x \in \mathbb{R}^p : |x_1| \leq f(x_{Q \cup Q'})^{1/2}, x_{Q \cup Q'} \in K\},$$

where K is convex with K° a nonempty connected component of $\{f \geq 0\}$. In the first case we have that the Hessian of $f(x_{Q \cup Q'})^{1/2}$ is positive semi-definite, while in the second case the Hessian is negative semi-definite. Now suppose $Q' \neq \emptyset$. We show that in that case $\#Q \leq 1$. Let $x_{Q \cup Q'} \in K^\circ$. For $i \in Q'$, we have

$$\frac{\partial^2}{\partial x_i^2} f(x_{Q \cup Q'})^{1/2} = -f(x_{Q \cup Q'})^{-1/2} - x_i^2 f(x_{Q \cup Q'})^{-3/2} < 0,$$

hence, the Hessian is negative semi-definite. For $i \in Q$, we have

$$\frac{\partial^2}{\partial x_i^2} f(x_{Q \cup Q'})^{1/2} = f(x_{Q \cup Q'})^{-1/2} - x_i^2 f(x_{Q \cup Q'})^{-3/2},$$

which, therefore, also has to be negative. Now if $i, j \in Q$ and $i \neq j$, then

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x_{Q \cup Q'})^{1/2} = -f(x_{Q \cup Q'})^{-3/2} x_i x_j,$$

from which we deduce that

$$\det \begin{pmatrix} \frac{\partial^2}{\partial x_i^2} & \frac{\partial^2}{\partial x_i \partial x_j} \\ \frac{\partial^2}{\partial x_j \partial x_i} & \frac{\partial^2}{\partial x_j^2} \end{pmatrix} f(x_{Q \cup Q'})^{-1/2}$$

$$= f(x_{Q \cup Q'})^{-1/2} \left(f(x_{Q \cup Q'})^{-1/2} - (x_i^2 + x_j^2) f(x_{Q \cup Q'})^{-3/2} \right) < 0.$$

This contradicts the negative semi-definiteness of the Hessian. Thus, it holds that $\#Q \leq 1$. $\qquad\qquad\square$

If (2.1) is an affine SDE with drift μ, diffusion matrix θ and state space $\mathscr{X} = \{\theta \geq 0\}$ with non-empty quadratic boundary $\partial \mathscr{X} \subset \{\Phi = 0\}$, then stochastic invariance of $\mathscr{X}$ yields

$$\nabla\Phi(x)\theta(x) = 0, \quad \text{for all } x \in \partial\mathscr{X}. \tag{5.4}$$

by Proposition 3.1 and the remark preceding Proposition 3.4. This excludes that Φ is of the form $\Phi(x) = \sum_{i=1}^q x_i^2 + d$ or $\Phi(x) = x_1^2 - \sum_{i=2}^q x_i^2 + d$ with $d \neq 0$. Indeed, suppose $\Phi(x) = \sum_{i=1}^p x_i^2 + d$ with $d < 0$ (for simplicity we take $q = p$), then (5.4) reads

$$x^\top \theta(x) = 0, \quad \text{for all } x \quad \text{such that } \Phi(x) = 0.$$

Regarding $\Phi(x)$ as a univariate polynomial in x_1, we see that it has distinct roots

$$\pm \left(-d - \sum_{j=2}^p x_j^2 \right)^{1/2},$$

for x close enough to zero. Therefore, these are also the roots of $x^\top \Psi_i(x)$. Since the latter has maximal degree 2, it follows that

$$x^\top \Psi_i(x) = P_i \Phi(x), \quad \text{for all } i \quad \text{and} \quad x \in \mathbb{R}^p,$$

for some constant P_i. Note that the right-hand side of the above display has a constant term $P_i d$. Since the left-hand side only contains multiples of x, this yields $P_i = 0$ for all i, so that $x^\top \Psi(x) = 0$ for all $x \in \mathbb{R}^p$. Lemma 5.2 yields that $\Psi(x) = 0$ for all x, which contradicts the assumption that $\mathscr{X} = \{\theta \geq 0\}$ and $\partial\mathscr{X} \neq \emptyset$. Likewise one can show that $\Phi(x) = x_1^2 - \sum_{i=2}^q x_i^2 + d$ with $d \neq 0$ is impossible.

Lemma 5.2. *Let* $\theta : \mathbb{R}^p \to \mathbb{R}^{p \times p} : x \mapsto A^0 + \sum_{k=1}^p A^k x_k$ *with* $A^k \in \mathbb{R}^{p \times p}$ *symmetric and assume* $x^\top \theta(x) = 0$ *for all* x. *Then* $\theta(x) = 0$ *for all* x.

Proof. It is clear that $A^0 = 0$. We show that $A^k = 0$ for $k > 0$. It holds that

$$0 = \left(x^\top \sum_{k=1}^p A^k x_k \right)_j = \sum_{k=1}^p \sum_{i=1}^p x_i x_k A_{ij}^k = \sum_{1 \leq i < k \leq p} x_i x_k (A_{ij}^k + A_{kj}^i)$$

$$+ \sum_{i=1}^p x_i^2 A_{ij}^i,$$

for all j. Hence, for all i, j we have $A^i_{ij} = 0$ and $A^k_{ij} = -A^i_{kj}$ for $k \neq i$. Since A^k is symmetric, the latter gives $A^k_{ij} = -A^i_{kj} = -A^i_{jk}$. So, if we permute the indices i, j, k by the cycle $(i \mapsto j, j \mapsto k, k \mapsto i)$, then A^k_{ij} gets a minus sign. Permuting the indices repeatedly we obtain

$$A^k_{ij} = -A^i_{jk} = A^j_{ki} = -A^k_{ij},$$

which implies $A^k_{ij} = 0$ for all i, j and $k \neq i$. Hence, $A^k = 0$ for all k, as we have already shown that $A^i_{ij} = 0$ for all i, j. $\qquad\square$

Thus, in order to characterize all affine diffusions with quadratic state space, there are two cases to consider, namely $\Phi(x) = x_1 - \sum_{i=2}^q x_i^2$ and $\Phi(x) = x_1^2 - \sum_{i=2}^q x_i^2$. In the next subsections we characterize for these two forms of Φ all possible θ which can act as a diffusion matrix of an affine SDE, i.e. which θ satisfy (5.4). Moreover, we are able to construct a square root σ of θ such that (2.1) is an affine SDE with quadratic state space $\mathscr{X}$, generalizing the two-dimensional setting as treated in [9, Section 12] and [12]. In particular we show existence and uniqueness of a strong solution.

5.1. *Parabolic State Space*

Assume Φ is of the form $\Phi(x) = x_1 - \sum_{i=2}^q x_i^2$, with $1 < q \leq p$. The state space $\mathscr{X}$ then necessarily equals $\mathscr{X} = \{\Phi \geq 0\}$. For $x \in \mathbb{R}^p$ we write $x = (x_1, y^\top, z) \in \mathbb{R}^1 \times \mathbb{R}^{q-1} \times \mathbb{R}^{p-q}$ and we define affine matrix-valued functions ζ and η by

$$\zeta(x) = \begin{pmatrix} 4x_1 & 2y^\top \\ 2y & I \end{pmatrix}, \quad \eta(x) = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ T_{12}(y) & T_{13}(y) & \ldots & T_{q-2,q-1}(y) \end{pmatrix},$$

with $T_{ij} : \mathbb{R}^{q-1} \to \mathbb{R}^{q-1}$ for $1 \leq i < j < q$ given by $T_{ij}(y)_i = y_j$, $T_{ij}(y)_j = -y_i$, $T_{ij}(y)_k = 0$ for $k \neq i, j$. Moreover, we write $Q = \{2, \ldots, q\}$. Now, condition (5.4) reads

$$\begin{pmatrix} 1 & -2y^\top & 0 \end{pmatrix} \theta(x) = 0 \quad \text{for all } x \in \mathbb{R}^p \text{ such that } x_1 = y^\top y. \tag{5.5}$$

We use the following lemmas to characterize those θ that satisfy this condition in Proposition 5.5.

Lemma 5.3. *Consider the linear space*

$$\mathscr{L} = \left\{ a : \mathbb{R}^p \to \mathbb{R}^q \text{ affine } \mid \begin{pmatrix} 1 & -2y^\top \end{pmatrix} a(x) = 0 \text{ for all } x \text{ with } x_1 = y^\top y \right\}. \tag{5.6}$$

Then a basis for $\mathscr{L}$ is formed by the columns of ζ and η.

Proof. Clearly these columns are linearly independent elements of $\mathscr{L}$. To prove that they span $\mathscr{L}$ we use a dimension argument. Let $\text{Aff}(\mathbb{R}^p, \mathbb{R}^q)$ denote the space of affine functions from $\mathbb{R}^p$ to $\mathbb{R}^q$ and let $\text{Quadr}(\mathbb{R}^p, \mathbb{R})/(x_1 - y^\top y)$ be the space of quadratic functions from $\mathbb{R}^p$ to $\mathbb{R}$, modulo $x_1 - y^\top y$ (that is, p and q are equivalent if $p(x) - q(x) = c(x_1 - y^\top y)$ for some constant c). Consider the linear operator

$$L : \text{Aff}(\mathbb{R}^p, \mathbb{R}^q) \to \text{Quadr}(\mathbb{R}^p, \mathbb{R})/(x_1 - y^\top y) : a(x) \mapsto \begin{pmatrix} 1 & -2y^\top \end{pmatrix} a(x),$$

and note that $\mathcal{L} = \ker L$. By the dimension theorem for linear operators, we have

$$\dim \mathrm{Aff}(\mathbb{R}^p, \mathbb{R}^q) = \dim \ker L + \dim \mathrm{im}\, L.$$

It holds that $\dim \mathrm{Aff}(\mathbb{R}^p, \mathbb{R}^q) = pq + q$. Since $x_1 \equiv y^\top y$, a basis for $\mathrm{im}\, L$ is given by

$$\{1, x_2, \ldots, x_p\} \cup \{x_i x_j : 2 \le i \le q, 1 \le j \le p\},$$

when

$$\dim \mathrm{im}\, L = p + (q-1)p - \binom{q-1}{2} = pq - \binom{q-1}{2}.$$

It follows that $\dim \ker L = q + \frac{1}{2}(q-1)(q-2)$, which is the number of columns in ζ and η. Thus, the columns span the kernel of L. ☐

Lemma 5.4. *Let $\mathcal{L}$ be defined by (5.6) and suppose $M : \mathbb{R}^p \to \mathbb{R}^{q \times q}$ is affine and $M(x)$ is symmetric for all $x \in \mathbb{R}^p$. If the columns of M are in $\mathcal{L}$, then $M = c\zeta$ for some $c \in \mathbb{R}$.*

Proof. By Lemma 5.3 there exist matrices A and B such that

$$M(x) = \zeta(x)A + \eta(x)B.$$

Write $T(y) = (T_{ij}(y))_{1 \le i < j < q}$ and $B = \begin{pmatrix} B^1 & \widetilde{B} \end{pmatrix}$. Then the above display reads

$$M(x) = \begin{pmatrix} 4x_1 A_{11} + 2y^\top A_{Q1} & 4x_1 A_{1Q} + 2y^\top A_{QQ} \\ 2y A_{11} + A_{Q1} + T(y)B^1 & 2y A_{1Q} + A_{QQ} + T(y)\widetilde{B} \end{pmatrix}$$

Since $M(x)$ is symmetric it immediately follows that $A_{1Q} = 0$ and $A_{Q1} = 0$. Define $N(x) = M(x) - A_{11}\zeta(x)$. Then N is also symmetric. We have

$$N(x) = \begin{pmatrix} 0 & 2y^\top C \\ T(y)B^1 & C + T(y)\widetilde{B} \end{pmatrix},$$

with $C = A_{QQ} - A_{11}I$. This yields $C = C^\top$ and $T(y)B^1 = 2Cy$. Since $y^\top T(y) = 0$, the latter implies $y^\top Cy = 0$, whence $C = 0$, as C is diagonalizable by an orthogonal matrix. Thus, $A_{QQ} = A_{11}I$ and it remains to show that $\widetilde{B} = 0$.

It holds that $T(y)\widetilde{B}$ is symmetric and $y^\top T(y)\widetilde{B} = 0$. Lemma 5.2 yields $T(y)\widetilde{B} = 0$, whence $\widetilde{B} = 0$ by linear independence of the columns of $T(y)$, as we needed to prove. ☐

Proposition 5.5. *If (5.5) holds, then necessarily θ is of the form*

$$\theta(x) = \begin{pmatrix} c\zeta(x) & A(x) \\ A(x)^\top & B(x) \end{pmatrix}, \tag{5.7}$$

for some $c \ge 0$, with $A(x) = \zeta(x)A_1 + \eta(x)A_2$ for some matrices A_1, A_2 and $B : \mathbb{R}^p \to \mathbb{R}^{(p-q) \times (p-q)}$ affine and symmetric. Moreover, if $c = 1$ and $A_1 = 0$, then it holds that

$$B(x) - A_2^\top \eta(x)^\top \eta(x) A_2 \ge 0, \tag{5.8}$$

for all $x \in \mathcal{X} = \{\theta \ge 0\}$.

Proof. The first part follows from Lemma 5.3 and Lemma 5.4. It remains to show (5.8).

Suppose $c = 1$ and $A_1 = 0$. By positive semi-definiteness of θ, we have

$$0 \leq \begin{pmatrix} v^\top & w^\top \end{pmatrix} \theta(x) \begin{pmatrix} v \\ w \end{pmatrix}$$

$$= v^\top \zeta(x) v + 2 v^\top \eta(x) A_2 w + w^\top B(x) w,$$

for all $v \in \mathbb{R}^q$, $w \in \mathbb{R}^{p-q}$, $x \in \mathscr{X}$. Fix $w \in \mathbb{R}^{p-q}$, $x \in \mathscr{X}$ arbitrarily and take $v = -\eta(x) A_2 w$. Noting that $\zeta(x)\eta(x) = \eta(x)$ for all $x \in \mathbb{R}^p$, the above display then reads

$$w^\top B(x) w - w^\top A_2^\top \eta(x)^\top \eta(x) A_2 w \geq 0,$$

which proves (5.8). $\qquad\square$

To show the existence of an affine diffusion with parabolic state space in Theorem 5.7, we need the following result. Its proof is based on a modification of a result by Yamada and Watanabe [23, Theorem 1].

Proposition 5.6. *There exists a unique strong solution to the SDE*

$$d \begin{pmatrix} X_{1,t} \\ Y_t \end{pmatrix} = \begin{pmatrix} a_{11} X_{1,t} + a_{1Q} Y_t + b_1 \\ a_{QQ} Y_t + b_Q \end{pmatrix} dt + \begin{pmatrix} 2\sqrt{|X_{1,t} - Y_t^\top Y_t|} & 2 Y_t^\top \\ 0 & I \end{pmatrix} dW_t.$$

Proof. By continuity of the coefficients and satisfaction of the linear growth condition (2.2), there exists a weak solution $((X_1, Y), W)$ on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ carrying a Brownian motion W. For strong existence and uniqueness it suffices to prove pathwise uniqueness. Therefore, assume $((\widetilde{X}_1, \widetilde{Y}), W)$ is another solution on the same probability space. We see that the equation for Y does not contain X_1 and in fact it is an SDE which has a unique strong solution, whence $Y_t = \widetilde{Y}_t$ a.s. for all $t \geq 0$. Write $Z = X_1 - Y^\top Y$ and $\widetilde{Z} = \widetilde{X}_1 - \widetilde{Y}^\top \widetilde{Y}$. Then it follows that

$$d(Z_t - \widetilde{Z}_t) = a_{11}(Z_{1,t} - \widetilde{Z}_{1,t}) dt + 2\left(\sqrt{|Z_t|} - \sqrt{|\widetilde{Z}_t|}\right) dW_{1,t}.$$

Arguing as in the proof of [17, Proposition 5.2.13] (for instance), we deduce that $Z_t = \widetilde{Z}_t$ a.s. for all $t \geq 0$. Thus, $X_{1,t} = Z_t + Y_t^\top Y_t = \widetilde{Z}_t + \widetilde{Y}_t^\top \widetilde{Y}_t = \widetilde{X}_{1,t}$ a.s. for all $t \geq 0$. $\qquad\square$

Theorem 5.7. *Let $\Phi(x) = x_1 - y^\top y$ and suppose $\{\theta \geq 0\} = \{\Phi \geq 0\}$. Then there exists an affine SDE with drift μ, diffusion matrix θ and state space $\mathscr{X} = \{\theta \geq 0\}$ if and only if*

$$\nabla \Phi(x) \theta(x) = 0, \tag{5.9}$$

$$\nabla \Phi(x) \mu(x) \geq tr(\theta_{QQ}(x)) \tag{5.10}$$

for all $x \in \partial \mathscr{X} = \{\Phi = 0\}$.

Proof. Conditions (3.6) and (3.7) reduce to (5.9) and (5.10). The "only if" part follows, as these boundary conditions are necessary for stochastic invariance of $\mathscr{X}$ by Proposition 3.4.

Suppose the boundary conditions (5.9) and (5.10) hold. Then, by Proposition 5.5, it holds that θ is of the form (5.7) for some $c \geq 0$. If $c = 0$, then necessarily $\mathscr{X} = \{B \geq 0\}$ and we take

$$\sigma(x) = \begin{pmatrix} 0 & 0 \\ 0 & |B(x)|^{1/2} \end{pmatrix}.$$

It is easy to see that (2.1) admits a unique strong solution and that $\mathscr{X}$ is invariant. If $c > 0$, then we apply the linear transformation $x_1 \mapsto c^{-1} x_1$, $y \mapsto c^{-1/2} y$, $z \mapsto z - c^{-1} A_1(x_1, y)$, so that we may assume $c = 1$ and $A_1 = 0$. Then (5.8) holds and as a square root of θ on $\mathscr{X}$ we take

$$\sigma(x) = \begin{pmatrix} \xi(x) & 0 \\ A_2^\top \eta(x)^\top & \rho(x) \end{pmatrix},$$

with ξ and ρ defined by

$$\xi(x) = \begin{pmatrix} 2\sqrt{|x_1 - y^\top y|} & 2y^\top \\ 0 & I \end{pmatrix},$$

$$\rho = |B - A_2^\top \eta^\top \eta A_2|^{1/2}.$$

To see that $\sigma(x)\sigma(x)^\top = \theta(x)$ for $x \in \mathscr{X}$, note that $\xi(x)\eta(x) = \eta(x)$ for all $x \in \mathscr{X}$. Proposition 3.4 gives that $\mathscr{X}$ is invariant for (2.1). It remains to prove existence and uniqueness of a strong solution.

The boundary condition (5.10) reads

$$\mu_1(x) - 2y^\top \mu_Q(y) \geq q - 1, \quad \text{for all} x \quad \text{such that } x_1 = y^\top y,$$

so necessarily $\mu_{\{1\} \cup Q}$ admits the form

$$\mu_{\{1\} \cup Q}(x) = \begin{pmatrix} a_{11} & a_{1Q} & 0 \\ 0 & a_{QQ} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ y \\ z \end{pmatrix} + b_{\{1\} \cup Q}. \tag{5.11}$$

Proposition 5.6 gives the result.   □

For stochastic invariance of a parabolic state space $\mathscr{X}$, we now give sufficient conditions on the diffusion matrix and sufficient and necessary conditions on the drift, analogous to the admissibility conditions for polyhedral state spaces in canonical form.

Proposition 5.8. *Condition* (5.8) *holds if B admits the form*

$$B(x) = (q - 2)x_1 A_2^\top A_2.$$

Proof. One can show using Cauchy-Schwarz that

$$\eta(x)_i \eta(x)_i^\top I \geq \eta(x)_i^\top \eta(x)_i, \quad \text{for all } x \text{ and } i.$$

Hence,

$$\left(\sum_i \eta(x)_i \eta(x)_i^\top \right) I \geq \sum_i \eta(x)_i^\top \eta(x)_i.$$

The right-hand side equals $\eta(x)^\top \eta(x)$, while the left-hand side is equal to $(q-2)\left(\sum_i y_i^2\right) I$, which is smaller than $(q-2)x_1 I$ for $x \in \mathscr{X}$. Hence,

$$(q-2)x_1 A_2^\top A_2 - A_2^\top \eta(x)^\top \eta(x) A_2 = A_2^\top ((q-2)x_1 I - \eta(x)^\top \eta(x)) A_2 \geq 0,$$

for $x \in \mathscr{X}$, which yields the result. $\qquad\square$

Proposition 5.9. *Suppose X is an affine diffusion with state space $\mathscr{X} = \{\theta \geq 0\} = \{x \in \mathbb{R}^p : x_1 \geq y^\top y\}$ and $\theta_{QQ} \neq 0$. Then X is a linear transformation of an affine diffusion with the same state space $\mathscr{X}$, diffusion matrix θ of the form (5.7) with $c = 1$ and with drift $\mu(x) = ax + b$ satisfying (5.11) as well as*

$$a_{11}I - 2a_{QQ} = diag(d_Q) \tag{5.12}$$

$$d_{Q_1} > 0 \tag{5.13}$$

$$d_{Q_2} = 0 \tag{5.14}$$

$$a_{1Q_2} = 2b_{Q_2} \tag{5.15}$$

$$b_1 \geq q - 1 + \sum_{i \in Q_1} \tfrac{1}{4} d_i^{-1}(a_{1i} - 2b_i)^2, \tag{5.16}$$

for some vector d and some disjoint Q_1 and Q_2 with $Q = Q_1 \cup Q_2$.

Proof. As in the proof of Theorem 5.7 we can assume θ is of the form (5.7) with $c = 1$. Then $\text{tr}(\theta_{QQ}(x)) = q - 1$ for all x, so the boundary condition (5.10) for the drift reads

$$y^\top (a_{11}I - 2a_{QQ})y + (a_{1Q} - 2b_Q^\top)y + b_1 - q + 1 \geq 0, \quad \text{for all } y \in \mathbb{R}^{p-1}.$$

For this it is necessary that $M := a_{11}I - 2a_{QQ}$ is positive semi-definite. Moreover, if y is in the kernel of M, then y should also be in the kernel of $a_{1Q} - 2b_Q^\top$. We can diagonalize M by an orthogonal matrix O, so $D = OMO^\top$ is diagonal with positive diagonal elements d_i for $i \in Q_1$ and $d_i = 0$ for $i \in Q_2 = Q \backslash Q_1$, for some $Q_1 \subset Q$. Applying the orthogonal transformation $y \mapsto Oy$, the above condition becomes

$$\sum_{i \in Q_1} d_i x_i^2 + \sum_{i \in Q_1} (a_{1i} - 2b_i)x_i + b_1 - q + 1 \geq 0, \quad \text{for all } x_{Q_1}.$$

We can write the left-hand side as

$$\sum_{i \in Q_1} d_i \left(x_i + \frac{1}{2} d_i^{-1}(a_{1i} - 2b_i) \right)^2 - \frac{1}{4} \sum_{i \in Q_1} d_i^{-1}(a_{1i} - 2b_i)^2 + b_1 - q + 1,$$

which is non negative for all x_{Q_1} if and only if

$$-\frac{1}{4}\sum_{i\in Q_1}d_i^{-1}(a_{1i}-2b_i)^2+b_1-q+1\geq 0.$$

This yields the result.   □

Remark 5.10. Note that (5.12) implies that a_{QQ} is diagonal. Hence, the coordinates of X_Q are mutually independent.

Proposition 5.11. *Consider the situation of Proposition 5.9. If we strengthen condition (5.16) to*

$$b_1\geq q+1+\sum_{i\in Q_1}\frac{1}{4}d_i^{-1}(a_{1i}-2b_i)^2, \tag{5.16$'$}$$

then $\mathcal{X}^\circ$ becomes invariant.

Proof. It suffices to verify condition (3.9) of Proposition 3.7. Recall that $\Phi(x)=x_1-y^\top y$ and $\theta(x)=A^0+\sum_{i=1}^p A^i x_i$ is of the form (5.7) with $c=1$. It follows that $\nabla\Phi(x)(\mu(x)-\frac{1}{2}\sum_{i=1}^p(A^i)^i)$ equals

$$a_{11}x_1+a_{1Q}y-2y^\top a_{QQ}y-2y^\top b_Q+b_1-q-1.$$

If in addition to conditions (5.12)–(5.16$'$) also $a_{11}\geq 0$ is imposed, then for $x\in\mathcal{X}$ the above display is bounded from below by

$$y^\top(a_{11}I-2a_{QQ})y+(a_{1Q}-2b_Q^\top)y+b_1-q-1.$$

This is non negative for all $y\in\mathbb{R}^{q-1}$ under the imposed assumptions, similar as in the proof of Proposition 5.9, which yields (3.9). The nonnegativity of a_{11} can be dispensed with, as shown as follows.

By applying a measure transformation with density $\mathscr{E}(\lambda^\top\sigma(X)\cdot W)$ for some $\lambda\in\mathbb{R}^p$ with $\lambda_i=0$ for $i\neq 1$ (which yields a probability measure by [20, Proposition 3.4]), we see that $\mathcal{X}$ is also invariant for the SDE with drift $\tilde{\mu}(x)=ax+b+\theta(x)\lambda=\tilde{a}x+b$, where $\tilde{a}_{11}=a_{11}+4\lambda$, $\tilde{a}_{QQ}=a_{QQ}+2\lambda I$ and the remaining coordinates unaltered. Note that $\tilde{a}_{11}I-2\tilde{a}_{QQ}=a_{11}I-2a_{QQ}$, so conditions (5.12)–(5.16$'$) are not affected by such a measure transformation. This gives the result.   □

Example 5.12. Here we give a concrete example of an affine diffusion on a parabolic state space, which is not of the kind discussed in [12]. Consider the state space $\mathcal{X}=\{(x_1,y,z)\in\mathbb{R}\times\mathbb{R}^2\times\mathbb{R}:x_1^2\geq y_1^2+y_2^2\}$. As diffusion matrix we take

$$\theta(x)=\begin{pmatrix}4x_1 & 2y_1 & 2y_2 & 0\\ 2y_1 & 1 & 0 & y_2\\ 2y_2 & 0 & 1 & -y_1\\ 0 & y_2 & -y_1 & x_1\end{pmatrix}$$

and we take the drift μ of the form

$$\mu(x) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & \frac{1}{2}(a_{11} - d_1) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(a_{11} - d_2) & 0 \\ * & * & * & * \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ * \end{pmatrix},$$

with $d_1, d_2 > 0$, b_1 satisfying (5.16) and $*$ arbitrary. Then the conditions of Proposition 5.9 are satisfied (with $Q_2 = \emptyset$), which yields the existence of an affine diffusion with state space $\mathscr{X}$, drift μ and diffusion matrix θ.

We further note that from this example one can easily construct affine diffusions on a parabolic state space with a *degenerate* (and non-zero) diffusion matrix. For this, take $\mathscr{X} \times \mathbb{R}$ and extend $\theta(x)$ to

$$\begin{pmatrix} 4x_1 & 2y_1 & 2y_2 & 0 & 0 \\ 2y_1 & 1 & 0 & y_2 & y_2 \\ 2y_2 & 0 & 1 & -y_1 & -y_1 \\ 0 & y_2 & -y_1 & x_1 & x_1 \\ 0 & y_2 & -y_1 & x_1 & x_1 \end{pmatrix}.$$

5.2. The Lorentz Cone

Let $p \geq q > 1$. For $x \in \mathbb{R}^p$ we write $x = (x_1, y^\top, z) \in \mathbb{R}^1 \times \mathbb{R}^{q-1} \times \mathbb{R}^{p-q}$. We consider the quadratic form

$$\Phi(x) = x_1^2 - \sum_{i \in Q} x_i^2,$$

where $Q = \{2, \ldots, q\}$. The closed convex set $\{\Phi \geq 0\}$ is known as the (closed) *Lorentz cone*. Define affine matrix-valued functions ζ and η by

$$\zeta(x) = \begin{pmatrix} x_1 & y^\top \\ y & x_1 I \end{pmatrix}, \quad \eta(x) = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ T_{12}(y) & T_{13}(y) & \ldots & T_{q-2,q-1}(y) \end{pmatrix},$$

with $T_{ij} : \mathbb{R}^{q-1} \to \mathbb{R}^{q-1}$ for $1 \leq i < j < q$ given by $T_{ij}(y)_i = y_j$, $T_{ij}(y)_j = -y_i$, $T_{ij}(y)_k = 0$ for $k \neq i, j$. By applying a reflection, we may assume the state space $\mathscr{X}$ is of the form $\mathscr{X} = \{\Phi \geq 0\} \cap \{x_1 \geq 0\}$. Analogously to Lemmas 5.3 and 5.4 we have the following.

Lemma 5.13. *Consider the linear space*

$$\mathscr{L} = \left\{ a : \mathbb{R}^p \to \mathbb{R}^q \text{ affine} \mid \begin{pmatrix} x_1 & -y^\top \end{pmatrix} a(x) = 0 \text{ for all } x \text{ with } x_1^2 = y^\top y \right\}. \quad (5.17)$$

Then a basis for $\mathscr{L}$ is formed by the columns of ζ and η.

Proof. Similar to the proof of Lemma 5.3. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 5.14. *Consider the linear space*

$$\mathscr{M} = \left\{ M : \mathbb{R}^p \to \mathbb{R}^{q \times q} \text{ affine} \mid M(x) \text{ symmetric and } M^i \in \mathscr{L} \text{ for all } x, i \right\}$$

with $\mathcal{L}$ defined by (5.17). *Then a basis for $\mathcal{M}$ is given by*

$$\mathcal{B} = \{\zeta, \rho(1), \ldots, \rho(q-1)\},$$

with $\rho(i)$ an affine symmetric-matrix valued function defined defined by its rows (with e_k the standard kth basis vector of $\mathbb{R}^p$)

$$\rho(i)_1 : x \mapsto \begin{pmatrix} y_i & 0 \end{pmatrix},$$
$$\rho(i)_{i+1} : x \mapsto \begin{pmatrix} x_1 & y^\top \end{pmatrix},$$
$$\rho(i)_j : x \mapsto -y_i e_j^\top + y_{j-1} e_{i+1}^\top \quad \text{for } j \neq 1, i+1.$$

Proof. Clearly the elements of $\mathcal{B}$ are linearly independent elements of $\mathcal{M}$. It remains to show that they span $\mathcal{M}$.

Let $M \in \mathcal{M}$ be arbitrary. By Lemma 5.13 there exist matrices A and B such that

$$M(x) = \zeta(x)A + \eta(x)B.$$

Write $Q = \{2, \ldots, q\}$, $T(y) = (T_{ij}(y))_{1 \leq i < j < q}$ and $B = \begin{pmatrix} B^1 & \widetilde{B} \end{pmatrix}$. Then the above display reads

$$M(x) = \begin{pmatrix} x_1 A_{11} + y^\top A_{Q1} & x_1 A_{1Q} + y^\top A_{QQ} \\ y A_{11} + x_1 A_{Q1} + T(y)B^1 & y A_{1Q} + x_1 A_{QQ} + T(y)\widetilde{B} \end{pmatrix}$$

Symmetry of $M(x)$ yields

$$A_{1Q} = A_{Q1}^\top,$$
$$A_{QQ} = A_{QQ}^\top,$$
$$y A_{11} + T(y)B^1 = A_{QQ}^\top y,$$
$$y A_{1Q} + T(y)\widetilde{B} = (y A_{1Q} + T(y)\widetilde{B})^\top.$$

Since $y^\top T(y) = 0$, the second equation together with the third gives

$$0 = y^\top T(y)B^1 = y^\top (A_{QQ} - A_{11}I)y,$$

which implies $A_{QQ} - A_{11}I = 0$, as $A_{QQ} - A_{11}I$ is symmetric and thus diagonalizable by an orthogonal matrix. Define

$$N = M - A_{11}\zeta - \sum_{i \in Q} A_{1i}\rho(i-1).$$

Then $N \in \mathcal{M}$ and N is of the form

$$N(x) = \begin{pmatrix} 0 & 0 \\ 0 & \sum_{k \in Q} C^k y_k \end{pmatrix},$$

for some symmetric $((q-1) \times (q-1))$-matrices C^k. By Lemma 5.2, it follows that $N = 0$. $\square$

Unlike the parabolic case, for a general Lorentz cone state space we are not able to find a square root such that strong existence and uniqueness for the resulting SDE can be proved (for showing existence and unqiueness of a weak solution though, one can apply the results of [4, 22]). An exception is the two-dimensional cone, as this is just a polyhedron which has already been covered in Section 4. The following example shows that problems appear for closed Lorentz cones in higher dimensions.

Example 5.15. For $p = q = 3$, a basis for $\mathcal{M}$ is given by

$$\zeta(x) = \begin{pmatrix} x_1 & y_1 & y_2 \\ y_1 & x_1 & 0 \\ y_2 & 0 & x_1 \end{pmatrix}, \quad \rho(1)(x) = \begin{pmatrix} y_1 & x_1 & 0 \\ x_1 & y_1 & y_2 \\ 0 & y_2 & -y_1 \end{pmatrix},$$

$$\rho(2)(x) = \begin{pmatrix} y_2 & 0 & x_1 \\ 0 & -y_2 & y_1 \\ x_1 & y_1 & y_2 \end{pmatrix}.$$

Note that not only ζ but also $\zeta + \rho(1)$ and $\zeta + \rho(2)$ are positive semi-definite on $\mathcal{X} = \{\Phi \geq 0\} \cap \{x_1 \geq 0\} = \{x \in \mathbb{R}^3 : x_1 \geq 0, x_1^2 \geq y^\top y\}$. The structure of these matrices appears to be too complex to compute a manageable square root.

However, Proposition 3.7 enables us to derive sufficient conditions for stochastic invariance of the *open* Lorentz cone $\{\zeta > 0\}$. This can be used to show existence of a unique strong solution for the affine SDE (2.1) with square root $\sigma = |\zeta|^{1/2}$, see the next theorem. Note that this approach is not applicable for $\zeta + \rho(1)$ and $\zeta + \rho(2)$ in Example 5.15, as these matrices are singular on the whole of $\mathbb{R}^3$. We leave the question of (strong) existence of an affine diffusion on a *closed* Lorentz cone open for further research.

Theorem 5.16. *There exists an affine SDE with drift* $\mu(x) = ax + b$, *diffusion matrix*

$$\theta(x) = \begin{pmatrix} x_1 & y^\top \\ y & x_1 I \end{pmatrix},$$

and Lorentz cone state space $\mathcal{X} = \{\theta > 0\} = \{x_1 > (y^\top y)^{1/2}\}$ *if*

$$a_{1Q} - a_{Q1}^\top = 0 \tag{5.18}$$

$$a_{11}I - a_{QQ} \geq 0 \tag{5.19}$$

$$b_1 - \tfrac{1}{2}p - \|b_Q\| \geq 0. \tag{5.20}$$

Proof. Let $\sigma = |\theta|^{1/2}$. Then σ is locally Lipschitz continuous on $\mathcal{X}$, so strong existence and uniqueness for (2.1) follows (see [17, Theorem 5.2.5]) as soon as we have shown stochastic invariance of $\mathcal{X}$. It holds that $\mathcal{X}$ is a connected component of $\{\Phi > 0\}$, with $\Phi(x) = x_1^2 - y^\top y$. Therefore, in view of Proposition 3.7, it suffices to prove (3.9) and (3.10). The first condition is immediate. For the second one, a calculation shows that $\nabla\Phi(x)(\mu(x) - \frac{1}{2}\sum_{i=1}^{p}(A^i)^i)$ equals

$$2\left(a_{11}x_1^2 + x_1(a_{1Q} - a_{Q1}^\top)y + \left(b_1 - \frac{1}{2}p\right)x_1 - y^\top a_{QQ}y - b_Q^\top y\right).$$

This is non negative for all $x \in \mathscr{X}$ if (5.18)–(5.20) hold and $a_{11} \geq 0$. Indeed, in that case we have for $x \in \mathscr{X} = \{x_1^2 > y^\top y\} \cap \{x_1 > 0\}$ that

$$a_{11} x_1^2 + x_1 (a_{1Q} - a_{Q1}^\top) y + \left(b_1 - \frac{1}{2} p\right) x_1 - y^\top a_{QQ} y - b_Q^\top y$$

$$\geq y^\top (a_{11} \mathrm{I} - a_{QQ}) y + \left(b_1 - \frac{1}{2} p\right) x_1 - \langle b_Q, y \rangle$$

$$\geq y^\top (a_{11} \mathrm{I} - a_{QQ}) y + \left(b_1 - \frac{1}{2} p - \|b_Q\|\right) x_1 \geq 0,$$

since $-\langle b_Q, y \rangle \geq -\|b_Q\| \|y\| \geq -\|b_Q\| x_1$ by Cauchy-Schwarz. The nonnegativity of a_{11} can be dispensed with, by the same arguments as in the proof of Proposition 5.11. □

References

1. Bhatia, R. 1997. *Matrix Analysis*. Vol. 169 of Graduate Texts in Mathematics. Springer-Verlag, New York.
2. Cheridito, P., Filipovic, D., and Kimmel, R. 2007. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics* 83:123–170.
3. Cheridito, P., Filipović, D., and Kimmel, R.L. 2010. A note on the Dai-Singleton canonical representation of affine term structure models. *Mathematical Finance* 20:509–519.
4. Cuchiero, C. 2011. Affine and polynomial processes. PhD thesis, ETH Zürich, No. 19652.
5. Cuchiero, C., Filipović, D., Mayerhofer, E., and Teichmann, J. 2011. Affine processes on positive semidefinite matrices. *Annals of Applied Probability* 21:397–463.
6. Da Prato, G., and Frankowska, H. 2007. Stochastic viability of convex sets. *Journal of Mathematical Analysis and Applications* 333:151–163.
7. Dai, Q., and Singleton, K.-J. 2000. Specification analysis of affine term structure models. *Journal of Finance* 55:1943–1978.
8. Duffee, G.-R. 2002. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57:405–443.
9. Duffie, D., Filipović, D., and Schachermayer, W. 2003. Affine processes and applications in finance. *Annals of Applied Probability* 13:984–1053.
10. Duffie, D., and Kan, R. 1996. A yield-factor model of interest rates. *Mathematical Finance* 6:379–406.
11. Filipović, D., and Mayerhofer, E. 2009. Affine diffusion processes: theory and applications. In *Advanced Financial Modelling*. Vol. 8 of Radon Series on Computational and Applied Mathematics. Walter de Gruyter, Berlin, pp. 125–164.
12. Gourieroux, C., and Sufana, R. 2006. Classification of affine term structure models. *Journal of Financial Econometrics* 4:31–52.
13. Gourieroux, C., and Sufana, R. 2011. Discrete time wishart term structure models. *Journal of Economic Dynamics and Control* 35:815–824.
14. Grasselli, M., and Tebaldi, C. 2008. Solvable affine term structure models. *Mathematical Finance* 18:135–153.
15. Ikeda, N., and Watanabe, S. 1981. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.
16. Kallenberg, O. 2002. *Foundations of Modern Probability*. 2nd ed. Probability and its Applications. Springer-Verlag, New York.
17. Karatzas, I., and Shreve, S.E. 1991. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York.

18. Mayerhofer, E., Pfaffel, O., and Stelzer, R. 2011. On strong solutions for positive definite jump diffusions. *Stochastic Processes and Their Applications* 121:2072–2086.

19. Milian, A. 1995. Stochastic viability and a comparison theorem. *Colloquium Mathematicum* 68:297–316.

20. Spreij, P.J.C., and Veerman, E. 2010. *The Affine Transform Formula for Affine Diffusions with Convex State Space.* arXiv:1005.1099v2.

21. Tappe, S. 2009. *Stochastic Invariance of Closed, Convex Sets with Respect to Jump-Diffusions.* Universität Wien working paper 18.

22. Veerman, E. 2011. *Affine Markov Processes on a General State Space.* PhD thesis, Universiteit van Amsterdam.

23. Watanabe, S., and Yamada, T. 1971. On the uniqueness of solutions of stochastic differential equations. *Mathematics of Kyoto University* 11:155–167.

Appendix A. Convex Analysis

In this Appendix, we state and prove the results on convex analysis applied in Section 4. Let $\mathscr{X}$ be given by (4.1) and in addition to an affine function u we are given an affine function d by

$$d : \mathbb{R}^p \to \mathbb{R} : x \mapsto ax + b,$$

for some $a \in \mathbb{R}^{1 \times p}$, $b \in \mathbb{R}$. Proposition A.1 below is the main result, which yields Proposition A.2 to tackle the drift and Proposition A.4 to tackle the diffusion matrix of affine diffusions with non-canonical polyhedral state space.

Proposition A.1. *Suppose $\mathscr{X} \subset \{d \geq 0\}$. Then there exist $c \geq 0$ and $\lambda \in \mathbb{R}_{\geq 0}^{1 \times q}$, such that*

$$d = \lambda u + c.$$

Proof. We give a proof by contradiction. Let

$$\mathscr{K} = \{(\lambda\gamma, \lambda\delta + c) : \lambda \in \mathbb{R}_{\geq 0}^{1 \times q}, c \geq 0\}.$$

Suppose $(a, b) \notin \mathscr{K}$. Since $\mathscr{K}$ is a closed convex set, (a, b) is strictly separated from $\mathscr{K}$ by the Separating Hyperplane Theorem. Therefore, there exist $y \in \mathbb{R}^p$ and $y_0 \in \mathbb{R}$ such that $\langle (y, y_0), (k, k_0) \rangle > \langle (y, y_0), (a, b) \rangle$ for all $(k, k_0) \in \mathscr{K}$, that is,

$$ky + k_0 y_0 > ay + by_0 \quad \text{for all } (k, k_0) \in \mathscr{K}.$$

In other words, for all $\lambda_i \geq 0$ and $c \geq 0$ we have

$$\sum_i \lambda_i(\gamma_i y + \delta_i y_0) + cy_0 > ay + by_0.$$

It easily follows that

$$ay + by_0 < 0 \tag{A.1}$$

$$\gamma_i y + \delta_i y_0 \geq 0 \tag{A.2}$$

$$y_0 \geq 0. \tag{A.3}$$

Using this we construct $x \in \mathscr{X}$ for which $d(x) < 0$. Suppose $y_0 > 0$. Then we take $x = y/y_0$. Indeed, $u_i(x) = (\gamma_i y + \delta_i y_0)/y_0 \geq 0$, so $x \in \mathscr{X}$. But $d(x) = (ay + by_0)/y_0 < 0$, which is a contradiction. Suppose $y_0 = 0$. Then we take an arbitrary $x_0 \in \mathscr{X}$ and let $x_N = x_0 + Ny$, with $N \in \mathbb{N}$. Then $u_i(x_N) = u_i(x_0) + N\gamma_i y \geq 0$ for all N, so $x_N \in \mathscr{X}$, but $d(x_N) = d(x_0) + Nay < 0$ for N big enough. $\qquad\square$

Proposition A.2. *Suppose $\partial\mathscr{X}_i \subset \{d \geq 0\}$ for some $i \in Q$. Then there exist $c \geq 0$ and $\lambda \in \mathbb{R}^{1\times q}$ with $\lambda_j \geq 0$ for $j \in Q\backslash\{i\}$, such that*

$$d = \lambda u + c.$$

Proof. Let $u_0 := -u_i$. Then $\partial\mathscr{X}_i = \bigcap_{j=0}^{q}\{u_j \geq 0\}$ and $d(x) \geq 0$ for $x \in \partial\mathscr{X}_i$. Hence we can apply Proposition A.1, which gives the existence of $\lambda_j \geq 0$ with $j = 0, \ldots, q$ and $c \geq 0$ such that

$$d(x) = \sum_{j=0}^{q}\lambda_j u_j(x) + c = \sum_{j=1}^{q}\tilde{\lambda}_j u_j(x) + c,$$

with $\tilde{\lambda}_j = \lambda_j \geq 0$ for $j \neq i$ and $\tilde{\lambda}_i = \lambda_i - \lambda_0$. $\qquad\square$

Lemma A.3. *Assume Q is minimal. It holds that $\partial\mathscr{X}_i \neq \emptyset$ for all $i \in Q$.*

Proof. Fix $i \leq q$. By minimality of Q we can choose $x \in \mathbb{R}^p$ such that $u_i(x) < 0$ and $u_j(x) \geq 0$ for all $j \neq i$. Since $\mathscr{X} \neq \emptyset$, we can choose $y \in \mathscr{X}$. Then $u_j(y) \geq 0$ for all j. For $t \in [0, 1]$ it holds that

$$u_j(tx + (1 - t)y) = tu_j(x) + (1 - t)u_j(y),$$

which is non negative for $j \neq i$. For $t = u_i(y)/(u_i(y) - u_i(x))$ we have $u_i(tx + (1 - t)y) = 0$, so $tx + (1 - t)y \in \partial\mathscr{X}_i$. $\qquad\square$

Proposition A.4. *Assume Q is minimal. Suppose $\partial\mathscr{X}_i \subset \{d = 0\}$ for some $i \in Q$. Then there exists $\lambda_i \in \mathbb{R}$ such that $v(x) = \lambda_i u_i(x)$ for $x \in \mathscr{X}$. If $\mathscr{X}° \neq \emptyset$, then $v(x) = \lambda_i u_i(x)$ for all $x \in \mathbb{R}^p$.*

Proof. We have $\partial\mathscr{X}_i \subset \{v \geq 0\}$ and $\partial\mathscr{X}_i \subset \{-v \geq 0\}$. Applying Proposition A.2 with $d = v$, respectively, $d = -v$, we derive that

$$v(x) = \sum_{j=1}^{q}\lambda_j u_j(x) + c_1$$

$$-v(x) = \sum_{j=1}^{q}\mu_j u_j(x) + c_2,$$

for some $\lambda, \mu \in \mathbb{R}^{1\times p}$ with $\lambda_j, \mu_j \geq 0$ for $j \neq i$ and $c_1, c_2 \geq 0$. Adding the equations in the above display gives

$$0 = \sum_{j=1}^{q}(\lambda_j + \mu_j)u_j(x) + c_1 + c_2.$$

By Lemma A.3 we can choose $x \in \partial \mathscr{X}_i$ and deduce that $c_1 = c_2 = 0$. So

$$-(\lambda_i + \mu_i)u_i(x) = \sum_{j \neq i}(\lambda_j + \mu_j)u_j(x).$$

By minimality of Q we can choose $x \in \mathbb{R}^p$ such that $u_i(x) < 0$ and $u_j(x) \geq 0$ for all $j \neq i$. This gives that $c := \lambda_i + \mu_i \geq 0$. If $c > 0$, then for $x \in \mathscr{X}$ we have

$$0 \leq u_i(x) = -c^{-1}\sum_{j \neq i}(\lambda_j + \mu_j)u_j(x) \leq 0,$$

when $u_i(x) = 0$ for $x \in \mathscr{X}$. So $\mathscr{X} = \partial \mathscr{X}_i \subset \{v = 0\}$ and $v(x) = u_i(x) = 0$ for $x \in \mathscr{X}$. If $c = 0$, then $\sum_{j \neq i}(\lambda_j + \mu_j)u_j(x) = 0$ for all x. This holds in particular for $x \in \mathscr{X}$, that is, for x such that $u_j(x) \geq 0$ for all j. Hence, for $x \in \mathscr{X}$ we have $\lambda_j u_j(x) = \mu_j u_j(x) = 0$ for all $j \neq i$, so

$$v(x) = \sum_{j=1}^{q} \lambda_j u_j(x) + c_1 = \lambda_i u_i(x), \tag{A.4}$$

for $x \in \mathscr{X}$. If $\mathscr{X}° \neq \emptyset$, then choosing $x \in \mathscr{X}°$ gives $u_j(x) > 0$ for all j, which implies $\lambda_j = 0$ for all $j \neq i$. Then (A.4) holds for all $x \in \mathbb{R}^p$. $\qquad \square$