Compound matrices in systems and control theory: a tutorial

Eyal Bar-Shalom · Omri Dalin · Michael Margaliot

Received: 1 April 2022 / Accepted: 1 March 2023 / Published online: 15 March 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

Abstract
The multiplicative and additive compounds of a matrix play an important role in several fields of mathematics including geometry, multi-linear algebra, graph theory, combinatorics, and the analysis of nonlinear time-varying dynamical systems. There is a growing interest in applications of these compounds, and their generalizations, in systems and control theory. The goal of this tutorial paper is to provide a gentle and self-contained introduction to these topics with an emphasis on the geometric interpretation of the compounds, and to describe some of their recent applications including several non-trivial generalizations of positive systems, cooperative systems, contracting systems, and more.

Keywords Contracting systems · Diagonal stability · Positive systems · Cooperative systems · Chaotic systems · Sign variation diminishing property · Volume of parallelotopes · Hankel $k$-positive systems

This paper is dedicated to Eduardo Sontag on the occasion of his 70th birthday.

This research was partially supported by research grants from the ISF and DFG. An abridged version of this paper has been presented in the 2021 IEEE Conference on Decision and Control [6].

Michael Margaliot
michaelm@tauex.tau.ac.il

1 School of Electrical Engineering, Tel-Aviv University, 69978 Tel-Aviv, Israel
1 Introduction

Let $A \in \mathbb{C}^{n \times n}$. Fix $k \in \{1, \ldots, n\}$. The $k$-multiplicative and $k$-additive compounds of $A$, denoted $A^{(k)}$ and $A^{[k]}$, respectively, are $\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right) \times \left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)$ matrices that play an important role in several fields of applied mathematics. These matrices have an interesting spectral property. If $\lambda_i$, $i = 1, \ldots, n$, denote the eigenvalues of $A$, then the eigenvalues of $A^{(k)}$ are the $\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)$ products:

$$ \left\{ \prod_{j=1}^{k} \lambda_{i_j} : 1 \leq i_1 < \cdots < i_k \leq n \right\}, $$

and those of $A^{[k]}$ are the $\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)$ sums:

$$ \left\{ \sum_{j=1}^{k} \lambda_{i_j} : 1 \leq i_1 < \cdots < i_k \leq n \right\}. $$
In particular, $A^{(1)} = A^{[1]} = A$, so the eigenvalues of these matrices are just the $\lambda_i$s, $A^{(n)} = \det(A) = \prod_{i=1}^{n} \lambda_i$, and $A^{[n]} = \text{trace}(A) = \sum_{i=1}^{n} \lambda_i$.

Recently, there is a growing interest in the applications of these compounds, and their generalizations in systems and control theory (see, e.g., [1, 2, 9, 29–33, 38, 48, 77, 78, 81–84]).

This tutorial paper reviews the $k$-compounds, focusing on their geometric interpretation, and surveys some of their recent applications in systems and control theory, including the introduction and analysis of $k$-positive systems, $k$-cooperative systems, $k$-contracting systems, $k$-diagonal stability, and Hankel $k$-positive input/output systems.

This paper is organized as follows: The next section provides some geometric motivation by relating the evolution of the volume of parallelotopes under a linear time invariant (LTI) system to sums of eigenvalues such as in (2). Section 3 introduces the multiplicative compound and reviews some of its properties and, in particular, the fundamental role of the $k$-multiplicative compound in computing the volume of $k$-dimensional parallelotopes, and in establishing sign variation diminishing properties. Section 4 describes the additive compound, and this sets the stage to explaining the role of the compounds in ordinary differential equations (ODEs) in Sect. 5. The following section describes a general and important principle for what we call $k$-generalizations of dynamical systems. The idea is to take a dynamical property, e.g., contraction in a nonlinear system, and require that it holds for $k$-dimensional bodies rather than 1-dimensional bodies (i.e., lines). For $k = 1$, this reduces to standard contraction. We then demonstrate how this general principle leads to interesting and non-trivial generalizations of contracting, positive, cooperative, and diagonally stable systems to $k$-contracting, $k$-positive, $k$-cooperative, and $k$-diagonally stable systems. Section 11 reviews the recently introduced concept of $\alpha$-compounds, with $\alpha$ being a real number in $[1, n]$. For $\alpha \in (k, k+1)$, the $\alpha$-compound can be interpreted as an interpolation between the $k$ and $k+1$ compounds. We show that this leads to the notion of $\alpha$-contracting systems. Any attractor of such a system has a Hausdorff dimension smaller than $\alpha$. Section 12 reviews a type of $k$-positivity in systems with an input and output. The final section concludes and describes several directions for future research.

1.1 Notation

We use standard notation. The positive orthant in $\mathbb{R}^n$ is $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0, \ i = 1, \ldots, n\}$, and its interior is $\mathbb{R}^n_{++} := \{x \in \mathbb{R}^n : x_i > 0, \ i = 1, \ldots, n\}$. $\mathbb{Z}$ denotes the set of integers, and $\mathbb{N}_0$ is the subset of non-negative integers. For a set $S$, $\text{int}(S)$ denotes the interior of $S$. For scalars $\lambda_i, i \in \{1, \ldots, n\}$, $\text{diag}(\lambda_1, \ldots, \lambda_n)$ is the $n \times n$ diagonal matrix with diagonal entries $\lambda_i$. Vectors [matrices] are denoted by small [capital] letters. For a matrix $A$, $A^T$ is the transpose of $A$. For a square matrix $B$, $\text{trace}(B)$ [$\det(B)$] is the trace [determinant] of $B$. $B$ is tri-diagonal if $b_{ij} = 0$ for all $i, j$ with $|i - j| > 1$. $B$ is Metzler if all its off-diagonal entries are non-negative. A matrix is called Toeplitz [Hankel] if the entries along any diagonal [anti-diagonal] are equal. A matrix is called Hurwitz if the real part of every eigenvalue of the matrix is negative. A matrix is called Schur if the absolute value of every eigenvalue of the matrix is
smaller than 1. A square binary matrix that has exactly one entry of 1 in each row and each column, and 0s elsewhere is called a permutation matrix. A matrix $A \in \mathbb{R}^{n \times n}$ is called irreducible if there exists a permutation matrix $P \in \{0, 1\}^{n \times n}$ such that $P A P^T = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, where 0 is an $r \times (n-r)$ zero block matrix for some $1 \leq r \leq n-1$. A matrix is called reducible if it is not irreducible. Inequalities between matrices $A$, $B \in \mathbb{R}^{n \times m}$ are interpreted component-wise, e.g., $A \preceq B$ if $a_{ij} \preceq b_{ij}$ for every $i$, $j$, and $A \gg B$ if $a_{ij} > b_{ij}$ for every $i$, $j$.

Compound matrices require notation for specifying the minors of a matrix. Let $Q(k, n)$ denote all the $\binom{n}{k}$ increasing sequences of $k$ integers from the set $\{1, \ldots, n\}$, ordered lexicographically. For example,

$$Q(3, 4) = ((1, 2, 3), (1, 2, 4), (1, 3, 4), (2, 3, 4)).$$

Let $A \in \mathbb{C}^{n \times m}$. Fix $k \in \{1, \ldots, \min(n, m)\}$. For $\alpha \in Q(k, n)$, $\beta \in Q(k, m)$, let $A[\alpha|\beta]$ denote the $k \times k$ submatrix obtained by taking the entries of $A$ in the rows indexed by $\alpha$ and the columns indexed by $\beta$. For example,

$$A[(2, 3)|(1, 2)] = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

The minor of $A$ corresponding to $\alpha$, $\beta$ is

$$A(\alpha|\beta) := \det(A[\alpha|\beta]).$$

For example, if $m = n$ then $Q(n, n)$ includes the single element $\alpha = (1, \ldots, n)$, $A[\alpha|\alpha] = A$, and $A(\alpha|\alpha) = \det(A)$.

Recall that a vector norm $|\cdot| : \mathbb{R}^n \to \mathbb{R}_+$ induces a matrix norm $||\cdot|| : \mathbb{R}^{n \times n} \to \mathbb{R}_+$ defined by $||A|| := \max_{|x|=1} |A x|$, and a matrix measure (also known as logarithmic norm) $\mu(\cdot) : \mathbb{R}^{n \times n} \to \mathbb{R}$ defined by

$$\mu(A) := \lim_{\epsilon \downarrow 0} \frac{||I + \epsilon A|| - 1}{\epsilon}$$

(see, e.g., [70, 74]). For the $L_1$, $L_2$, and $L_\infty$ norms, there exist closed-form expressions for the induced matrix norms and matrix measures (see Table 1).

---

Table 1: Closed-form expressions for common matrix norms and matrix measures

| Vector norm | Induced matrix norm | Induced matrix measure |
|-------------|---------------------|-----------------------|
| $|x|_1 = \sum_j |x_j|$ | $||A||_1 = \max_j \sum_i |a_{ij}|$ | $\mu_1(A) = \max_j |a_{jj} + \sum_{j \neq j} |a_{ij}|$ |
| $|x|_2 = \sqrt{\sum_j |x_j|^2}$ | $||A||_2 = \sqrt{\lambda_1(A^T A)}$ | $\mu_2(A) = \frac{1}{2} \lambda_1(A + A^T)$ |
| $|x|_\infty = \max_j |x_j|$ | $||A||_\infty = \max_i \sum_j |a_{ij}|$ | $\mu_\infty(A) = \max_i |a_{ii} + \sum_{j \neq i} |a_{ij}|$ |

For a symmetric matrix $S$, $\lambda_1(S)$ denotes the largest eigenvalue of $S$. 

---

\( \sum \) Springer
2 Geometric motivation

$k$-compound matrices provide information on the evolution of $k$-dimensional paralleleptopes subject to a linear time-varying dynamics. To explain this in the simplest setting, consider the LTI system:

\[ \dot{x}(t) = \text{diag}(\lambda_1, \lambda_2, \lambda_3)x(t), \quad (3) \]

with \( \lambda_i \in \mathbb{R} \) and \( x : \mathbb{R}_+ \rightarrow \mathbb{R}^3 \). Let \( e^i, i = 1, 2, 3 \), denote the \( i \)-th canonical vector in \( \mathbb{R}^3 \). For \( x(0) = e^i \), we have \( x(t) = \exp(\lambda_i t)x(0) \). Thus, \( \exp(\lambda_i t) \) describes the rate of evolution of the line between 0 and \( e^i \) subject to (3). What about 2D areas? Let \( S_{ij} \subset \mathbb{R}^3 \) denote the square generated by \( e^i \) and \( e^j \), with \( i \neq j \). Then, \( S(t) := x(t, S_{ij}) \) is the rectangle generated by \( \exp(\lambda_i t)e^j \) and \( \exp(\lambda_j t)e^i \), so the area of \( S(t) \) is \( \exp((\lambda_i + \lambda_j)t) \). Similarly, if \( B_{123} \subset \mathbb{R}^3 \) is the 3D cube generated by \( e^1, e^2, \) and \( e^3 \), then the volume of \( B(t) := x(t, B_{123}) \) is \( \exp((\lambda_1 + \lambda_2 + \lambda_3)t) \) (see Fig. 1). Since \( \exp(At) = \text{diag}(\exp(\lambda_1 t), \exp(\lambda_2 t), \exp(\lambda_3 t)) \), this discussion suggests that it may be useful to have a \( 3 \times 3 \) matrix whose eigenvalues are the sums of any two eigenvalues of \( \exp(At) \), and a \( 1 \times 1 \) matrix whose eigenvalue is the sum of the three eigenvalues of \( \exp(At) \). With this geometric motivation in mind, we turn to review the multiplicative and additive compounds of a matrix. For more details and proofs, see, e.g., [23, Ch. 6] [62].

3 Multiplicative compound

The \( k \)-multiplicative compound of a matrix \( A \) is a matrix that collects all the \( k \)-minors of \( A \).

Definition 1 Let \( A \in \mathbb{C}^{n \times m} \) and fix \( k \in \{1, \ldots, \min(n, m)\} \). The \( k \)-multiplicative compound of \( A \), denoted \( A^{(k)} \), is the \( \binom{n}{k} \times \binom{m}{k} \) matrix that contains all the \( k \)-minors of \( A \) ordered lexicographically.

For example, if \( n = m = 3 \) and \( k = 2 \), then

\[
A^{(2)} = \begin{bmatrix}
A((1, 2)|(1, 2)) & A((1, 2)|(1, 3)) & A((1, 2)|(2, 3)) \\
A((1, 3)|(1, 2)) & A((1, 3)|(1, 3)) & A((1, 3)|(2, 3)) \\
A((2, 3)|(1, 2)) & A((2, 3)|(1, 3)) & A((2, 3)|(2, 3))
\end{bmatrix}.
\]

In particular, Definition 1 implies that \( A^{(1)} = A \), and if \( n = m \), then \( A^{(n)} = \det(A) \). Note also that by definition \( (A^T)^{(k)} = (A^{(k)})^T \). In particular if \( A \) is symmetric, i.e., \( A = A^T \), then \( (A^{(k)})^T = (A^T)^{(k)} = A^{(k)} \), so \( A^{(k)} \) is also symmetric.

3.1 The Cauchy–Binet formula and its implications

The next result, known as the Cauchy–Binet Formula (see, e.g., [28]), justifies the term multiplicative compound.
Fig. 1 Evolution of lines, areas, and volumes under the LTI (3) with $\lambda_1 > \lambda_2 > \lambda_3$. 

$B(t) = \exp(\lambda_1 + \lambda_2 + \lambda_3)t)B(0)$

$S(t) = \exp(\lambda_1 + \lambda_2)t)S(0)$

$L(t) = \exp(\lambda_1)t)L(0)$
Theorem 1 Let $A \in \mathbb{C}^{n \times m}, B \in \mathbb{C}^{m \times p}$. Fix $k \in \{1, \ldots, \min(n, m, p)\}$. Then,

$$(AB)^{(k)} = A^{(k)}B^{(k)}. \tag{4}$$

For $n = m = p$, Eq. (4) with $k = n$ reduces to the familiar formula

$$\det(AB) = \det(A) \det(B). \tag{5}$$

For the sake of completeness, we include a proof of Theorem 1 in Appendix. We now describe some immediate implications of the Cauchy–Binet Formula. Let $I_s$ denote the $s \times s$ identity matrix. Definition 1 implies that $I_s^{(k)} = I_r$, where $r := \binom{n}{k}$. Hence, if $A \in \mathbb{R}^{n \times n}$ is non-singular, then $(AA^{-1})^{(k)} = (A^{-1}A)^{(k)} = I_r$ and combining this with (4) yields $(A^{-1})^{(k)} = (A^{(k)})^{-1}$. In particular, if $A$ is non-singular, then so is $A^{(k)}$. Another implication of (4) is that if $T \in \mathbb{R}^{n \times n}$ is non-singular, then

$$(TAT^{-1})^{(k)} = T^{(k)}A^{(k)}(T^{-1})^{(k)} = T^{(k)}A^{(k)}(T^{(k)})^{-1}. \tag{6}$$

The Cauchy–Binet formula also yields a closed-form expression for the spectral properties of the $k$-multiplicative compound $A^{(k)}$ in terms of the spectral properties of $A$.

Proposition 1 For $A \in \mathbb{C}^{n \times n}$, let $\lambda_i$, $i = 1, \ldots, n$, denote the eigenvalues of $A$. Fix $k \in \{1, \ldots, n\}$. Then, the eigenvalues of $A^{(k)}$ are all the $\binom{n}{k}$ products:

$$\left\{ \prod_{\ell=1}^{k} \lambda_{i_{\ell}} : 1 \leq i_1 < i_2 < \cdots < i_k \leq n \right\}. \tag{7}$$

Furthermore, if for every eigenvalue $\lambda_i$, there exists a corresponding eigenvector $v^i$ and

$$[v^{i_1} \ldots v^{i_k}]^{(k)}$$

is not the zero vector, then it is the eigenvector of $A^{(k)}$ corresponding to the eigenvalue $\prod_{\ell=1}^{k} \lambda_{i_{\ell}}$.

To prove Proposition 1 note that

$$A \begin{bmatrix} v^{i_1} & \cdots & v^{i_k} \end{bmatrix} = \begin{bmatrix} v^{i_1} & \cdots & v^{i_k} \end{bmatrix} \text{diag}(\lambda_{i_1}, \ldots, \lambda_{i_k}),$$

and applying the Cauchy–Binet formula gives

$$A^{(k)} \begin{bmatrix} v^{i_1} & \cdots & v^{i_k} \end{bmatrix}^{(k)} = \begin{bmatrix} v^{i_1} & \cdots & v^{i_k} \end{bmatrix}^{(k)} (\text{diag}(\lambda_{i_1}, \ldots, \lambda_{i_k}))^{(k)} = \left( \prod_{\ell=1}^{k} \lambda_{i_{\ell}} \right) \begin{bmatrix} v^{i_1} & \cdots & v^{i_k} \end{bmatrix}^{(k)}.$$
Example 1 Suppose that \( n = 3 \) and \( A \) is upper-triangular: \( A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \). Then, the eigenvalues of \( A \) are \( \lambda_i = a_{ii}, i = 1, 2, 3 \). Assume that \( c := a_{22} - a_{11} \neq 0 \). Then, the first two corresponding eigenvectors are \( v^1 = [1 \ 0 \ 0]^T, v^2 = [a_{12} \ c \ 0]^T \), and \( v^1 \) and \( v^2 \) are linearly independent. Now,

\[
\begin{bmatrix} v^1 & v^2 \end{bmatrix}^{(2)} = \begin{bmatrix} 1 & a_{12} \\ 0 & c \\ 0 & 0 \end{bmatrix}^{(2)} = c \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.
\]

A direct calculation gives

\[
A^{(2)} = \begin{bmatrix} a_{11}a_{22} & a_{11}a_{23} & a_{12}a_{23} - a_{13}a_{22} \\ 0 & a_{11}a_{33} & a_{12}a_{33} \\ 0 & 0 & a_{22}a_{33} \end{bmatrix},
\]

so the eigenvalues of \( A^{(2)} \) are of the form (7), and the first eigenvector of \( A^{(2)} \) is indeed \( \begin{bmatrix} v^1 & v^2 \end{bmatrix}^{(2)} \).

The next result describes another simple and useful application of the multiplicative compound.

Proposition 2 Suppose that \( X, Y \in \mathbb{C}^{n \times k} \), with \( k \leq n \). Then,

\[
\det(Y^T X) = (Y^{(k)})^T X^{(k)}.
\]

Note that \( X^{(k)} \) (and \( Y^{(k)} \)) has dimensions \( (n \choose k) \times (k \choose k) \), i.e., it is a column vector, so the right-hand side of (8) is the inner product of two vectors. In the particular case where \( k = n \), Eq. (8) reduces to \( \det(Y^T X) = \det(Y) \det(X) \). However, (8) holds also for non-square matrices.

Proof Using the fact that \( Y^T X \in \mathbb{R}^{k \times k} \) and the Cauchy–Binet formula yields

\[
\det(Y^T X) = (Y^T X)^{(k)} = (Y^{(k)})^T X^{(k)} = (Y^{(k)})^T X^{(k)},
\]

and this completes the proof.

Proposition 2 has found many applications. For example, Ref. [10] applied it to determine the sensitivity of the natural modes of an electrical circuit to modifications in the circuit elements and topology. As we will see in the next section, Proposition 2 also implies that the \( k \)-multiplicative compound can be used to describe the volume of \( k \)-dimensional parallelotopes.
Another implication of the Cauchy–Binet Formula is that certain sign properties of the minors of a matrix are preserved under matrix multiplication. To explain this, we require the following definition.

**Definition 2** [20] Let \( A \in \mathbb{R}^{n \times m} \), and fix \( r \in \{1, \ldots, \min(n, m)\} \). Then, \( A \) is called **totally non-negative of order** \( r \) (TN\(r\)) if every \( s \times s \) minor with size \( s \leq r \) of \( A \) is non-negative. \( A \) is called **totally positive of order** \( r \) (TP\(r\)) if every \( s \times s \) minor with size \( s \leq r \) of \( A \) is positive. \( A \in \mathbb{R}^{n \times m} \) is called **totally positive** (TP) if all its minors are positive.

Equivalently, \( A \in \mathbb{R}^{n \times m} \) is TP if it is TP\(r\) with \( r = \min(n, m) \). Another equivalent definition is that for any \( k \in \{1, \ldots, \min(n, m)\} \) the \( k \)-multiplicative compound matrix \( A^{(k)} \) has positive entries. For example, for \( A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \), we have \( A^{(1)} = A \gg 0 \) and \( A^{(2)} = \begin{bmatrix} 5 & 7 & 1 \end{bmatrix} \gg 0 \), so \( A \) is TP. If \( n = m \), i.e., \( A \) is square, then \( A \) is TP if and only if \( A^{(k)} \) maps \( \mathbb{R}^{(k)}_+ \setminus \{0\} \) to \( \mathbb{R}^{(k)}_+ \) for every \( k \in \{1, \ldots, n\} \).

Such matrices admit sign variations diminishing properties that have important applications to dynamical systems (see Sect. 8 below). In general, TN\(r\) and TP\(r\) are not preserved under natural matrix operations.

**Example 2** Consider the matrices \( A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \) and \( B = I_3 \). A direct calculation of all the minors shows that both these matrices are TN\(3\). However, \( A + B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \) admits a minor of order two that is negative, as \( \det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} = -1 \), so \( A + B \) is not TN\(3\) (and not even TN\(2\)).

However, the Cauchy–Binet Formula immediately implies that the product of two TN\(r\) [TP\(r\)] matrices is TN\(r\) [TP\(r\)].

### 3.2 Multiplicative compounds and the volume of parallelotopes

We now review an important interpretation of the \( k \)-multiplicative compound based on the presentation in [25, Chapter IX] and [27]. For a vector \( v \in \mathbb{R}^n \), let \( ||v||_2 := \sqrt{v^Tv} \) denote the \( L_2 \) norm of \( v \). Fix \( k \in \{1, \ldots, n\} \) and vectors \( x^1, \ldots, x^k \in \mathbb{R}^n \). The parallelotope generated by these vectors (and the zero vertex) is the set:

\[
P := \left\{ \sum_{i=1}^{k} r_i x^i : r_i \in [0, 1] \right\}.
\]

Note that this implies that \( 0 \in \mathbb{R}^n \) is a vertex of \( P \). This can always be assured by a simple translation. We can also interpret \( P \) as the image of the unit \( k \)-cube under the
The volume of the parallelotope $P(x^1, x^2)$ is $\text{vol}(P(x^1))h = |x^1|^2h$.

The Gram matrix associated with $x^1, \ldots, x^k$ is the $k \times k$ symmetric matrix:

$$G(x^1, \ldots, x^k) := X^T X$$

$$= \begin{bmatrix}
(x^1)^T x^1 & (x^1)^T x^2 & \cdots & (x^1)^T x^k \\
(x^2)^T x^1 & (x^2)^T x^2 & \cdots & (x^2)^T x^k \\
\vdots & \vdots & \ddots & \vdots \\
(x^k)^T x^1 & (x^k)^T x^2 & \cdots & (x^k)^T x^k
\end{bmatrix}. \quad (9)$$

For example, for $k = 1$, we have $G(x^1) = |x^1|^2$, and for $k = 2$ we have

$$G(x^1, x^2) = \begin{bmatrix}
|x^1|^2_2 & (x^1)^T x^2 \\
(x^2)^T x^1 & |x^2|^2_2
\end{bmatrix}.$$ 

It follows from (9) that for any $s \in \mathbb{R}^k$, we have

$$\left| \sum_{i=1}^{k} s_i x^i \right|^2_2 = s^T G(x^1, \ldots, x^k) s,$$

so $G(x^1, \ldots, x^k)$ is non-negative definite, and it is positive-definite iff $x^1, \ldots, x^k$ are linearly independent.

The volume of $P(x^1, \ldots, x^k)$ is denoted $\text{vol}(P(x^1, \ldots, x^k))$. For $k = 1$, the parallelotope is just the line connecting the origin and $x^1$, so $\text{vol}(P(x^1)) = |x^1|^2$. For $k = 2$, the parallelotope $P(x^1, x^2)$ is depicted in Fig. 2. In general, $\text{vol}(P(x^1, \ldots, x^k))$ is defined in a recursive manner.

**Definition 3** For $k = 1$, $\text{vol}(P(x^1)) := |x^1|^2$. For any $k > 1$,

$$\text{vol}(P(x^1, \ldots, x^k)) := \text{vol}(P(x^1, \ldots, x^{k-1}))h, \quad (10)$$

where $P(x^1, \ldots, x^{k-1})$ is the $(k-1)$-dimensional “base” of $P(x^1, \ldots, x^k)$, and the altitude $h$ is the distance from $x^k$ to the base (see Fig. 3).

The next result relates $\text{Vol}(P)$ to the $k$-multiplicative compound of the matrix $X$. 
Fig. 3 The volume of the parallelotope \( P(x^1, \ldots, x^k) \) is \( \text{vol}(P(x^1, \ldots, x^{k-1}))h \).

**Proposition 3** The volume of \( P(x^1, \ldots, x^k) \) satisfies:

\[
\text{vol}(P(x^1, \ldots, x^k)) = |X^{(k)}|_2. \tag{11}
\]

Note that since \( X \in \mathbb{R}^{n \times k} \), \( X^{(k)} \) is a column vector.

**Proof** We first prove a simple algebraic expression for the volume in terms of the Gram matrix, namely,

\[
\text{vol}(P(x^1, \ldots, x^k)) = \sqrt{\det(G(x^1, \ldots, x^k))).} \tag{12}
\]

To prove this, let \( y \) denote the “foot” of the altitude \( h \), that is,

\[
y = \sum_{i=1}^{k-1} r_i x^i, \quad \text{with } r_i \in \mathbb{R}, \tag{13}
\]

and

\[
0 = (x^k - y)^T x^j, \quad j = 1, \ldots, k - 1. \tag{14}
\]

Then, the altitude \( h \) satisfies

\[
h^2 = |x^k - y|_2^2 = (x^k - y)^T (x^k - y) = (x^k - y)^T x^k. \tag{15}
\]

Equations (14) and (15) yield the linear system

\[
\begin{bmatrix}
(x^1)^T x^1 & (x^2)^T x^1 & \cdots & (x^{k-1})^T x^1 & 0 \\
(x^1)^T x^{k-1} & (x^2)^T x^{k-1} & \cdots & (x^{k-1})^T x^{k-1} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
(x^1)^T x^k & (x^2)^T x^k & \cdots & (x^{k-1})^T x^k & 1 \\
\end{bmatrix}
\begin{bmatrix}
r_1 \\
r_2 \\
\vdots \\
r_{k-1} \\
h^2
\end{bmatrix}
= \begin{bmatrix}
(x^1)^T x^1 \\
\vdots \\
(x^{k-1})^T x^{k-1} \\
(x^k)^T x^k
\end{bmatrix}. \tag{16}
\]
If \( x^1, \ldots, x^k \) are linearly dependent, that is, \( x^k \) lies in the \((k - 1)\)-dimensional subspace that contains the “base” of the parallelotope then \( h = 0 \), so (10) gives \( \text{vol}(P(x^1, \ldots, x^k)) = 0 \) and (12) gives the same value.

If \( x^1, \ldots, x^k \) are linearly independent, then applying Cramer’s rule \([34, \text{Ch. 0}]\) to (16) gives

\[
 h^2 = \frac{\det \begin{pmatrix} (x^1)^T x^1 & (x^2)^T x^1 & \cdots & (x^{k-1})^T x^1 & (x^k)^T x^1 \\ (x^1)^T x^k & (x^2)^T x^k & \cdots & (x^{k-1})^T x^k & (x^k)^T x^k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (x^1)^T x^{k-1} & (x^2)^T x^{k-1} & \cdots & (x^{k-1})^T x^{k-1} & (x^k)^T x^{k-1} \end{pmatrix} \det \begin{pmatrix} (x^1)^T x^1 & (x^2)^T x^1 & \cdots & (x^{k-1})^T x^1 & 0 \\ (x^1)^T x^k & (x^2)^T x^k & \cdots & (x^{k-1})^T x^k & 1 \end{pmatrix}}{
\det(G(x^1, \ldots, x^k)) \det(G(x^1, \ldots, x^{k-1}))}
\]

that is,

\[
 h^2 = \frac{\det(G(x^1, \ldots, x^k))}{\det(G(x^1, \ldots, x^{k-1}))},
\]

and combining this with (10) and Definition 3 proves (12). Now, using the fact that \( \det(G(x^1, \ldots, x^k)) = \det(X^T X) \), and using Proposition 2 proves (11).

Note that in the special case where \( k = n \), the matrix \( X \) is a square matrix, and (11) gives \((\text{vol}(P(x^1, \ldots, x^n)))^2 = (\det(X))^2\). Thus, for \( k = n \) we recover the well-known formula

\[
 \text{vol}(P(x^1, \ldots, x^n)) = |\det([x^1 \ldots x^n])|.
\]

### 3.3 Multiplicative compounds and the sign variation diminishing property

One reason for the usefulness of compound matrices in systems and control theory is that they allow to track the evolution of volumes along the solutions of differential equations. Another important application of compound matrices is in the context of sign variation diminishing properties. We review this topic in the classical setting of totally positive matrices \([20, 26, 57]\). Unfortunately, this field suffers from non-uniform notation. We follow the notation in [20].

#### 3.3.1 Totally positive matrices

TP matrices satisfy a beautiful sign variation diminishing property. For a vector \( x \in \mathbb{R}^n \setminus \{0\} \), let \( s^-(x) \) denote the number of sign variations in \( x \) after deleting all its zero entries. For example, \( s^-([{-1 \quad 0 \quad 0 \quad 2 \quad -3}]^T) = 2 \). We define \( s^-(0) := 0 \). For a vector \( x \in \mathbb{R}^n \), let \( s^+(x) \) denote the maximal possible number of
sign variations in $x$ after setting every zero entry in $x$ to either $-1$ or $+1$. For example, $s^+[\begin{bmatrix} -1 & 0 & 0 & 2 & -3 \end{bmatrix}^T] = 4$. These definitions imply that

$$0 \leq s^-(x) \leq s^+(x) \leq n - 1, \text{ for all } x \in \mathbb{R}^n.$$  \hfill (17)

There is a useful duality relation between $s^-$ and $s^+$ that is straightforward to prove. Let $D_{\pm} := \text{diag}(1, -1, 1, \ldots, (-1)^{n-1})$. Then,

$$s^-(x) + s^+(D_{\pm}x) = n - 1, \text{ for any } x \in \mathbb{R}^n.$$  \hfill (18)

The next result is useful when studying the number of sign variations in a vector that is the limit of a sequence of vectors.

**Proposition 4** [57, Chapter 3] Suppose that $x^1, x^2, \ldots$ is a sequence of vectors in $\mathbb{R}^n$ that converges to a limit $x^* := \lim_{i \to \infty} x^i$. Then,

$$s^-(x^*) \leq \liminf_{i \to \infty} s^-(x^i) \leq \limsup_{i \to \infty} s^+(x^i) \leq s^+(x^*).$$  \hfill (18)

Intuitively speaking, the number of sign changes in the limit vector $x^*$ can change only if there exists an index $k$ such that $x^*_k = 0$. Equation (18) follows from the fact that zero entries are ignored in $s^-$, but may lead to an increase in $s^+$. For example, for $x^i = [\begin{bmatrix} -1 & 2^i & -1 \end{bmatrix}^T$ we have $s^-(x^i) = 2$ for any $i$, $s^-(x^*) = 0$, and $s^+(x^*) = 2$.

The next result describes the sign variation diminishing property (SVDP) of TP matrices.

**Theorem 2** [58] Suppose that $A \in \mathbb{R}^{n \times n}$ is TP. Then, for any $x \in \mathbb{R}^n \setminus \{0\}$, we have

$$s^+(Ax) \leq s^-(x).$$  \hfill (19)

Furthermore, if

$$s^+(Ax) = s^-(x),$$  \hfill (20)

then the sign of the first [last] component of $Ax$ (if zero, the sign given in determining $s^+(Ax)$) agrees with the sign of the first [last] nonzero component of $x$.

In other words, multiplying a vector by a TP matrix can never increase the number of sign variations between $s^-(x)$ and $s^+(Ax)$, and if the number of sign changes remains the same, then in some sense $x$ and $Ax$ have the same “orientation”.

**Example 3** Consider $A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix}$, with

$b, c, d > 0$ and $d > bc$.  \hfill (21)

Then, $A$ is TP. Fix $x \in \mathbb{R}^2 \setminus \{0\}$. Then, one of the following cases holds.
**Case 1.** Suppose that \( s^-(x) = s^+(x) = 0 \). We may assume without loss of generality (wlog) that \( x_1, x_2 > 0 \). Then,

\[
y := Ax = \begin{bmatrix} x_1 + bx_2 \\ cx_1 + dx_2 \end{bmatrix}
\]
satisfies \( y_1, y_2 > 0 \), so \( s^+(y) = 0 = s^-(x) \), and the sign of the first [last] component of \( y \) agrees with the sign of the first [last] nonzero component of \( x \).

**Case 2.** Suppose that \( s^-(x) = 0 \) and \( s^+(x) = 1 \). We may assume wlog that \( x_1 = 0, x_2 = 1 \). Then,

\[
y = \begin{bmatrix} b \\ d \end{bmatrix}
\]
satisfies \( y_1, y_2 > 0 \), so \( s^+(y) = 0 = s^-(x) \), and the sign of the first [last] component of \( y \) agrees with the sign of the first [last] nonzero component of \( x \).

**Case 3.** Suppose that \( s^-(x) = s^+(x) = 1 \). We may assume wlog that \( x_1 = 1, x_2 < 0 \). Then,

\[
y = \begin{bmatrix} 1 + bx_2 \\ c + dx_2 \end{bmatrix}
\]
and (21) implies that \( y_2 < cy_1 \). If \( y_1 < 0 \), then \( y_2 < 0 \), so \( s^+(y) < s^-(x) \). If \( y_1 = 0 \), then \( y_2 < 0 \), so \( s^+(y) = s^-(x) \), and the sign given to \( y_1 \) in determining \( s^+(y) \) is plus. If \( y_1 > 0 \), then either: (1) \( y_2 > 0 \) and then, \( s^+(y) < s^-(x) \); or (2) \( y_2 = 0 \) and then, \( s^+(y) = s^-(x) \) and the sign given to \( y_2 \) in determining \( s^+(y) \) is minus; or (3) \( y_2 < 0 \) and then, \( s^+(y) = s^-(x) \).

Thus, we see that in each case the assertions in Theorem 2 hold.

As we will see in Sect. 8, the SVDP has important implications in the asymptotic analysis of linear and nonlinear dynamical systems [48].

### 3.3.2 Recognition of totally positive matrices

Since the minors of a matrix are not independent, verifying that a matrix is TP does not require checking that all minors are positive. This fact will play an important role in Sect. 12, so we review one result on the recognition of TP matrices. For more details, see [20].

**Definition 4** An index set \( \alpha \in Q(k, n) \) is called *contiguous* if it consists of only consecutive numbers, i.e., \( \alpha_1 = p, \alpha_2 = p + 1, \ldots, \alpha_k = p + k - 1 \) for some integer \( p \). If \( \alpha, \beta \in Q(k, n) \) are two contiguous index sets, then the corresponding submatrix \( A[\alpha | \beta] \) is called a *contiguous submatrix of* \( A \), and the \( k \)-minor \( A(\alpha | \beta) \) is called a *contiguous minor*. A contiguous minor is called *initial* if at least one of the two sets \( \alpha \) or \( \beta \) is \( \{1, 2, \ldots, k\} \).

\( \text{Springer} \)
The next result shows that if certain initial and contiguous minors are positive, then all minors are positive.

**Proposition 5** [21] Let $A \in \mathbb{R}^{m \times n}$. If all the initial minors of $A$ up to order $k - 1$ are positive, and all its contiguous $k$-minors are positive, then $A$ is $TP_k$.

**Example 4** Consider the matrix $A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \\ 1 & 3 & 10 \end{bmatrix}$. We apply Proposition 5 to verify that $A^{(2)} \gg 0$, i.e., that all the 2-minors of $A$ are positive. The initial minors up to order 1 are the entries in the first row and first column of $A$, and these are positive. The contiguous 2-minors are $A((1, 2) \mid (1, 2))$, $A((1, 2) \mid (2, 3))$, $A((2, 3) \mid (1, 2))$ and $A((2, 3) \mid (2, 3))$, which are 1, 1, 5, and 1, respectively, so we conclude that $A^{(2)} \gg 0$.

### 4 Additive compound

The matrix $\exp(At)$ describes the evolution of lines under the LTI dynamics $\dot{x} = Ax$. We will see below that, more generally, the evolution of $k$-dimensional parallelotopes is described by $(\exp(At))^{(k)}$. This naturally leads to the following question: what is the derivative with respect to (w.r.t.) time of $(\exp(At))^{(k)}$? To address this, we require the $k$-additive compound of $A$.

**Definition 5** Let $A \in \mathbb{C}^{n \times n}$. The $k$-additive compound matrix of $A$ is the $\binom{n}{k} \times \binom{n}{k}$ matrix defined by:

$$A^{[k]} := \frac{d}{d\epsilon} (I_n + \epsilon A)^{(k)}|_{\epsilon = 0}. \quad (22)$$

The derivative here is well-defined, as every entry of $(I_n + \epsilon A)^{(k)}$ is a polynomial in $\epsilon$. Note that this definition implies that

$$A^{[k]} = \frac{d}{d\epsilon} (\exp(A\epsilon))^{(k)}|_{\epsilon = 0}, \quad (23)$$

and also that

$$(I_n + \epsilon A)^{(k)} = I_r + \epsilon A^{[k]} + o(\epsilon), \quad (24)$$

where $r := \binom{n}{k}$, and $o(\epsilon)$ denotes a function $f(\epsilon)$ such that $\lim_{\epsilon \to 0} \frac{f(\epsilon)}{\epsilon} = 0$. In other words, $A^{[k]}$ is the coefficient of the first-order term in the Taylor expansion of $(I_n + \epsilon A)^{(k)}$. The definition also implies that

$$\frac{d}{dt} (\exp(At))^{(k)} = A^{[k]} (\exp(At))^{(k)} \quad (25)$$
(see the more general result in Proposition 8 below). Thus, \((\exp(At))^{(k)}\) satisfies an LTI with the matrix \(A^{[k]}\). Note that for \(k = 1\), Eq. (25) reduces to \(\frac{d}{dt} \exp(At) = A \exp(At)\), whereas for \(k = n\) it becomes \(\frac{d}{dt} \det(\exp(At)) = \text{trace}(A) \det(\exp(At))\).

Equation (22) and the Cauchy–Binet formula can be used to determine how \(A^{[k]}\) changes under a similarity transformation of \(A\). If \(T\) is non-singular, then

\[
(TAT^{-1})^{[k]} = \frac{d}{d\epsilon} (I_n + \epsilon TAT^{-1})^{(k)} |_{\epsilon=0} = \frac{d}{d\epsilon} (T(I_n + \epsilon A)T^{-1})^{(k)} |_{\epsilon=0} = T^{(k)} A^{[k]} (T^{(k)})^{-1}.
\]

The next result describes the spectral properties of the additive compound. Its proof follows from combining Proposition 1 and (24).

**Proposition 6** For \(A \in \mathbb{C}^{n \times n}\), let \(\lambda_i, i = 1, \ldots, n\), denote the eigenvalues of \(A\), and let \(v^i, i = 1, \ldots, n\), denote the eigenvector corresponding to \(\lambda_i\). Fix \(k \in \{1, \ldots, n\}\). Then, the eigenvalues of \(A^{[k]}\) are all the \(\binom{n}{k}\) sums:

\[
\left\{ \sum_{\ell=1}^{k} \lambda_{i_\ell} : 1 \leq i_1 < i_2 < \cdots < i_k \leq n \right\}.
\]

Furthermore, if

\[
[v^1 \ldots v^k]^{(k)}
\]

is not the zero vector, then it is the eigenvector of \(A^{[k]}\) corresponding the eigenvalue \(\sum_{\ell=1}^{k} \lambda_{i_\ell}\).

In particular, if \(A\) is positive-definite (so it is symmetric and all its eigenvalues are real and positive), then \(A^{[k]}\) is symmetric and all its eigenvalues are real and positive, so \(A^{[k]}\) is positive-definite.

Another important implication of the definitions above is that for any \(A, B \in \mathbb{C}^{n \times n}\) we have

\[
(A + B)^{[k]} = A^{[k]} + B^{[k]}.
\]

This justifies the term additive compound. Moreover, the mapping \(A \rightarrow A^{[k]}\) is linear. To prove (27), note that (24) gives

\[
I_r + \epsilon(A + B)^{[k]} = (I_n + \epsilon(A + B))^{(k)} + o(\epsilon) = ((I_n + \epsilon A)(I_n + \epsilon B))^{(k)} + o(\epsilon) = (I_n + \epsilon A)^{(k)}(I_n + \epsilon B)^{(k)} + o(\epsilon) = \left(I_r + \epsilon A^{[k]}\right)\left(I_r + \epsilon B^{[k]}\right) + o(\epsilon)
\]
\[ = I_r + \epsilon(A^{[k]} + B^{[k]}) + o(\epsilon) \]

and using the continuity of the mapping \( A \rightarrow A^{[k]} \) implies (27).

The next result gives a useful explicit formula for \( A^{[k]} \) in terms of the entries \( a_{ij} \) of \( A \). Recall that any entry of \( A^{(k)} \) is a minor \( A(\alpha|\beta) \). Thus, it is natural to index the entries of \( A^{(k)} \) and \( A^{[k]} \) using \( \alpha, \beta \in Q(k,n) \).

**Proposition 7** Fix \( \alpha, \beta \in Q(k,n) \) and let \( \alpha = \{i_1, \ldots, i_k\} \) and \( \beta = \{j_1, \ldots, j_k\} \). Then, the entry of \( A^{[k]} \) corresponding to \( (\alpha, \beta) \) is equal to:

1. \( \sum_{\ell=1}^{k} a_{i_i i_{\ell}} \) if \( i_\ell = j_\ell \) for all \( \ell \in \{1, \ldots, k\} \);
2. \( (-1)^{\ell+m} a_{i_{\ell} j_m} \), if all the indices in \( \alpha \) and \( \beta \) agree, except for a single index \( i_{\ell} \neq j_m \);
3. 0, otherwise.

Note that the first case in the proposition corresponds to the diagonal entries of \( A^{[k]} \). Also, the proposition implies in particular that \( A^{(n)} = \sum_{\ell=1}^{n} a_{\ell\ell} = \text{trace}(A) \).

We prove Proposition 7 for the case \( k = 2 \). The proof when \( k > 2 \) is similar. Let \( B := (I_n + \epsilon A)^{(2)} \). Denote

\[ \delta_{pq} := \begin{cases} 1, & \text{if } p = q, \\ 0, & \text{otherwise}. \end{cases} \]

Then, for any \( 1 \leq i_1 < i_2 \leq n \) and \( 1 \leq j_1 < j_2 \leq n \),

\[ B((i_1, i_2)|(j_1, j_2)) = b_{i_1 j_1} b_{i_2 j_2} - b_{i_1 j_2} b_{i_2 j_1} \]

\[ = (\delta_{i_1 j_1} + \epsilon a_{i_1 j_1})(\delta_{i_2 j_2} + \epsilon a_{i_2 j_2}) - (\delta_{i_1 j_2} + \epsilon a_{i_1 j_2})(\delta_{i_2 j_1} + \epsilon a_{i_2 j_1}) \]

\[ = \epsilon(\delta_{i_1 j_1} a_{i_2 j_2} + \delta_{i_2 j_2} a_{i_1 j_1} - \delta_{i_1 j_2} a_{i_2 j_1} - \delta_{i_2 j_1} a_{i_1 j_2}) + c + o(\epsilon), \]

where \( c \) does not depend on \( \epsilon \). Applying (24) yields

\[ A^{[2]}(\alpha|\beta) = \delta_{i_1 j_1} a_{i_2 j_2} + \delta_{i_2 j_2} a_{i_1 j_1} - \delta_{i_1 j_2} a_{i_2 j_1} - \delta_{i_2 j_1} a_{i_1 j_2}, \]

and this agrees with the expressions in Proposition 7.

**Example 5** For \( A \in \mathbb{R}^{4 \times 4} \) and \( k = 3 \), Proposition 7 yields

\[
A^{[3]} = \begin{bmatrix}
(1, 2, 3) & (1, 2, 4) & (1, 3, 4) & (2, 3, 4) \\
\{a_{11} + a_{22} + a_{33} \} & \{a_{34} - a_{24} \} & \{-a_{14} \} & \{a_{14} \} \\
\{a_{43} \} & \{a_{11} + a_{22} + a_{44} \} & \{a_{23} - a_{13} \} & \{-a_{12} \} \\
\{-a_{42} \} & \{a_{32} \} & \{a_{11} + a_{33} + a_{44} \} & \{a_{12} \} \\
\{a_{41} \} & \{-a_{31} \} & \{-a_{21} \} & \{a_{22} + a_{33} + a_{44} \}
\end{bmatrix}
\]

where the indexes \( \alpha \in Q(3, 4) [\beta \in Q(3, 4)] \) are marked on right-hand side [above] of the matrix. For example, the entry in the second row and fourth column of \( A^{[3]} \) corresponds to \( (\alpha|\beta) = ((1, 2, 4)|(2, 3, 4)) \). As \( \alpha \) and \( \beta \) agree in all indices except for the index \( i_1 = 1 \) and \( j_2 = 3 \), this entry is equal to \((-1)^{1+2}a_{13} = -a_{13}\).
Example 6 For $A \in \mathbb{R}^{3 \times 3}$ and $k = 2$, Proposition 7 yields

$$A^{[2]} = \begin{bmatrix} (1, 2) & (1, 3) & (2, 3) \\ a_{11} + a_{22} & a_{23} & -a_{13} \\ a_{32} & a_{11} + a_{33} & a_{12} \\ -a_{31} & a_{21} & a_{22} + a_{33} \end{bmatrix} \cdot (1, 2) \quad (1, 3) \quad (2, 3)$$

(28)

For example, the entry in the second row and third column of $A^{[3]}$ corresponds to $(\alpha | \beta) = ((1, 3)|(2, 3))$. As $\alpha$ and $\beta$ agree in all indices except for the index $i_1 = 1$ and $j_1 = 2$, this entry is equal to $(-1)^{1+1}a_{12} = a_{12}$.

The next section describes applications of compound matrices for dynamical systems described by ODEs. For more details and proofs, see [51, 62].

5 Compound matrices and ODEs

Fix a time interval $[\tau_0, \tau_1]$. Let $A : [\tau_0, \tau_1] \to \mathbb{R}^{n \times n}$ be a continuous matrix function, and consider the linear time-varying (LTV) system:

$$\dot{x}(t) = A(t)x(t), \quad x(\tau_0) = x_0.$$  

(29)

The solution at time $t$ is $x(t) = \Phi(t, \tau_0)x_0$, where $\Phi(t, \tau_0)$ is the solution at time $t$ of the matrix differential equation

$$\frac{d}{ds} \Phi(s) = A(s)\Phi(s), \quad \Phi(\tau_0) = I_n.$$  

(30)

Fix $k \in \{1, \ldots, n\}$ and let $r := \binom{n}{k}$. A natural question is: how do the $k$-order minors of $\Phi(t)$ evolve in time?

5.1 $k$-Compound system

The next result provides an elegant formula for the evolution of $\Phi^{(k)}(t) := (\Phi(t))^{(k)}$.

Proposition 8 If $\Phi$ satisfies (30), then

$$\frac{d}{ds} \Phi^{(k)}(s) = A^{[k]}(s)\Phi^{(k)}(s), \quad \Phi^{(k)}(\tau_0) = I_r,$$  

(31)

where $A^{[k]}(s) := (A(s))^{[k]}$.

Proof For any $\varepsilon > 0$, we have

$$\Phi^{(k)}(s + \varepsilon) = (\Phi(s) + \varepsilon A(s)\Phi(s) + o(\varepsilon))^{(k)}$$

$$= ((I_n + \varepsilon A(s))\Phi(s))^{(k)} + o(\varepsilon)$$
\[
= (I_n + \varepsilon A(s))^k \Phi^k(s) + o(\varepsilon)
= (I + \varepsilon A[k](s))\Phi^k(s) + o(\varepsilon),
\]
where the last step follows from using (24). Thus,
\[
\frac{\Phi^k(s + \varepsilon) - \Phi^k(s)}{\varepsilon} = A[k](s)\Phi^k(s) + o(\varepsilon),
\]
and this completes the proof.

Equation (31) is the \textit{k-compound system} of (30) and is the basis of all the \textit{k}-generalizations of dynamical systems described in this paper.

The dynamics of the \textit{k}-compound system implies that \(k\)-minors of \(\Phi\) also satisfy an LTV. In particular, if \(A(t) \equiv A\) and \(\tau_0 = 0\), then \(\Phi(t) = \exp(At)\) so \(\Phi^k(t) = (\exp(At))^k\), and (31) gives
\[
(\exp(At))^k = \exp(A[k]t).
\]
This identity has interesting implications. For example, recall that \(\exp(Bt) \geq 0\) (where the inequality is component-wise) for all \(t \geq 0\) iff \(B\) is Metzler. Equation (32) implies that \((\exp(At))^k \geq 0\) for all \(t \geq 0\) iff \(A[k]\) is Metzler. This is the basis for the notion of \(k\)-positive systems described in Sect. 8.

Equation (32) also has some Lie-algebraic implications [71]. The next result describes one such application.

\textbf{Lemma 1} Let \(A, B \in \mathbb{R}^{n \times n}\). Then,
\[
(BA - AB)^{[k]} = B^{[k]}A^{[k]} - A^{[k]}B^{[k]}.
\]
In particular, if \(A\) and \(B\) commute, then \(A^{[k]}\) and \(B^{[k]}\) commute.

\textbf{Proof} For any \(t \in \mathbb{R}\), we have
\[
(\exp(-At)B \exp(At))^{[k]} = (\exp(-At))^{[k]} B^{[k]}(\exp(At))^{[k]}
= \exp(-A^{[k]}t)B^{[k]} \exp(A^{[k]}t).
\]

Expanding both sides as a Taylor series gives
\[
B^{[k]} + (BA - AB)^{[k]}t + o(t) = B^{[k]} + (B^{[k]}A^{[k]} - A^{[k]}B^{[k]})t + o(t),
\]
and this completes the proof.

Note also that for the special case \(k = n\), Proposition 8 yields
\[
\frac{d}{dt} \det(\Phi(t)) = \text{trace}(A(t)) \det(\Phi(t)),
\]

\text{Springer}
which is the Abel–Jacobi–Liouville identity (see, e.g., [15, p. 152]).

Proposition 8 implies that under the LTV dynamics (29), \( k \)-dimensional paralleloptopes evolve according to the dynamics (31). Indeed, consider \( k \) initial conditions \( a^1, \ldots, a^k \in \mathbb{R}^n \). Under the LTV dynamics, and assuming for simplicity that \( t_0 = 0 \), the corresponding solutions satisfy \( x(t, a^i) = \Phi(t)a^i \). Let \( X(t) := [x(t, a^1) \ldots x(t, a^k)] \). Then, the volume of the paralleloptope \( P(x(t, a^1), \ldots, x(t, a^k)) \) is \( |X(k)(t)|_2 \). Now,

\[
X(k)(t) = [\Phi(t)a^1 \ldots \Phi(t)a^k]^{(k)} \\
= (\Phi(t) [a^1 \ldots a^k])^{(k)} \\
= \Phi(k)(t) [a^1 \ldots a^k]^{(k)} \\
= \Phi(k)(t)X(k)(0),
\]

(33)

where the third equation follows from the Cauchy Binet formula. Using (31) gives

\[
\frac{d}{dt} \left( X(k)(t) \right) = A_{ab}^{[k]}(t)X(k)(0) \\
= A_{ab}^{[k]}(t)X(k)(t).
\]

(34)

Note that this implies in particular that if \( X(k)(0) \neq 0 \) then \( X(k)(t) \neq 0 \) for all \( t \). In other words, if \( a^1, \ldots, a^k \) are linearly independent then \( x(t, a^1), \ldots, x(t, a^k) \) are linearly independent for all \( t \).

The compounds can also be used in the analysis of nonlinear dynamical systems. Consider the time-varying system

\[
\dot{x}(t) = f(t, x).
\]

(35)

For the sake of simplicity, we assume that the initial time is zero, and that the system admits a convex and compact state-space \( \Omega \). We also assume that \( f \in C^1 \). The Jacobian of the vector field \( f \) is \( J(t, x) := \frac{\partial}{\partial x} f(t, x) \). Compound matrices can be used to analyze (35) by using an LTV called the variational equation associated with (35). To define it, fix \( a, b \in \Omega \). Let \( z(t) := x(t, a) - x(t, b) \), and for \( s \in [0, 1] \), let \( \gamma(s) := sx(t, a) + (1-s)x(t, b) \), i.e., the line connecting \( x(t, a) \) and \( x(t, b) \). Then,

\[
\dot{z}(t) = f(t, x(t, a)) - f(t, x(t, b)) \\
= \int_0^1 \frac{\partial}{\partial s} f(t, \gamma(s)) ds,
\]

and this gives the variational equation:

\[
\dot{z}(t) = A_{ab}^{[k]}(t)z(t),
\]

(36)
where

\[ A^{ab}(t) := \int_0^1 J(t, \gamma(s)) ds. \] (37)

Note that the variational equation (36) is an LTV.

In the next section, we use the results above to describe a general principle for deriving useful “k-generalizations” of important classes of nonlinear dynamical systems including cooperative systems [66], contracting systems [3, 14, 44], and diagonally stable systems [37].

6 k-generalizations of dynamical systems

Suppose that the LTV (29) satisfies a specific property guaranteed by a related condition on A(t). For example, property may be that the LTV is positive (i.e., A(t) is Metzler for all t) or that the LTV is contracting (guaranteed if \( \mu(A(t)) \leq -\eta < 0 \) for all t, where \( \mu : \mathbb{R}^{n \times n} \to \mathbb{R} \) is some matrix measure). Fix \( k \in \{1, \ldots, n\} \). We say that the LTV satisfies k-property if the k-compounded system satisfies the related property. For example, the LTV is k-positive if \( A^{[k]}(t) \) is Metzler for all t; the LTV is k-contracting if \( \mu(A^{[k]}(t)) \leq -\eta < 0 \) for all t, and so on.

This generalization principle was first suggested in [6]. It makes sense for two reasons. First, when \( k = 1 \), the k-compounded system reduces to the original system, so k-property is clearly a generalization of property. Second, the k-compounded system has a clear geometric meaning: it describes the evolution of k-dimensional parallelotopes along the dynamics. Also, the k-compounded system describes the evolution of k minors, and the sign of these minors is important in establishing an SVDP.

The same principle can be applied to the nonlinear system (35) using the variational equation (36). For example, if \( \mu(J(t,x)) \leq -\eta < 0 \) for all \( t \geq 0 \) and all \( x \in \Omega \), then (35) is contracting: the distance between any two solutions (in the norm that induced \( \mu \)) decays at an exponential rate. If we replace this by the condition \( \mu(J^{[k]}(t,x)) \leq -\eta < 0 \) for all \( t \geq 0 \) and all \( x \in \Omega \) (i.e., the same condition but now for the k-compounded system), then (35) is called k-contracting. Roughly speaking, this means that the volume of k-dimensional parallelotopes decays to zero exponentially along the flow of the nonlinear system. We now turn to describe several such k-generalizations in more detail.

7 k-contracting systems

The term k-contracting systems was introduced in [82]. For \( k = 1 \), these reduce to contracting systems. This generalization of contracting systems is motivated in part by the seminal work of Muldowney [51] who analyzed nonlinear systems that, in the new terminology, are 2-contracting (see also the unpublished preprint [46] for some preliminary ideas). Muldowney derived several interesting results for 2-contracting systems. For example, every bounded trajectory of a \textit{time-invariant} 2-contracting
system converges to an equilibrium (but, unlike in the case of contracting systems, the equilibrium is not necessarily unique). These results have found many applications in models from epidemiology, see, e.g., [41]. Such models typically have two equilibrium points corresponding to the disease free and the endemic steady-states. The existence of more than a single equilibrium point implies that the system is not contracting (that is, not 1-contracting) w.r.t. any norm.

In this tutorial, we focus on explaining \( k \)-contraction in LTVs. The case of nonlinear systems follows by applying similar ideas to the variational equation (36), see [82].

**Definition 6** [82] Fix \( k \in \{1, \ldots, n\} \). The LTV (29) (with initial time \( \tau_0 = 0 \)) is called \( k \)-contracting if there exist \( \eta > 0 \) and a vector norm \( |\cdot| \) such that for any \( a^1, \ldots, a^k \in \mathbb{R}^n \), the mapping \( X(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times k} \) defined by \( X(t) := [x(t, a^1) \ldots x(t, a^k)] \) satisfies

\[
|X^{(k)}(t)| \leq \exp(-\eta t)|X^{(k)}(0)|, \quad \text{for all } t \geq 0.
\]  

(38)

In other words, the volume of any \( k \)-parallelotope exponentially decays to zero under the dynamics. Another interpretation is that the initial condition \( X^{(k)}(0) = [a^1 \ldots a^k]^{(k)} \) is “forgotten”.

For \( k = 1 \), Eq. (38) reduces to the requirement that \( |x(t, a)| \leq \exp(-\eta t)|a| \) for any \( a \in \mathbb{R}^n \), and since the dynamics is linear, this implies that \( |x(t, a) - x(t, b)| \leq \exp(-\eta t)|a - b| \) for any \( a, b \in \mathbb{R}^n \), i.e., standard contraction.

Equation (33) implies that \( k \)-contraction is equivalent to

\[
||\Phi^{(k)}(t)|| \leq \exp(-\eta t)||\Phi^{(k)}(0)|| = \exp(-\eta t), \quad \text{for all } t \geq 0.
\]

Using Coppel’s inequality [16] and (34) provides a simple sufficient condition for \( k \)-contraction in terms of the \( k \) additive compound of \( A \).

**Proposition 9** If there exist \( \eta > 0 \) and a matrix measure \( \mu : \mathbb{R}^{n \times n} \rightarrow \mathbb{R} \) such that

\[
\mu(A^{[k]}(t)) \leq -\eta < 0, \quad \text{for all } t \geq 0,
\]

(39)

then (29) is \( k \)-contracting.

Note that for \( k = 1 \) condition (39) reduces to the standard infinitesimal condition for contraction [3]. For \( k = n \), condition (39) becomes trace\( (A(t)) \leq -\eta < 0 \) for all \( t \geq 0 \). It was shown in [82] that if (39) holds for some \( L_p \) norm, with \( p \in \{1, 2, \infty\} \) then for any integer \( \ell \geq k \) we have \( \mu(A^{[\ell]}(t)) \leq -\eta < 0 \) for all \( t \geq 0 \), so the system is also \( \ell \)-contracting w.r.t. the \( L_p \) norm (see Fig. 4).
Example 7 Consider the LTI system $\dot{x} = Ax$ with $A = \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}$, $c \in \mathbb{R}$. Since $\frac{d}{dt}(x_1^2(t) + x_2^2(t)) \equiv 0$, the solution for any initial condition $x(0)$ is a circle with radius $r := \sqrt{x_1^2(t) + x_2^2(t)} \equiv \sqrt{x_1^2(0) + x_2^2(0)}$. Fix $a^1, a^2 \in \mathbb{R}^2$ and consider $X(t) := [x(t, a^1) \ x(t, a^2)]$. Equation (34) gives

$$\frac{d}{dt} \left( X^{(2)}(t) \right) = A^{[2]} X^{(2)}(t) = \text{trace}(A)X^{(2)}(t) = 0.$$ 

This shows that the LTI is “on the verge” of being 2-contracting. This makes sense because the matrix $A$ represents a purely rotational dynamics, so the area of the parallelogram generated by $x(t, a^1)$ and $x(t, a^2)$ is a constant of time for any value $c$.

For the $L_p$ norms, with $p \in \{1, 2, \infty\}$, combining Proposition 7 with the expressions in Table 1 provides the following explicit expressions for $\mu_p(A^{[k]})$ [51]:

$$\mu_1(A^{[k]}) = \max_{\alpha \in Q(k,n)} \left( \sum_{p=1}^k a_{\alpha_p, \alpha_p} + \sum_{j \notin \alpha} (|a_j, \alpha_1| + \cdots + |a_j, \alpha_k|) \right),$$

$$\mu_2(A^{[k]}) = \sum_{i=1}^k \lambda_i(A + A^T)/2,$$

$$\mu_\infty(A^{[k]}) = \max_{\alpha \in Q(k,n)} \left( \sum_{p=1}^k a_{\alpha_p, \alpha_p} + \sum_{j \notin \alpha} (|a_{\alpha_1, j}| + \cdots + |a_{\alpha_k, j}|) \right),$$

where $\lambda_i(A + A^T)$, $i = 1, \ldots, n$, denote the eigenvalues of the symmetric matrix $A + A^T$ ordered such that $\lambda_1 \geq \cdots \geq \lambda_n$.

Example 8 Consider the LTV (29) with $n = 2$, $\tau_0 = 0$, and $A(t) = \begin{bmatrix} -1 & 0 \\ -2 \cos(t) & 0 \end{bmatrix}$. 

$\text{Springer}$
The corresponding transition matrix is: \( \Phi(t) = \begin{bmatrix} \exp(-t) & 0 \\ -1 + \exp(-t)(\cos(t) - \sin(t)) & 1 \end{bmatrix} \).

This implies that the LTV is uniformly stable, and that for any \( x(0) \in \mathbb{R}^2 \) we have

\[
\lim_{t \to \infty} x(t, x(0)) = \begin{bmatrix} 0 \\ x_2(0) - x_1(0) \end{bmatrix}.
\]

The LTV is not contracting w.r.t. any norm, as it admits more than a single equilibrium. However, \( A[2](t) = \text{trace}(A(t)) \equiv -1 \), so the system is 2-contracting. Let \( S_0 \subset \mathbb{R}^2 \) denote the unit square, and let \( S(t) := x(t, S_0) \), that is, the evolution at time \( t \) of the unit square under the dynamics. Figure 5 depicts \( S(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} 2t \) for several values of \( t \geq 0 \), where the shift by \( 2t \) is for the sake of clarity. It may be seen that the area of \( S(t) \) decays with \( t \), and that \( S(t) \) converges to a line.

Note that convergence to sets that are more general than an equilibrium point is a common requirement in applications. For example, in consensus algorithms [49] the requirement is that the state \( x(t) \) converges to the subspace \( \text{span}(\begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}^T) \).

The next result describes another interesting implication of \( k \)-contraction of an LTV.

**Proposition 10** [51] Suppose that the LTV (29) is uniformly stable. Fix \( k \in \{1, \ldots, n\} \). The following conditions are equivalent:

(a) The LTV admits an \((n - k + 1)\)-dimensional linear subspace \( \mathcal{X} \subseteq \mathbb{R}^n \) such that

\[
\lim_{t \to \infty} x(t, x_0) = 0 \text{ for any } x_0 \in \mathcal{X}.
\]

(b) Every solution of

\[
\dot{y}(t) = A[k](t)y(t),
\]
with \( y(\cdot) \in \mathbb{R}^{\binom{n}{k}} \), converges to the origin.

For \( k = 1 \), the equivalence of conditions (a) and (b) is obvious. For \( k = n \), condition (a) becomes the existence of a one-dimensional subspace \( \mathcal{X} \subseteq \mathbb{R}^n \) such that any solution emanating from \( \mathcal{X} \) converges to the origin, whereas (b) requires that every solution of the scalar system \( \dot{y}(t) = \text{trace}(A(t))y(t) \) converges to the origin.

Note that if \( \mu(A[k](t)) \leq -\eta < 0 \) for all \( t \geq 0 \), then condition (b) holds, and thus, condition (a) holds.

**Example 9** Consider the special case where the LTV (29) reduces to an LTI \( \dot{x}(t) = Ax(t) \). The uniform stability requirement implies that the real part of every eigenvalue of \( A \) is non-positive. Condition (b) is equivalent to the requirement that the sum of every \( k \) eigenvalues of \( A \) has a negative real part. This implies that at least \( (n - k + 1) \) eigenvalues of \( A \) have a negative real part, and thus, condition (a) holds.

**Example 10** Consider again the LTV in Example 8. Here, \( n = 2 \) and the LTV is uniformly stable and \( k \)-contracting for \( k = 2 \), so Proposition 10 implies that the LTV admits a one-dimensional linear subspace \( \mathcal{X} \subseteq \mathbb{R}^2 \) such that (41) holds. Indeed, it follows from (40) that \( \text{span}(\begin{bmatrix} 1 & 1 \end{bmatrix}^T) \) is this subspace.

As noted above, the definition of \( k \)-contractivity for nonlinear systems follows by applying the same approach to the variational equation which is an LTV. We refer to [82] for the details. Muldowney and his colleagues [42, 51] proved that time-invariant 2-contracting systems have a “well-ordered” asymptotic behavior, and this has been used to derive a global asymptotic analysis of important nonlinear dynamical models from epidemiology (see, e.g., [41]). A recent paper [53] extended some of these results to systems that are not necessarily 2-contracting, but can be represented as the serial interconnections of \( k \)-contracting systems, with \( k \in \{1, 2\} \). Reference [54] studied the series connection of two systems and derived a sufficient condition for \( k \)-contraction of the overall system. This is based on a new formula for the \( k \)-compounds of a block-diagonal matrix. The recent paper [52] derives sufficient conditions for a \( k \)-contractivity of a Lurie system.

The notion of \( k \)-contracting systems is based on the relation between \( k \)-compounds and the volume of \( k \)-dimensional parallelotopes. However, the \( k \)-minors of a matrix have another important application related to the sign variation diminishing property. Using this allows to introduce \( k \)-generalizations of the important classes of linear positive systems and nonlinear cooperative systems.

**8 \( k \)-positive systems**

Positive linear systems [22] and cooperative nonlinear systems [66] are characterized by the fact that every state-variable takes non-negative values. This is the case in many real-world systems where the state-variables represent quantities such as densities in queues, probabilities, concentration of molecules, etc. An important property of such systems is that many system and control problems scale well with the system size [60, 72].
Reference [77] introduced the notions of $k$-positive and $k$-cooperative systems. The LTV (29) is called $k$-positive if $A^{[k]}(t)$ is Metzler for all $t \in [\tau_0, \tau_1]$. In other words, we require the standard condition for positivity, but on the $k$-compound system. For $k = 1$, this reduces to requiring that $A(t)$ is Metzler for all $t \in [\tau_0, \tau_1]$. In this case, the system is positive, i.e., the flow maps $\mathbb{R}^n_+$ to $\mathbb{R}^n_+$ and also $\mathbb{R}^n_- := -\mathbb{R}^n_+$ to $\mathbb{R}^n_-$. In other words, the flow maps the set of vectors with zero sign variations to itself. The flow of $k$-positive systems maps the set of vectors with up to $(k - 1)$ sign variations to itself. To explain this, we use the definitions and results introduced in Sect. 3.3.

For any $k \in \{1, \ldots, n\}$, define the sets [24][40, p. 71]:

$$P^k_- := \{ z \in \mathbb{R}^n : s^-(z) \leq k - 1 \},$$
$$P^k_+ := \{ z \in \mathbb{R}^n : s^+(z) \leq k - 1 \}.$$  \(43\)

In other words, these are the sets of all vectors with up to $(k - 1)$ sign variations.

Then, $P^k_-$ is closed, and it can be shown that $P^k_+ = \text{int}(P^k_-)$. For example,

$$P^1_- = \mathbb{R}^n_+ \cup \mathbb{R}^n_-,$$  \(44\)
$$P^1_+ = \text{int}(\mathbb{R}^n_+) \cup \text{int}(\mathbb{R}^n_-).$$

Since multiplying a vector by a positive (or a negative) constant does not change the number of sign variations in the vector, the sets $P^k_-$ and $P^k_+$ are cones. However, they are not convex cones. For example, for $n = 2$, $x = [2 \ 1]^T$ and $y = [-1 \ -2]^T$, we have $s^-(x) = s^-(y) = s^+(x) = s^+(y) = 0$, but $s^-(\frac{x+y}{2}) = s^+(\frac{x+y}{2}) = 1$. For an analysis of the geometric structure of these sets, see [77].

**Definition 7** The LTV (29) is called $k$-positive on the interval $[\tau_0, \tau_1]$ if for any $\tau_0 < t_0 < \tau_1$,

$$x_0 \in P^k_- \implies x(t, t_0, x_0) \in P^k_+ \text{ for all } t_0 \leq t < \tau_1,$$

and is called strongly $k$-positive if

$$x_0 \in P^k_- \setminus \{0\} \implies x(t, t_0, x_0) \in P^k_+ \text{ for all } t_0 < t < \tau_1.$$

In other words, the sets of up to $(k - 1)$ sign variations are invariant sets of the dynamics.

Recall the important SVDP of TP matrices (19), namely, multiplying a vector by a TP matrix can never increase the number of sign variations. For our purposes, we need a more specialized result.

**Definition 8** A matrix $A \in \mathbb{R}^{n \times m}$ is called sign-regular of order $k$ (denoted $SR_k$) if its minors of order $k$ are all non-positive or all non-negative. It is called strictly sign-regular of order $k$ (denoted $SSR_k$) if its minors of order $k$ are all positive or all negative. In this case, we use $\epsilon_k$ to denote the signature of the $k$-minors, that is, $\epsilon_k = 1$ [$\epsilon_k = -1$] if all the $k$-minors are positive [negative].
For example, $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is SSR$_1$ with $\epsilon_1 = 1$ as every entry (i.e., every 1-minor) is positive, and it is SSR$_2$ with $\epsilon_2 = -1$ since $\det(A) < 0$, and this is the only 2-minor of $A$.

**Proposition 11** (see, e.g., [9]) Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix. Fix $k \in \{1, \ldots, n\}$. Then, the following two conditions are equivalent:

1. For any $x \in \mathbb{R}^n$ with $s^-(x) \leq k - 1$, we have $s^-(Ax) \leq k - 1$.
2. $A$ is SR$_k$.

Also, the following two conditions are equivalent:

(a) For any $x \in \mathbb{R}^n \setminus \{0\}$ with $s^-(x) \leq k - 1$, we have $s^+(Ax) \leq k - 1$.
(b) $A$ is SSR$_k$.

**Example 11** Consider the non-singular matrix $A = \begin{bmatrix} 3 & 2 & -1 \\ 3 & 5 & -1 \\ 3 & 5 & 0 \end{bmatrix}$. Then, $A^{(2)} = \begin{bmatrix} 9 & 0 & 3 \\ 9 & 3 & 5 \\ 0 & 3 & 5 \end{bmatrix}$, so $A$ is SR$_2$. Fix

$$x \in \mathbb{R}^3 \setminus \{0\} \text{ with } s^-(x) \leq 1. \quad (44)$$

Proposition 11 implies that $s^-(Ax) \leq 1$. We verify this directly. Seeking a contradiction, assume that $s^-(Ax) > 1$, i.e., $s^-(Ax) = 2$. Then, we may assume, wlog, that $y := Ax$ satisfies $y_1, y_3 > 0$ and $y_2 < 0$, that is,

$$3x_1 + 2x_2 - x_3 > 0,$$
$$3x_1 + 5x_2 - x_3 < 0,$$
$$3x_1 + 5x_2 > 0.$$ 

The first two equations give $x_2 < 0$, and the last two equations give $x_3 > 0$. Now, the first equation implies that $x_1 > 0$, and this contradicts (44).

Note that the statements in Proposition 11 do not directly compare $s^-(Ax)$ and $s^-(x)$. However, the results in this proposition imply the following.

**Corollary 1** Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix. Fix $k \in \{1, \ldots, n\}$. Then, the following two conditions are equivalent:

1. For any $x \in \mathbb{R}^n$ with $s^-(x) \leq k - 1$, we have $s^-(Ax) \leq s^-(x)$.
2. $A$ is SR$_j$ for every $j \in \{1, \ldots, k\}$.

Also, the following two conditions are equivalent:

(a) For any $x \in \mathbb{R}^n \setminus \{0\}$ with $s^-(x) \leq k - 1$, we have $s^+(Ax) \leq s^-(x)$.
(b) $A$ is SSR$_j$ for every $j \in \{1, \ldots, k\}$.
Suppose that $A$ is $SR_j$ for every $j \in \{1, \ldots, k\}$. Fix $x \in \mathbb{R}^n$ with $s^-(x) \leq k - 1$. Then, there exists a $j \in \{1, \ldots, k\}$ such that $s^-(x) = j - 1$, and Proposition 11 implies that $s^-(Ax) \leq j - 1$, so $s^-(Ax) \leq s^-(x)$.

Now, assume that there exists some $j \in \{1, \ldots, k\}$ such that $A$ is not sign-regular of order $j$. Then, Proposition 11 implies that there exists $x \in \mathbb{R}^n$ such that $s^-(x) \leq j - 1$ and $s^-(Ax) > j - 1$, so $s^-(Ax) > s^-(x)$. This proves the equivalence of the first two assertions. The remainder of the proof is similar and thus, omitted. 

Remark 1 Note that Corollary 1 implies the following. Let $A \in \mathbb{R}^{n \times n}$ be non-singular. If $A$ is $TN_k$, then $s^-(x) \leq k - 1$ implies that $s^-(Ax) \leq s^-(x)$. If $A$ is $TP_k$, then for any $x \neq 0$ with $s^-(x) \leq k - 1$ we have $s^+(Ax) \leq s^-(x)$.

Using these tools allows to characterize $k$-positive LTVs.

Theorem 3 [77] The LTV (29) is $k$-positive on $[\tau_0, \tau_1]$ iff $A^{[k]}(t)$ is Metzler for all $t \in (\tau_0, \tau_1)$. It is strongly $k$-positive on $[\tau_0, \tau_1]$ iff $A^{[k]}(t)$ is Metzler for all $t \in (\tau_0, \tau_1)$, and $A^{[k]}(t)$ is irreducible for all $t \in (\tau_0, \tau_1)$ except, perhaps, at isolated time points.

The proof is simple. Consider for example the second assertion in the theorem. The Metzler and irreducibility assumptions imply that the matrix differential system (31) is a positive linear system, and furthermore, all the entries of $\Phi^{(k)}(t, t_0)$ are positive for all $t > t_0$ (recall that the initial condition is $\Phi(t_0) = I \geq 0$). In other words, $\Phi(t, t_0)$ is $SSR_k$ for all $t > t_0$. Since $x(t, \tau_0, x(\tau_0)) = \Phi(t, t_0)x(\tau_0)$, applying Proposition 11 completes the proof.

This line of reasoning demonstrates the general principle in Sect. 6, namely, given conditions on $A^{[k]}$ we can apply standard tools from dynamical systems theory to the $k$-compound system (31) and deduce results on the behavior of the solution $x(t)$ of the original system (29).

8.1 Sign conditions for $k$-positivity

A natural question is: when is $A^{[k]}$ a Metzler matrix? This can be answered using Proposition 7 in terms of sign pattern conditions on the entries $a_{ij}$ of $A$. This is useful, as in fields like chemistry and systems biology, exact values of various parameters are typically unknown, but their signs may be inferred from various properties of the system [68].

Proposition 12 [77] Let $A \in \mathbb{R}^{n \times n}$ with $n \geq 3$. Then,

1. $A^{[n-1]}$ is Metzler iff $a_{ij} \geq 0$ for all $i, j$ with $i - j$ odd, and $a_{ij} \leq 0$ for all $i \neq j$ with $i - j$ even;
2. for any odd $k$ in the range $1 < k < n - 1$, $A^{[k]}$ is Metzler iff $a_{1n}, a_{n1} \geq 0$, $a_{ij} \geq 0$ for all $|i - j| = 1$, and $a_{ij} = 0$ for all $1 < |i - j| < n - 1$;
3. for any even $k$ in the range $1 < k < n - 1$, $A^{[k]}$ is Metzler iff $a_{1n}, a_{n1} \leq 0$, $a_{ij} \geq 0$ for all $|i - j| = 1$, and $a_{ij} = 0$ for all $1 < |i - j| < n - 1$. 

Springer
**Example 12** For \( n = 5 \), the corresponding sign patterns are as follows. The matrix \( A^{[4]} \) is Metzler iff

\[
A = \begin{bmatrix}
  * & + & - & + & - \\
  + & * & + & - & - \\
  - & + & * & + & - \\
  + & - & * & + & - \\
  - & + & - & + & * \\
\end{bmatrix},
\]

where + denotes a non-negative entry, − denotes a non-positive entry, and * denotes “don’t care”. The matrix \( A^{[3]} \) is Metzler iff

\[
A = \begin{bmatrix}
  * & + & 0 & 0 & + \\
  + & * & + & 0 & 0 \\
  0 & + & * & + & 0 \\
  0 & 0 & + & * & + \\
  + & 0 & 0 & + & * \\
\end{bmatrix},
\]

and \( A^{[2]} \) is Metzler iff

\[
A = \begin{bmatrix}
  * & + & 0 & 0 & - \\
  + & * & + & 0 & 0 \\
  0 & + & * & + & 0 \\
  0 & 0 & + & * & + \\
  - & 0 & 0 & 0 & + \\
\end{bmatrix}.
\]

We consider the three cases in Proposition 12 in more detail. In Case 1, \( A^{[n-1]} \) is Metzler. Let \( U \in \{-1, 1\}^{n \times n} \) be the antidiagonal matrix with entries

\[
u_{ij} = \begin{cases} 
(−1)^{j+1}, & \text{if } i + j = n + 1, \\
0, & \text{otherwise.}
\end{cases}
\]

Then, \( U^T = U^{-1} \), and it can be shown that \( A^{[n-1]} \) Metzler implies that \((-UAU^{-1})\) is Metzler [77]. For example, when \( n = 5 \) a calculation gives

\[
UAU^{-1} = \begin{bmatrix}
  a_{55} & -a_{54} & a_{53} & -a_{52} & a_{51} \\
  -a_{45} & a_{44} & -a_{43} & a_{42} & -a_{41} \\
  a_{35} & -a_{34} & a_{33} & -a_{32} & a_{31} \\
  -a_{25} & a_{24} & -a_{23} & a_{22} & -a_{21} \\
  a_{15} & -a_{14} & a_{13} & -a_{12} & a_{11} \\
\end{bmatrix},
\]

and (45) implies that \((-UAU^{-1})\) is Metzler. This is in fact a special case of a duality relation between additive compounds of a matrix, see [17]. In other words, in Case 1 the coordinate transformation \( y := UX \) gives \( \dot{y} = UAU^{-1}y \), and this is a competitive dynamical system [66]. Thus, \( k \)-positive systems, with \( k \in \{1, \ldots, n-1\} \),
may be viewed as a kind of interpolation from cooperative systems (when \( k = 1 \)) to competitive systems (when \( k = n - 1 \)).

In Case 2), \( A \) is in particular Metzler. The dynamical behavior in Case 3) is illustrated in the next example.

**Example 13** Consider the case \( n = 3 \) and \( A = \begin{bmatrix} 2 & 1 & -0.5 \\ 0 & -1 & 0.5 \\ -1 & 0 & 5 \end{bmatrix} \). Note that \( A \) is not Metzler, yet \( A^{[2]} = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0 & 7 & 1 \\ 1 & 0 & 4 \end{bmatrix} \) is Metzler (and irreducible). Theorem 3 guarantees that for any \( x_0 \) with \( s^-(x_0) \leq 1 \), we have

\[
\begin{align*}
  s^-(x(t, x_0)) &\leq 1 \\
  \text{for all } t \geq 0.
\end{align*}
\]  

(46)

Fig. 6 depicts \( s^-(x(t, x_0)) = s^-(\exp(At)x_0) \) for \( x_0 = \begin{bmatrix} 2 & -30 & -6 \end{bmatrix}^T \). Note that \( s^-(x_0) = 1 \). It may be seen that \( s^-(x(t, x_0)) \) decreases and then, increases, but always satisfies the bound (46).

Fix \( k \in \{2, \ldots, n - 1\} \). Proposition 12 implies that a system is \( k \)-positive with \( k \) even iff it is 2-positive, and if it is \( k \)-positive with \( k \) odd, then it is 1-positive.

### 8.2 Totally positive differential systems

We call \( A \in \mathbb{R}^{n \times n} \) a Jacobi matrix if \( A \) is tri-diagonal with positive entries on the super- and sub-diagonals. An immediate implication of Proposition 12 is that \( A^{[k]} \) is Metzler and irreducible for all \( k \in \{1, \ldots, n - 1\} \) iff \( A \) is Jacobi. It then follows that for any \( t > 0 \) the matrices \( (\exp(At))^{(k)}, k = 1, \ldots, n, \) are (component-wise) positive, that is, \( \exp(At) \) is TP for all \( t > 0 \). Combining this with Theorem 3 yields the following.

\( \square \) Springer
Proposition 13 [62] The following two conditions are equivalent.
1. A is Jacobi.
2. for any \( x_0 \in \mathbb{R}^n \setminus \{0\} \) the solution of the LTI \( \dot{x}(t) = Ax(t), x(0) = x_0 \), satisfies
\[
s^-(x(t, x_0)) \leq s^+(x(t, x_0)) \leq s^-(x_0) \text{ for all } t > 0.
\]

The left-hand side inequality follows from (17), and the right-hand inequality from the SVDP of TP systems in (19). Proposition 13 implies that \( s^-(x(t, x_0)) \) and also \( s^+(x(t, x_0)) \) are non-increasing functions of \( t \) and may thus be considered as piece-wise constant Lyapunov functions for the dynamics.

Proposition 13 was proved by Schwarz [62], who only considered linear systems. It was recently shown [48] that important results on the asymptotic behavior of time-invariant and periodic time-varying nonlinear systems with a Jacobian that is a Jacobi matrix for all \( t \), \( x \) [63, 65] follow from the fact that the associated variational equation is a totally positive LTV.

9 \( k \)-cooperative systems

We now review the applications of \( k \)-positivity to the time-invariant nonlinear system:
\[
\dot{x} = f(x),
\]
(47)
with \( f \in C^1 \). Let \( J(x) := \frac{\partial}{\partial x} f(x) \). We assume that the trajectories of (47) evolve on a convex and compact state-space \( \Omega \subseteq \mathbb{R}^n \).

Recall that (47) is called cooperative if \( J(x) \) is Metzler for all \( x \in \Omega \) [66]. In other words, the variational equation associated with (47) is positive. The system (47) is called strongly cooperative if \( J(x) \) is Metzler and irreducible for all \( x \in \Omega \). Strong cooperativity has far reaching implications. By Hirsch’s quasi-convergence theorem [66], almost every bounded trajectory converges to the set of equilibria.

It is natural to generalize cooperativity to \( k \)-cooperativity by requiring that the variational equation associated with (47) is \( k \)-positive.

Definition 9 [77] The nonlinear system (47) is called \( \text{[strongly]} \ k \)-cooperative if the associated LTV (36) is \( \text{[strongly]} \ k \)-positive for any \( a, b \in \Omega \).

Note that for \( k = 1 \) this reduces to the definition of a cooperative [strongly cooperative] dynamical system.

One immediate implication of Definition 9 is the existence of certain invariant sets of the dynamics.

Proposition 14 Suppose that (47) is \( k \)-cooperative. Pick \( a, b \in \Omega \). Then,
\[
a - b \in P^k_- \implies x(t, a) - x(t, b) \in P^k_- \text{ for all } t \geq 0.
\]
If, furthermore, \( 0 \in \Omega \) and \( 0 \) is an equilibrium point of (47), i.e., \( f(0) = 0 \), then
\[
a \in P^k_- \implies x(t, a) \in P^k_- \text{ for all } t \geq 0.
\]
The sign pattern conditions in Proposition 12 yields simple sufficient conditions for [strong] $k$-cooperativity of (47). Indeed, if $J(x)$ satisfies a sign pattern condition for all $x \in \Omega$, then the integral of $J$ in the variational equation (36) satisfies the same sign pattern, and thus, so does $A^{ab}$. The next example, adapted from [77], illustrates this.

**Example 14** Elkhader [19] studied the nonlinear system

\[
\begin{align*}
\dot{x}_1 &= f_1(x_1, x_n), \\
\dot{x}_i &= f_i(x_{i-1}, x_i, x_{i+1}), & i &= 2, \ldots, n-1, \\
\dot{x}_n &= f_n(x_{n-1}, x_n), \\
\end{align*}
\]

under the following assumptions: the state-space $\Omega \subseteq \mathbb{R}^n$ is convex, $f_i \in C^{n-1}$, $i = 1, \ldots, n$, and there exist $\delta_i \in \{-1, 1\}$, $i = 1, \ldots, n$, such that

\[
\begin{align*}
\delta_1 \frac{\partial}{\partial x_n} f_1(x) &> 0, \\
\delta_2 \frac{\partial}{\partial x_1} f_2(x), \delta_3 \frac{\partial}{\partial x_3} f_2(x) &> 0, \\
&\vdots \\
\delta_{n-1} \frac{\partial}{\partial x_{n-2}} f_{n-1}(x), \delta_n \frac{\partial}{\partial x_{n-1}} f_{n-1}(x) &> 0, \\
\delta_n \frac{\partial}{\partial x_{n-1}} f_n(x) &> 0,
\end{align*}
\]

for all $x \in \Omega$. This is a generalization of the monotone cyclic feedback system analyzed in [45]. As noted in [19], we may assume without loss of generality that $\delta_2 = \delta_3 = \cdots = \delta_n = 1$ and $\delta_1 \in \{-1, 1\}$. Then, the Jacobian of (48) has the form

\[
J(x) = \begin{bmatrix}
* & 0 & 0 & 0 & \cdots & 0 & 0 & \text{sgn}(\delta_1) \\
> 0 & * & > 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & > 0 & * & > 0 & \cdots & 0 & 0 & 0 \\
& \vdots & & & & & & \\
0 & 0 & 0 & 0 & \cdots & 0 & > 0 & *
\end{bmatrix},
\]

for all $x \in \Omega$. Here, $*$ denotes “don’t care”. Note that $J(x)$ is irreducible for all $x \in \Omega$.

If $\delta_1 = 1$, then $J(x)$ is Metzler and irreducible, so the system is strongly 1-cooperative. If $\delta_1 = -1$, then $J(x)$ satisfies the sign pattern in Case 3 in Proposition 12, so the system is strongly 2-cooperative. (If $n$ is odd, then $J(x)$ also satisfies the sign pattern in Case 1, so there is a coordinate transformation for which the system is also strongly competitive.)

The main result in [77] is that strongly 2-cooperative systems satisfy a strong Poincaré–Bendixson property.

Springer
Theorem 4 (Poincaré–Bendixson Property) [77] Suppose that (47) is strongly 2-cooperative. Pick \( a \in \Omega \). If the omega limit set \( \omega(a) \) does not include an equilibrium, then it is a closed orbit.

The proof of this result is based on the seminal results of Sanchez [61] on dynamical systems that admit an invariant cone of rank \( k \). Yet, Theorem 4 is considerably stronger than the main result in [61], as it applies to any trajectory emanating from \( \Omega \) and not only to pseudo-ordered trajectories (see the definition in [61]).

The Poincaré–Bendixson property is useful because often it can be combined with a local analysis near the equilibrium points to provide a global picture of the dynamics. For applications of a Poincaré–Bendixson property to models from systems biology, synthetic biology, and epidemiology, see [47, 75, 82].

Summarizing, the use of \( k \)-compounds allows to generalize positive linear systems [cooperative nonlinear systems] to \( k \)-positive linear systems [\( k \)-cooperative nonlinear systems]. In particular, 2-cooperative systems, of any order \( n \), “behave like” 2-dimensional systems.

The next section describes the use of compound matrices to extend the important concept of diagonal stability to \( k \)-diagonal stability [83]. We consider the case of discrete-time (DT) systems. For a symmetric matrix \( S \in \mathbb{R}^{n \times n} \), we use \( S \succ 0 \) [\( S \prec 0 \)] to denote that \( S \) is positive-definite [negative-definite].

10 \( k \)-diagonal stability

Recall that the discrete-time LTI system

\[
x(j+1) = Ax(j),
\]

is called \textit{diagonally stable} if there exists a diagonal matrix \( D \succ 0 \) such that \( A^TDA \prec D \). In other words, \( V(x) := x^T Dx \) is a diagonal Lyapunov function for (49).

Diagonal stability of both discrete-time and continuous-time LTI systems has attracted considerable interest, as it has important implications to nonlinear dynamical systems [5, 7, 37, 80]. For example, let \( S^1 \) denote the class of functions \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) such that \( f(x) = [f_1(x_1) \ldots f_n(x_n)]^T \), each \( f_i : \mathbb{R} \rightarrow \mathbb{R} \) is continuous and satisfies

\[
0 < |f_i(s)| \leq |s|, \quad \text{for all } s \in \mathbb{R} \setminus \{0\}.
\]

Note that \( f(0) = 0 \). Consider the DT nonlinear system

\[
x(j+1) = Af(x(j)), \quad f \in S^1.
\]

If (49) is diagonally stable, then it is straightforward to verify that \( V(x) := x^T Dx \) is also a Lyapunov function for the nonlinear system (50).

Stable LTIs always admit a quadratic Lyapunov function, but not necessarily a diagonal Lyapunov function (DLF). However, stable \textit{positive} LTIs always admit a DLF.
Lemma 2 [11, 59] If \( A \in \mathbb{R}^{n \times n} \) with \( A \geq 0 \), then the following statements are equivalent:

1. The matrix \( A \) is Schur, i.e., \( \rho(A) < 1 \).
2. There exists \( \xi \in \mathbb{R}^n \) with \( \xi \gg 0 \) such that \( A\xi \ll \xi \).
3. There exists \( z \in \mathbb{R}^n \) with \( z \gg 0 \) such that \( A^Tz \ll z \).
4. There exists a diagonal matrix \( D > 0 \) such that \( A^TDA < D \).
5. The matrix \((I - A)\) is non-singular and \((I - A)^{-1} \geq 0\).

Remark 2 Suppose that \( A \geq 0 \) is Schur. Fix \( x, y \in \mathbb{R}^n \), with \( x, y \gg 0 \). Then,

\[
\xi := (I - A)^{-1}x, \\
z := (I - A^T)^{-1}y, \\
D := \text{diag}\left(\frac{z_1}{\xi_1}, \ldots, \frac{z_n}{\xi_n}\right)
\]

satisfy conditions (2), (3), and (4) in Lemma 2, respectively. This provides a constructive recipe for determining a DLF \( D \). Note that if \( A \in \mathbb{R}^{n \times n} \) is Schur and \( A \leq 0 \), then \((-A)\) is a non-negative Schur matrix, and Lemma 2 guarantees that there exists a diagonal matrix \( D > 0 \) such that \( A^TDA < D \).

10.1 Discrete-time \( k \)-diagonal stability

Fix \( k \in \{1, \ldots, n\} \). To study the evolution of volumes of \( k \)-parallelotopes under the DT LTI (49), fix \( k \) initial conditions \( a^1, \ldots, a^k \in \mathbb{R}^n \). Then,

\[
\begin{align*}
[x(j, a^1) \ldots x(j, a^k)]^{(k)} &= [A^ja^1 \ldots A^ja^k]^{(k)} \\
&= (A^j)^{(k)} [a^1 \ldots a^k]^{(k)} \\
&= (A^{(k)})^j [a^1 \ldots a^k]^{(k)}.
\end{align*}
\]

This may be interpreted as the \( k \)-compound system of (49). It shows that the evolution of \( k \)-parallelotopes follows a DT LTI with the matrix \( A^{(k)} \). This naturally leads to the following definition.

Definition 10 [83] Consider the DT LTI (49) with \( A \in \mathbb{R}^{n \times n} \). Fix \( k \in \{1, \ldots, n\} \) and let \( r := \binom{n}{k} \). Then, (49) is called \( k \)-diagonally stable if there exists a matrix \( D = \text{diag}(d_1, \ldots, d_r) > 0 \) such that

\[
(A^{(k)})^TDA^{(k)} < D.
\]

Note that for \( k = 1 \) this reduces to standard diagonal stability, and for \( k = n \) this becomes the requirement \((\det(A))^2 < 1\). Note also that if \( A^{(k)} \) is Schur and \( A^{(k)} \geq 0 \), then Lemma 2 implies that (49) is \( k \)-diagonally stable.

It is natural to expect that diagonal stability implies \( k \) diagonal stability for any \( k \). The next result shows that this is indeed so.
Proposition 15 [83] If $A$ is diagonally stable, then it is $k$-diagonally stable for any $k \in \{1, 2, \ldots, n\}$.

Proof Let $\mathbb{D}^{n \times n}$ denote the set of positive diagonal $n \times n$ matrices. There exists a $P \in \mathbb{D}^{n \times n}$ such that $A^TPA < P$. Multiplying this inequality by $P^{-1/2}$ on the right- and left-hand side yields:

$$P^{-1/2}A^TPAP^{-1/2} < I_n,$$

so the symmetric matrix $H := P^{-1/2}A^TPAP^{-1/2}$ is Schur. Fix $k \in \{1, 2, \ldots, n\}$, and let $D := P^{(k)}$. Then, $D \in \mathbb{D}^{r \times r}$, where $r := \binom{n}{k}$. Now,

$$H^{(k)} = D^{-1/2}(A^{(k)})^TDA^{(k)}D^{-1/2},$$

and since $H^{(k)}$ is symmetric and any eigenvalue of $H^{(k)}$ is the product of $k$ eigenvalues of $H$, we also have $H^{(k)} < I_r$. We conclude that

$$(A^{(k)})^TDA^{(k)} < D,$$

i.e., $A$ is $k$-diagonally stable. \qed

We note that, in general, $k$-diagonal stability, with $k > 1$, does not imply diagonal stability (see the specific examples in [83]).

It is interesting to find general classes of matrices that are $k$-diagonally stable for some $k$. Recall that $A \in \mathbb{R}^{n \times n}$ is called cyclic if

$$A = \begin{bmatrix}
\alpha_1 & \beta_1 & 0 & \ldots & 0 & 0 \\
0 & \alpha_2 & \beta_2 & \ldots & 0 & 0 \\
0 & 0 & \alpha_3 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \alpha_{n-1} & \beta_{n-1} \\
(-1)^{\ell+1}\beta_n & 0 & 0 & \ldots & 0 & \alpha_{n}\n\end{bmatrix},$$

with $\alpha_i, \beta_i \geq 0$, $i = 1, \ldots, n$, and $\ell \in \mathbb{Z}$. The DT LTI system $x(j + 1) = Ax(j)$ is called cyclic if $A$ is cyclic. This represents a dynamical system where each $x_i$, $i = 1, \ldots, n-1$, receives positive feedback from its “cyclic neighbours” $x_i, x_{i+1}$, with the exception that $x_n$ receives negative feedback from $x_1$ if $\ell$ is even.

Theorem 5 [83] Suppose that $A \in \mathbb{R}^{n \times n}$ is cyclic for some $\ell \in \{1, \ldots, n-1\}$. Then, $A$ is $SR_\ell$ with signature $\epsilon_\ell = 1$. Furthermore, if $\ell$ is odd, then $A$ is diagonally stable iff $A$ is Schur. If $\ell$ is even, then $A$ is $\ell$-diagonally stable iff $A^{(\ell)}$ is Schur.

Example 15 Consider the case $n = 3$, i.e.,

$$A = \begin{bmatrix}
\alpha_1 & \beta_1 & 0 \\
0 & \alpha_2 & \beta_2 \\
(-1)^{\ell+1}\beta_3 & 0 & \alpha_3
\end{bmatrix}. $$
Then,
\[
A^{(2)} = \begin{bmatrix}
\alpha_1 \alpha_2 & \alpha_1 \beta_2 & \beta_1 \beta_2 \\
(-1)^{\ell} \beta_1 \beta_3 & \alpha_1 \alpha_3 & \beta_3 \alpha_1 \\
(-1)^{\ell} \alpha_2 \beta_3 & (-1)^{\ell} \beta_2 \beta_3 & \beta_3 \alpha_2
\end{bmatrix}.
\]

Thus, for \(\ell = 1\) [\(\ell = 2\)], \(A^{SR_1} [SR_2]\) with signature \(\epsilon_1 = 1\) [\(\epsilon_2 = 1\)] since all the entries of \(A [A^{(2)}]\) are non-negative.

As an application of these notions, Ref. [83] describes a class of nonlinear functions \(S^k\) such that \(k\)-diagonal stability of (49) implies that the evolution of \(k\)-parallelotopes under the nonlinear dynamics
\[
x(j + 1) = Af(x(j)), \quad f \in S^k.
\]
is asymptotically stable. For \(k = 1\), \(S^k\) is the set of functions \(S^1\) defined above.

The \(k\)-compound matrices are defined only for integer values of \(k\), as they are based on \(k \times k\) sub-matrices. However, it turns out that there are good reasons to generalize the notion of \(k\)-compounds, with \(k\) an integer, to \(\alpha\)-compounds, with \(\alpha\) a positive real number. In particular, this allows to introduce the notion of \(\alpha\)-contracting systems.

### 11 \(\alpha\)-compounds and \(\alpha\)-contracting systems

A recent paper [84] defined a generalization called the \(\alpha\)-multiplicative compound and \(\alpha\)-additive compound of a matrix, where \(\alpha\) is a real number.

Recall that the Kronecker product of two matrices \(A \in \mathbb{C}^{n \times m}\) and \(B \in \mathbb{C}^{p \times q}\) is
\[
A \otimes B := \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1m}B \\
a_{21}B & a_{22}B & \cdots & a_{2m}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1}B & a_{n2}B & \cdots & a_{nm}B
\end{bmatrix},
\]
where \(a_{ij}\) denotes the \(ij\)th entry of \(A\). Hence, \(A \otimes B \in \mathbb{C}^{(np) \times (mq)}\). The Kronecker sum of two square matrices \(X \in \mathbb{C}^{n \times n}\) and \(Y \in \mathbb{C}^{m \times m}\) is
\[
X \oplus Y := X \otimes I_m + I_n \otimes Y.
\]

**Definition 11** Let \(A \in \mathbb{C}^{n \times n}\) be non-singular. If \(\alpha = k + s\), where \(k \in \{1, \ldots, n-1\}\) and \(s \in (0, 1)\), then the \(\alpha\)-multiplicative compound of \(A\) is defined by:
\[
A^{(\alpha)} := (A^{(k)})^{1-s} \otimes (A^{(k+1)})^{s}.
\]

Note that \(A^{(\alpha)} \in \mathbb{C}^{r \times r}\), where \(r := \binom{n}{k} \binom{n}{k+1}\), and that \(A^{(\alpha)}\) may be complex (non-real) even if \(A\) is real. Since \(A\) is non-singular, \(A^{(\ell)}\) is non-singular for all \(\ell \in \{1, \ldots, n\}\), so \((A^{(k)})^{1-s}\) and \((A^{(k+1)})^{s}\) in (56) are well-defined.
The matrix $A^{(\alpha)}$ is a kind of “multiplicative interpolation” between $A^{(k)}$ and $A^{(k+1)}$. For example, $A^{(2,2)} = (A^{(2)})^{0.8} \otimes (A^{(3)})^{0.2}$.

**Example 16** Let $D = \text{diag}(d_1, \ldots, d_4)$ with $d_i \neq 0$ for all $i$. Fix $\alpha \in (2, 3)$, so that $k = 2$ and $s = \alpha - 2 \in (0, 1)$. Then,

$$D^{(\alpha)} = (D^{(2)})^{1-s} \otimes (D^{(3)})^s$$

$$= \text{diag}((d_1d_2)^{1-s}, (d_1d_3)^{1-s}, \ldots, (d_3d_4)^{1-s})) \otimes \text{diag}((d_1d_2d_3)^s, (d_1d_2d_4)^s, (d_1d_3d_4)^s, (d_2d_3d_4)^s)$$

$$= \text{diag}(d_1d_2d_3^s, d_1d_2d_4^s, \ldots, d_2^sd_3d_4),$$

so, any eigenvalue of $D^{(\alpha)}$ is a “multiplicative interpolation” between eigenvalues of $D^{(2)}$ and $D^{(3)}$.

Just like the $k$-additive compound, the $\alpha$-additive compound is defined using the $\alpha$-multiplicative compound.

**Definition 12** Let $A \in \mathbb{C}^{n \times n}$ be non-singular. If $\alpha = k + s$, where $k \in \{1, \ldots, n-1\}$ and $s \in (0, 1)$ then the $\alpha$ additive compound matrix of $A$ is

$$A^{[\alpha]} := \frac{d}{d\epsilon} (I + \epsilon A)^{(\alpha)}|_{\epsilon=0}.$$ 

It was shown in [84] that this yields

$$A^{[\alpha]} = ((1-s)A^{[k]}) \oplus (sA^{[k+1]}). \quad (57)$$

**Example 17** Let $A \in \mathbb{R}^{n \times n}$ and fix $\alpha \in (n-1, n)$ so that $\alpha = k + s$, with $k = n-1$ and $s \in (0, 1)$. Then, (57) gives

$$A^{[\alpha]} = ((1-s)A^{[n-1]}) \oplus (sA^{[n]})$$

$$= ((1-s)A^{[n-1]}) \oplus (s \text{trace}(A))$$

$$= ((1-s)A^{[n-1]}) \otimes I_1 + I_n \otimes (s \text{trace}(A))$$

$$= (1-s)A^{[n-1]} + s \text{trace}(A)I_n, \quad (58)$$

so in this particular case the Kronecker sum simplifies to a standard matrix sum.

Consider the time-varying nonlinear dynamical system (35) where $f$ is $C^1$. Let $x(t, t_0, x_0)$ denote the solution of (35) at time $t$ with $x(t_0) = x_0$. We assume from here on that $t_0 = 0$, and let $x(t, x_0) := x(t, 0, x_0)$. We also assume that the system admits a convex invariant set $\Omega \subseteq \mathbb{R}^n$, that is, for any $x_0 \in \Omega$ we have $x(t, x_0) \in \Omega$ for all $t \geq 0$. Let $J(t, x) := \frac{d}{dx} f(t, x)$.

Recall that the $k$-additive compound was used to define the notion of $k$-contraction. The next definition follows the same line of reasoning.
Definition 13 [84] The system (35) is called *infinitesimally* $\alpha$-contracting w.r.t. the norm $|\cdot|$ if

$$\mu(J^{[\alpha]}(t, x)) \leq -\eta < 0,$$

for all $t \geq 0$ and all $x \in \Omega$ [84].

A set $K \subseteq \Omega$ is called a strongly invariant set of (35) if

$$K = x(t, K) \text{ for all } t \geq 0.$$  

In other words, if we take all the points in $K$ as initial conditions of the dynamics, then the union of the resulting solutions at time $t$ is $K$.

For example, an equilibrium or a limit cycle is strongly invariant sets. More generally, for any $a \in \Omega$, the omega limit set $\omega(a)$ is a strongly invariant set.

Using these notions, it is possible to restate a seminal result of Douady and Oesterlé [18] (see also [67]) in terms of $\alpha$ contraction.

Theorem 6 [84] Suppose that (35) is $\alpha$-contracting for some $\alpha \in [1, n]$. Then, any compact and strongly invariant set of the dynamics has a Hausdorff dimension smaller than $\alpha$.

Roughly speaking, the dynamics contracts sets with a larger Hausdorff dimension, so any strongly invariant set must have a Hausdorff dimension small than $\alpha$.

The next example, adapted from [84], shows how these notions can be used to design a feedback controller that “de-chaotifies” a nonlinear dynamical system.

Example 18 Thomas’ cyclically symmetric attractor [8, 73] is a popular example for a chaotic system. It is described by:

$$\begin{align*}
\dot{x}_1 &= \sin(x_2) - bx_1, \\
\dot{x}_2 &= \sin(x_3) - bx_2, \\
\dot{x}_3 &= \sin(x_1) - bx_3,
\end{align*}$$

where $b > 0$ is the dissipation constant. The convex and compact set $\Omega := \{x \in \mathbb{R}^3 : \max_i |x_i| \leq b^{-1}\}$ is an invariant set of the dynamics, and we consider only initial conditions in this set.

For $b > 1$, the origin is the single stable equilibrium of (61). As $b$ is decreased, the dynamics becomes more complicated. Figure 7 depicts the solution of the system emanating from $x(0) = [1 \ -2 \ 1]^T$ for $b = 0.1$. Note the symmetric strange attractor.

The Jacobian $J_f$ of the vector field in (61) is

$$J_f(x) = \begin{bmatrix} -b & \cos(x_2) & 0 \\
0 & -b & \cos(x_3) \\
\cos(x_1) & 0 & -b \end{bmatrix}.$$
and thus, $J_f^{[3]} = \text{trace}(J(x)) = -3b$. Since $b > 0$, this implies that the system is 3-contracting w.r.t. any norm. Let $\alpha = 2 + s$, with $s \in (0, 1)$. Then, combining (58) and (28) gives

$$J_f^{[\alpha]}(x) = (1 - s)J_f^{[2]}(x) \oplus sJ_f^{[3]}(x)$$

$$= \begin{bmatrix}
-(2 + s)b & (1 - s)\cos(x_3) & 0 \\
0 & -(2 + s)b & (1 - s)\cos(x_2) \\
-(1 - s)\cos(x_1) & 0 & -(2 + s)b
\end{bmatrix}.$$ 

Thus,

$$\mu_1(J_f^{[\alpha]}(x)) \leq 1 - 2b - s(b + 1), \text{ for all } x \in \Omega.$$ 

We conclude that for any $b \in (0, 1/2)$ the system is $(2 + s)$-contracting w.r.t. the $L_1$ norm for any $s > \frac{1 - 2b}{1 + b}$.

We now show how $\alpha$-contraction can be used to design a partial-state controller for the system guaranteeing that the closed-loop system has a “well-ordered” behavior. Suppose that the closed-loop system is:

$$\dot{x} = f(x) + g(x),$$

where $f$ is the vector field in (61), and $g$ is the controller. Let $\alpha = 2 + s$, with $s \in (0, 1)$. The Jacobian of the closed-loop system is $J_{cl} := J_f + J_g$, so

$$\mu_1(J_{cl}^{[\alpha]}) = \mu_1(J_f^{[\alpha]} + J_g^{[\alpha]})$$

$$\leq \mu_1(J_f^{[\alpha]}) + \mu_1(J_g^{[\alpha]})$$

$$\leq 1 - 2b - s(b + 1) + \mu_1(J_g^{[\alpha]}),$$
Fig. 8 Several trajectories of the closed-loop system. The circles denote the initial conditions of the trajectories.

where the first inequality follows from the fact that any matrix measure is sub-additive. This implies that the closed-loop system is $\alpha$-contracting if

$$\mu_1(J^{[\alpha]}_g(x)) < s(b + 1) + 2b - 1 \text{ for all } x \in \Omega.$$ \hspace{1cm} (63)

Consider, for example, the controller $g(x_1, x_2) := c \text{ diag}(1, 1, 0)x$, with gain $c < 0$. Then, $J^{[\alpha]}_g = c \text{ diag}(2, 1 + s, 1 + s)$ and for any $c < 0$ condition (63) becomes

$$(1 + s)c < s(b + 1) + 2b - 1.$$ \hspace{1cm} (64)

This provides a simple recipe for determining the gain $c$ so that the closed-loop system is $(2 + s)$-contracting. For example, when $s \to 0$, Eq. (64) yields $c < 2b - 1$, and this guarantees that the closed-loop system is 2-contracting. Recall that in a time-invariant 2-contracting system every bounded trajectory converges to the set of equilibria, thus ruling out chaotic attractors and even non-trivial limit cycles [42]. Figure 8 depicts the behavior of the closed-loop system (62) with $b = 0.15$ and $c = 2b - 1.15$. The closed-loop system is thus 2-contracting, and as expected every solution converges to an equilibrium.

Summarizing, the notion of $\alpha$-contraction allows to add to the chaotic system the “correct amount” of feedback contraction, that may be a non-integer, to obtain a 2-contracting system.

The notion of $\alpha$-contraction has another interesting implication. It was shown in [82] that if

$$\mu(J^{[\alpha]}) \leq -\eta < 0,$$

with $\mu$ a matrix measure induced by an $L_p$ norm, with $p \in \{1, 2, \infty\}$, then

$$\mu(J^{[\alpha+\varepsilon]}) \leq -\eta < 0, \text{ for any } \varepsilon \geq 0.$$ \hspace{1cm} This monotonicity property implies the following. If the system is $(n - 1)$-contracting, then there always exists a real value $\alpha^*$ such that the system is $\alpha$ contracting for
any $\alpha > \alpha^*$. Thus, contraction is not a binary, yes-no property, that is, a system is either contracting or not. Rather, contraction holds for any $\alpha > \alpha^*$ (see Fig. 9).

The next example demonstrates this in the context of the well-known linear consensus algorithm [49].

**Example 19** Consider the LTI system

$$\dot{x} = -Lx,$$  

(65)

where $L \in \mathbb{R}^{n \times n}$ is the Laplacian of a (directed or undirected) weighted graph with a globally reachable vertex. For any $c \in \mathbb{R}$, we have that $c1_n$ is an equilibrium of (65), so the system cannot be 1-contracting w.r.t. any norm.

We now show that for any $s > 0$ there exists a vector norm $\| \cdot \|$ such that (65) is $(1+s)$-contracting w.r.t. to $\| \cdot \|$ (in other words, although the system is not 1-contracting, it is a “arbitrarily close” to being 1-contracting). It is enough to show that this holds for any $s \in (0, 1)$. Fix such an $s$. Let $\lambda_i$ denote the eigenvalues of $L$ ordered such that

$$\text{Re}(\lambda_1) \leq \text{Re}(\lambda_2) \leq \cdots \leq \text{Re}(\lambda_n),$$  

(66)

Then, $\lambda_1 = 0$ and $\text{Re}(\lambda_2) > 0$. Equation (57) gives

$$L^{[1+s]} = ((1-s)L) \oplus (sL^{[2]}).$$

The eigenvalues of $L^{[2]}$ are the sums $\lambda_i + \lambda_j$, $i < j$, so in particular the eigenvalue of $L^{[2]}$ with the minimal real part is $\lambda_2$. Recall that the eigenvalues of $A \oplus B$ are the pairwise sums of the spectrum of $A$ and $B$, so the minimal real part of an eigenvalue of $L^{[1+s]}$ is $s \text{Re}(\lambda_2) > 0$, i.e., $-L^{[1+s]}$ is Hurwitz. It is well-known [3, 43] that this implies that there exists a matrix measure $\mu$, induced by a scaled $L_2$ norm, such that $\mu(-L^{[1+s]}) < 0$. Thus, (65) is $(1+s)$-contracting.

Note that combining this with Theorem 6 implies that any compact and strongly invariant set of (65) has a Hausdorff dimension smaller or equal to one. Indeed, the one-dimensional “diagonal set” $\{x \in \mathbb{R}^n : x_1 = \cdots = x_n\}$ is an invariant set of the dynamics.

So far we only considered autonomous systems, i.e., systems with no inputs. The next section briefly reviews recent work by Grussler, Sepulchre and their colleagues.
on single-input single-output (SISO) DT LTI systems that are \( k \)-positive. This leads to the interesting notion of the \( k \)-compound of a transfer function. For the sake of consistency in this paper, we modify some of the terminology and notations. Recall that a square matrix \( A \) is called Hankel if every entry \( a_{ij} \) depends only on \( i + j \). In other words, the entries along any anti-diagonal are equal.

#### 12 Hankel \( k \)-positivity of discrete-time SISO LTI systems

Consider the SISO DT LTI system
\[
x(j + 1) = Ax(j) + bu(j), \\
y(j) = c^T x(j),
\]
with \( A \in \mathbb{R}^{n \times n} \), and \( b, c \in \mathbb{R}^n \). The impulse response of this system is:
\[
g(j) := c^T A^{j-1} b \mathbb{1}(j - 1),
\]
where \( \mathbb{1}(\cdot) \) is the Heaviside step function. We assume throughout that \( \sum_k |g(k)| < \infty \), and consider controls satisfying \( \sup_k |u(k)| < \infty \).

Let
\[
G(z) := \sum_{k=0}^{\infty} g(k)z^{-k}
\]
denote the transfer function of (67). We always assume that the state-space representation (67) is minimal, so there are no zero-pole cancellations in \( G(z) \).

Section 8 considered systems without inputs satisfying a SVDP w.r.t. to the mapping from a past state vector to a future state vector. For an input–output system like (67), it is natural to consider the relation between the number of sign variations in a sequence of consecutive inputs and the number of sign variations in the corresponding sequence of consecutive outputs. In certain applications, e.g., smoothing filters, it is natural to require an SVDP between such sequences. In general, establishing such an SVDP for (67) is a difficult problem because any sequence of outputs is the sum of two terms: the first [second] representing the effect of the initial condition [input sequence]. Bounding the number of sign changes in the sum of two vectors is non-trivial. For example, for \( a = [2 \ 2 \ 2]^T \) and \( b = [-1 \ -3 \ -1]^T \), we have \( s^-(a) = s^-(b) = 0 \), yet \( s^-(a + b) = 2 \).

Thus, it is customary to make some simplifying assumptions. From here on, we consider the so-called Hankel case defined by the assumptions that
\[
0 = u(0) = u(1) = u(2) = \ldots,
\]
and that the relevant sub-sequence of outputs is
\[
y(0), \ y(1), \ y(2), \ldots.
\]
Intuitively speaking, the control values $u(\ell)$, with $\ell < 0$, determine an initial condition $x(0)$, and (69) implies that for any $j \geq 0$ the system (67) reduces to the autonomous system

$$
x(j + 1) = Ax(j),
$$
$$
y(j) = c^T x(j).
$$

Under these assumptions, for any $j \geq 0$ we have

$$
y(j) = (Hgu)(j) := \sum_{\tau = 1}^{\infty} g(j + \tau)u(-\tau),
$$

where $H_g$ is called the \textit{Hankel operator} corresponding to $g$. The variation-diminishing properties of such convolution operators are a central theme in the theory of total positivity [36] and have found applications in various fields including statistics and interpolation theory.

In matrix notation,

$$
\begin{bmatrix}
y(0) \\
y(1) \\
y(2) \\
\vdots
\end{bmatrix} =
\begin{bmatrix}
g(1) & g(2) & g(3) & \ldots \\
g(2) & g(3) & g(4) & \ldots \\
g(3) & g(4) & g(5) & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
u(-1) \\
u(-2) \\
u(-3) \\
\vdots
\end{bmatrix}.
$$

(71)

Thus, the mapping from the input sequence $\{u(-1), u(-2), \ldots\}$ to the output sequence $\{y(0), u(1), \ldots\}$ is described by an infinite-dimensional Hankel matrix. Furthermore, for any input sequence $u$ such that $s^- (u), s^- (y) < \infty$, there exists $N > 0$ such that the subsequences $\{u(-N), u(-N + 1), \ldots u(-1)\}$ and $\{y(0), y(1), \ldots, y(N - 1)\}$ include all the sign variations in $u$ and $y$, respectively. Using certain limiting arguments, the mapping between the number of sign variations in these subsequences can be reduced to analyzing the total positivity properties of a square \textit{finite-dimensional} Hankel matrix. Thus, studying an SVDP in this context can be done using the tools described in Sect. 8 (and some limit arguments like Proposition 4), but there are two twists: (1) we only need to consider Hankel matrices, that are known to have special total positivity properties [21]; and (2) the SVDP can be related to the impulse response and/or transfer function of the system.

The Hankel operator $H_g$ in (70) can also be expressed as

$$
H_g u = O(A, c)(C(A, b)u),
$$

with

$$
C(A, b)u := \sum_{k = -\infty}^{-1} A^{-(k+1)}bu(k),
$$

$$
(O(A, c)z)(j) := c^T A^j z.
$$
Note that $C(A, b)u$ is just $x(0)$, i.e., the state of (67) at time 0.

Any contiguous $q \times q$ submatrix of the infinite Hankel matrix in (71) is of the form

$$H_g(p, q) := \begin{bmatrix}
g(p) & g(p + 1) & \cdots & g(p + q - 1) 
g(p + 1) & g(p + 2) & \cdots & g(p + q) 
\vdots & \vdots & \ddots & \vdots 
g(p + q - 1) & g(p + q) & \cdots & g(p + 2q - 2)
\end{bmatrix},$$

where $p, q \geq 1$. The matrix $H_g(p, q)$ can also be associated with the state-space representation of the system. Define the matrices $C_p(A, b) \in \mathbb{R}^{n \times p}$ and $O_p(A, c) \in \mathbb{R}^{p \times n}$ by

$$C_p(A, b) := \begin{bmatrix} b & Ab & \cdots & A^{p-1}b \end{bmatrix}, \quad O_p(A, c) := \begin{bmatrix} c^T 
c^T A 
\vdots 
c^T A^{p-1} \end{bmatrix}.$$  

Then, a calculation shows that

$$H_g(p, q) = O_q(A, c)A^{p-1}C_q(A, b).$$

The SVDP property for the input–output system is defined as follows.

**Definition 14 [30]** Fix $k \geq 0$. The system (67) is called **Hankel $k$-positive** if for any input sequence $u$ with $s^-(u) \leq k - 1$ the corresponding output sequence $y$ satisfies

1. $s^-(y) \leq s^-(u)$, and
2. if $s^-(y) = s^-(u)$, then the signs of the first nonzero element in $u$ and in $y$ are the same.

Note that here the requirement is for any input sequence $u$ with $s^-(u) \leq k - 1$, and not only for sequences $u$ such that $s^-(u) = k - 1$. It should be clear from the discussion above and Remark 1 that this requirement is closely related to $TN_k$ of the Hankel matrix.

Recall that in general total positivity properties are not preserved under matrix sum (see Example 2). However, if $A, B$ are both Hankel matrices and $TN_r$, then $A + B$ is Hankel and $TN_r$ [21]. This can be used to prove the following.

**Proposition 16 [30]** If $H_{g_1}$ and $H_{g_2}$ are Hankel $k$-positive, then so is $H_{g_1} + H_{g_2}$.

In other words, the parallel interconnection of systems preserves Hankel $k$-positivity.

We now consider several important special cases of Definition 14.
12.1 Hankel 1-positive systems

In a Hankel 1-positive system, the output corresponding to any sequence $u$ with $s^-(u) = 0$ satisfies

$$s^-(y) = 0,$$

and the signs of the first nonzero element in $u$ and in $y$ agree. Assume that at least one entry of $u$ is positive. Then, all the entries of $u$ are non-negative, and all the entries of $y$ are non-negative. Such a system is called externally positive, where the term externally refers to the fact that the I/O mapping is positive, but there are no requirements on the sign variations evolution in the state vector $x$.

It is straightforward to characterize Hankel 1-positive systems in terms of the impulse response $g$. Equation (70) implies that a necessary and sufficient condition for Hankel 1-positivity is that $g(k) \geq 0$ for all $k \geq 1$.

Given $b, c \in \{0, 1\}^n$ and $A \in \{0, 1\}^{n \times n}$, consider the sequence of non-negative numbers $\gamma_k := c^T A^k b$, $k = 0, 1, \ldots$. Note that every $\gamma_k$ satisfies either $\gamma_k = 0$ or $\gamma_k \geq 1$. It is known that determining if this sequence includes a zero is NP-hard in the dimension $n$ [12]. This implies that given the state-space representation (67), it is NP-hard to determine if the system is Hankel 1-positive.

The next example illustrates the difficulty in determining Hankel 1-positivity from the transfer function.

**Example 20** Consider the first order lag

$$G(z) = \frac{r}{z - p}. \quad (72)$$

If $p = 0$, then $g(k) = r \delta(k - 1)$. If $p \neq 0$, then $g(k) = rp^{k-1} 1(k - 1)$. Thus, (72) is Hankel 1-positive iff $p, r \geq 0$, i.e., iff (72) is a positive first-order lag.

As a less trivial example, consider the transfer function:

$$G(z) = \frac{r_1}{z - p_1} + \frac{r_2}{z - p_2},$$

where $p_1$ is a dominant pole, i.e., $|p_1| > |p_2|$. Since

$$g(k) = (r_1 p_1^{k-1} + r_2 p_2^{k-1}) 1(k - 1),$$

we have $g(k) \approx r_1 p_1^{k-1}$ for large values of $k$. Thus, a necessary condition for Hankel 1-positivity is that $p_1 > 0$, $r_1 \geq 0$, but this is not a sufficient condition.

12.2 Hankel 2-positivity and unimodality

The first-order difference of a sequence $s$ is $(\Delta^1 s)(k) := s(k + 1) - s(k)$, i.e., the “discrete-time derivative” of $s$, and the $j$th-order difference for $j > 1$ is defined
inductively by

\[(\Delta^{(j)}s)(k) := (\Delta^{(j-1)}s)(k+1) - (\Delta^{(j-1)}s)(k).\]

For example,

\[(\Delta^{(2)}s)(k) = (\Delta^{(1)}s)(k+1) - (\Delta^{(1)}s)(k) = s(k+2) - s(k+1) - (s(k+1) - s(k)) = s(k+2) - 2s(k+1) + s(k).\]

The linearity of (67) allows to transform an SVDP between \(u\) and \(y\) to an SVDP between \(\Delta^{(j)}u\) and \(\Delta^{(j)}y\). For example,

\[(\Delta^{(1)}y)(k) = y(k+1) - y(k) = (g \ast u)(k+1) - (g \ast u)(k) = g \ast (u(k+1) - u(k)) = g \ast (\Delta^{(1)}u)(k), \tag{73}\]

where \(\ast\) denotes the convolution operator.

A sequence \(u\) is called \emph{unimodal} if \(s^- (\Delta^{(1)}u) \leq 1\) (see, e.g., [35]). For example, \(\delta\) is unimodal, as

\[(\Delta^{(1)}\delta)(k) = \begin{cases} 1, & k = -1, \\ -1, & k = 0, \\ 0, & \text{otherwise}. \end{cases}\]

Equation (73) implies that if (67) is Hankel 2-positive, then it maps any unimodal input to a unimodal output. In particular, since \(\delta\) is unimodal, \(g\) must be unimodal.

\subsection{12.3 Hankel \(\infty\)-positive systems}

Definition 14 implies that for a Hankel \(\infty\)-positive system every input/output pair satisfies \(s^- (y) \leq s^- (u)\), and if \(s^- (y) = s^- (u)\), then the signs of the first nonzero element in \(u\) and in \(y\) are the same. Such systems are also called Hankel totally positive or relaxation systems [79] and have special control-theoretic properties [56]. These systems also admit a simple characterization in terms of their transfer function, namely, a system is Hankel \(\infty\)-positive iff its transfer function has the form

\[G(z) = \sum_{i=1}^{n} \frac{r_i}{z - p_i}\]

with \(r_i, p_i \geq 0\), i.e., it is the parallel interconnection of positive lags.
An important tool in the analysis of Hankel $k$-positivity is the Hankel $k$-compound system.

**Definition 15** [33] Given an impulse response $g$ and an integer $k \geq 1$, the **Hankel $k$-compound system** is the system with impulse response

$$g^{(k)}(j) = \det(H_g(j, k)), \quad j = 1, 2, \ldots$$

The transfer function associated with $g^{(k)}$ is denoted by $G^{(k)}$.

**Example 21** The Hankel 1-compound system has impulse response

$$g^{(1)}(j) = \det(H_g(j, 1)) = g(j),$$

i.e., it is just the original system, and $G^{(1)}(z) = G(z)$. The Hankel 2-compound system has impulse response

$$g^{(2)}(j) = \det(H_g(j, 2))$$

$$= \det\begin{bmatrix} g(j) & g(j + 1) \\ g(j + 1) & g(j + 2) \end{bmatrix}$$

$$= g(j)g(j + 2) - (g(j + 1))^2.$$

Recall from Proposition 5 that in general verifying total positivity of a matrix requires checking only contiguous and initial minors. In general, these results do not extend to checking total non-negativity. For example, all the contiguous minors of $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ are non-negative, yet $A$ is not $TN_3$ as it admits non-contiguous minors that are negative. For the Hankel case, things are simpler.

**Proposition 17** [30] Let $A \in \mathbb{R}^{n \times n}$ be Hankel. If all the contiguous minors of $A$ up to order $r$ are non-negative, then $A$ is $TN_r$.

Using this, it is possible to relate Hankel $k$-positivity to external positivity of the Hankel $k$-compound system.

**Proposition 18** [30] A DT SISO LTI system with transfer function $G$ is Hankel $k$-positive iff $G^{(k)}$ is externally positive.

In this respect, the construction of $G^{(k)}$ is another demonstration of the general principal described in Sect. 6.

### 13 Discussion

We introduced a principle that allows to generalize, in a non-trivial way, important classes of linear and nonlinear dynamical systems. If the LTV system (29) satisfies a specific property, we say that the LTV satisfies $k$-property if the associated
\( k \)-compound system satisfies this property. This has been used to generalize important classes of systems including contracting, positive, cooperative, and diagonally stable systems into \( k \)-contracting, \( k \)-positive, \( k \)-cooperative, and \( k \)-diagonally stable systems (see Fig. 10). This approach makes sense for at least two reasons. First, for \( k = 1 \) the generalization reduces to the original class of systems, as the 1-compound system is just the original system. Second, the \( k \)-compound system is based on \( k \)-compound matrices, and these matrices play an important role in describing the evolution of \( k \)-parallelotopes and the evolution of sign changes along the dynamics.

In some cases, these generalizations lead to a natural hierarchy. For example, systems that are \( k \)-contracting w.r.t. to the \( L_1 \), \( L_2 \) or \( L_\infty \) norms are also \( \ell \)-contracting w.r.t. to the same norm for any integer \( \ell \geq k \).

We believe that the ideas described in this tutorial paper may lead to interesting research directions. We now describe several possible directions.

### 13.1 Linear systems with inputs

In the context of systems and control theory, it is important to consider systems with inputs and outputs. We reviewed one possible approach for analyzing an SVDP for such systems in Sect. 12. A possible alternative is to use a formula for analyzing the evolution of compounds in systems with an input from [50]. Consider the set of time-varying non-autonomous ODEs:

\[
\dot{x}^i(t) = A(t)x^i(t) + f^i(t), \quad i = 1, \ldots, k. \tag{74}
\]

Then, it is shown in [50] that:

\[
\frac{d}{dt} \left[ x^1 \ldots x^k \right]^{(k)} = A^{[k]} \left[ x^1 \ldots x^k \right]^{(k)} + \sum_{i=1}^{k} \left[ x^1 \ldots x^{i-1} f^i x^{i+1} \ldots x^k \right]^{(k)}. \tag{75}
\]

Note that in the autonomous case, i.e., when \( f^i \equiv 0 \) for all \( i \), this reduces to (31).

Many of the notions described in this paper may perhaps be extended to such systems. For example, Angeli and Sontag extended cooperative systems (more generally, monotone systems) to systems with inputs [4]. It may be of interest to extend the notion of \( k \)-positive systems and \( k \)-cooperative systems to the case of systems with inputs.

As another example, a fundamental notion in systems and control theory is input-to-state stability (ISS) [69]. It may be of interest to consider a \( k \) generalization of this notion. To explain the basic idea, consider the LTI

\[
\dot{x} = Ax + 12u, \tag{76}
\]

with \( A = \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix} \) and \( 12 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T \). Since \( A \) is not Hurwitz, this system is not ISS. Fix two initial conditions \( a, b \in \mathbb{R}^2 \), and let \( \Phi(t) := \begin{bmatrix} x(t, a) & x(t, b) \end{bmatrix} \). Applying (75)
Fig. 10 Generalizing various types of dynamical systems using the $k$-compound system.
with $k = 2$ gives

$$\frac{d}{dt} \det(\Phi) = -2 \det(\Phi) + \left[ x(t, a) \right]^{(2)} + \left[ 12u \right]^{(2)} + \left[ x(t, b) \right]^{(2)},$$

and substituting the solution of (76) gives

$$\frac{d}{dt} \det(\Phi) = -2 \det(\Phi) + (a_1 - a_2 + b_2 - b_1) \exp(-4t)u.$$

This suggests that two-dimensional volumes are in some sense input to state stable (ISS) (see [69]), but with a gain function that depends on the initial conditions. An interesting research topic is formulating these notions rigorously, and studying their implications for interconnected systems.

13.2 $k$-contraction on networks

There is considerable interest in dynamical system over networks motivated in part by applications in multi-agent systems, the electric grid, neuroscience, epidemiology, and more [13, 39, 49].

Wang and Slotine [76] suggested an approach for the stability analysis and control synthesis of networked systems using contraction theory. It may be of interest to extend these results using $k$-contraction and, more generally, to formulate a suitable notion of the $k$-compound system of such networks and their associated connection graphs.

13.3 Kernels satisfying an SVDP and applications to PDEs

We focused here on matrices satisfying an SVDP and its implications to dynamical systems described by ODEs. There is, however, a rich theory on kernels that satisfy an SVDP [36]. These have found some applications in the analysis of PDEs (see, e.g., [55, 64] and the references therein), but we believe that this field is still largely unexplored.

Acknowledgements We thank the anonymous reviewers for many helpful comments. It is written in the Mishna that one who learns from his fellow a single letter must treat him with great honor. The third author would like to thank E. D. Sontag for teaching him a whole language.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

Appendix: Proof of Theorem 1

We begin with an auxiliary result.
Lemma 3 Let $A \in \mathbb{C}^{n \times m}$, $B \in \mathbb{C}^{m \times n}$. Then,
\[
\det(AB) = \begin{cases} 
\sum_{\alpha \in Q(n,m)} A([1, \ldots, n]|\alpha) B(\alpha|[1, \ldots, n]), & \text{if } n \leq m, \\
0, & \text{if } n > m.
\end{cases}
\]

Proof Let $D := \begin{bmatrix} 0 & A \\ B & I_m \end{bmatrix} \in \mathbb{C}^{(n+m) \times (n+m)}$. Since
\[
\begin{bmatrix} I_n & -A \\ 0 & I_m \end{bmatrix} D = \begin{bmatrix} -AB & 0 \\ B & I_m \end{bmatrix},
\]
we have
\[
\det(D) = \det(-AB) = (-1)^n \det(AB). \tag{77}
\]

The Laplace expansion of $D$ using minors that include the first $n$ rows gives:
\[
\det(D) = \sum_{\alpha \in Q(n,n+m)} D([1, \ldots, n]|\alpha) \cof(D([1, \ldots, n]|\alpha)), \tag{78}
\]
where $\cof(D(\alpha|\beta)) := (-1)^{\alpha_1+\cdots+\alpha_n+\beta_1+\cdots+\beta_n} D(([1, \ldots, n + m]|\alpha)|([1, \ldots, n + m]|\beta))$ is the cofactor of $D$ corresponding to $(\alpha|\beta)$. If $m < n$ then any submatrix in the form $D([1, \ldots, n]|\alpha)$ includes a column of zeros, so $D([1, \ldots, n]|\alpha) = 0$ and $\det(D) = 0$. Combining this with (77) yields $\det(AB) = 0$.

Now, suppose that $m \geq n$. If $\alpha$ includes any of the indexes $1, \ldots, n$, then $D([1, \ldots, n]|\alpha)$ includes a column of zeros. Using this and the fact that rows $n+1, \ldots, n+m$ of $D$ includes rows of $B$ and of $I_m$ in (78) gives
\[
\det(D) = (-1)^n \sum_{\beta \in Q(n,m)} A([1, \ldots, n]|\beta) B(\beta|[1, \ldots, n]),
\]
and combining this with (77) completes the proof of Lemma 3. \qed

We can now prove Theorem 1. Let $P \in \mathbb{C}^{n \times m}$, $Q \in \mathbb{C}^{m \times s}$, and fix $k \in \{1, \ldots, \min\{n,m,s\}\}$. Let $C := PQ \in \mathbb{C}^{n \times s}$. Fix $\alpha \in Q(k,n)$ and $\beta \in Q(k,s)$. Since $c_{ij} = \sum_{\ell=1}^{m} p_{i\ell} q_{\ell j}$,
\[
C(\alpha|\beta) = P(\alpha|[1, \ldots, m]) Q([1, \ldots, m]|\beta).
\]

Let
\[
A := P(\alpha|[1, \ldots, m]) \in \mathbb{C}^{k \times m}, \tag{79}
B := Q([1, \ldots, m]|\beta) \in \mathbb{C}^{m \times k}.
\]
so that $C[\alpha|\beta] = AB$. Applying Lemma 3 gives

$$C(\alpha|\beta) = \det(AB) = \sum_{\gamma \in Q(k,m)} A([1, \ldots, k]|\gamma) B(\gamma|[1, \ldots, k]),$$

and using (79) gives

$$C(\alpha|\beta) = \sum_{\gamma \in Q(k,m)} P(\alpha|\gamma) Q(\gamma|\beta).$$

This completes the proof of Theorem 1.

References

1. Alseidi R, Margaliot M, Garloff J (2019) On the spectral properties of nonsingular matrices that are strictly sign-regular for some order with applications to totally positive discrete-time systems. J Math Anal Appl 474:524–543
2. Alseidi R, Margaliot M, Garloff J (2021) Discrete-time $k$-positive linear systems. IEEE Trans Autom Control 66(1):399–405
3. Aminzare Z, Sontag ED (2014) Contraction methods for nonlinear systems: a brief introduction and some open problems. In: Proceedings of 53rd IEEE conference on decision and control, Los Angeles, CA, pp 3835–3847
4. Angeli D, Sontag ED (2003) Monotone control systems. IEEE Trans Autom Control 48(9):1531–1537
5. Bar-Shalom E, Margaliot M (2021) Compound matrices in systems and control theory. In: Proceedings of 60th IEEE conference on decision and control, pp 5802–5807
6. Barker G, Berman A, Plemmons RJ (1978) Positive diagonal solutions to the Lyapunov equations. Linear Multilinear Algebra 5(4):249–256
7. Basios V, Antonopoulos CG, Latifi A (2020) Labyrinth chaos: revisiting the elegant, chaotic, and hyperchaotic walks. Chaos 30(11):113129
8. Ben-Avraham T, Sharon G, Zarai Y, Margaliot M (2020) Dynamical systems with a cyclic sign variation diminishing property. IEEE Trans Autom Control 65:941–954
9. Berger T, Halikias G, Karcanias N (2015) Effects of dynamic and non-dynamic element changes in RC and RL networks. Int J Circuit Theory Appl 43(1):36–59
10. Berman A, Plemmons RJ (1987) Nonnegative matrices in the mathematical sciences. SIAM
11. Blondel VD, Portier N (2002) The presence of a zero in an integer linear recurrent sequence is NP-hard to decide. Linear Algebra Appl 351–352:91–98
12. Bullo F (2018) Lectures on network systems. CreateSpace Independent Publishing Platform
13. Bullo F (2022) Contraction theory for dynamical systems. Kindle Direct Publishing. http://motion.me.ucsb.edu/book-ctds
14. Chicone C (2006) Ordinary differential equations with applications, 2nd edn. Springer, New York
15. Coppel WA (1965) Stability and asymptotic behavior of differential equations. D. C. Heath, Boston
16. Dalin O, Ofir R, Bar-Shalom E, Ovseevich A, Bullo F, Margaliot M (2022) Verifying $k$-contraction without computing $k$-compounds. arXiv:2209.01046
17. Douady A, Oesterlé J (1980) Dimension de Hausdorff des attracteurs. C R Acad Sci Paris 290:1135–1138
18. Elkhader AS (1992) A result on a feedback system of ordinary differential equations. J Dyn Differ Equ 4(3):399–418
19. Fallat SM, Johnson CR (2011) Totally nonnegative matrices. Princeton University Press, Princeton

Springer
21. Fallat S, Johnson CR, Sokal AD (2017) Total positivity of sums, Hadamard products and Hadamard powers: results and counterexamples. Linear Algebra Appl 520:242–259
22. Farina L, Rinaldi S (2000) Positive linear systems: theory and applications. Wiley
23. Fiedler M (2008) Special matrices and their applications in numerical mathematics, 2nd edn. Dover Publications, Mineola
24. Fusco G, Oliva WM (1988) Jacobi matrices and transversality. Proc R Soc Edinb Sect A 109(3–4):231–243
25. Gantmacher FR (1960) The theory of matrices, vol I. Chelsea Publishing Company
26. Gantmacher FR, Krein MG (2002) Oscillation matrices and kernels and small vibrations of mechanical systems. American Mathematical Society, Providence (translation based on the 1941 Russian original)
27. Gover E, Krikorian N (2010) Determinants and the volumes of parallelotopes and zonotopes. Linear Algebra Appl 433(1):28–40
28. Grinberg D (2020) Notes on the combinatorial fundamentals of algebra. arXiv:2008.09862
29. Grussler C, Sepulchre R (2019) Strongly unimodal systems. In: Proceedings of 18th European Control Conference, pp 3273–3278
30. Grussler C, Sepulchre R (2020) Variation diminishing Hankel operators. In: Proceedings of 59th IEEE conference on decision and control, pp 4529–4534
31. Grussler C, Burghi T, Sojoudi S (2022) Internally Hankel k-positive systems. SIAM J Control Optim 60(4):2373–2392
32. Grussler C, Damm T, Sepulchre R (2022) Balanced truncation of k-positive systems. IEEE Trans Autom Control 67(1):526–531
33. Grussler C, Sepulchre R (2022) Variation diminishing linear time-invariant systems. Automatica 136:109985
34. Horn RA, Johnson CR (2013) Matrix analysis, 2nd edn. Cambridge University Press
35. Ibragimov I (1956) On the composition of unimodal distributions. Theory Probab Appl 1(2):255–260
36. Kaszkurewicz E, Bhaya A (2012) Matrix diagonal stability in systems and computation. Springer
37. Katz R, Margaliot M, Fridman E (2020) Entrainment to subharmonic trajectories in oscillatory discrete-time systems. Automatica 116:108919
38. Kiss IZ, Miller JC, Simon PL (2017) Mathematics of epidemics on networks: from exact to approximate models. Springer, Cham
39. Krasnoselskii M, Lifshits JA, Sobolev AV (1989) Positive linear systems: the method of positive operators. Heldermann Verlag
40. Li MY, Muldowney JS (1995) Global stability for the SEIR model in epidemiology. Math Biosci 125(2):155–164
41. Li MY, Muldowney JS (1995) On R. A. Smith’s autonomous convergence theorem. Rocky Mt J Math 25(1):365–378
42. Li MY, Wang L (1998) A criterion for stability of matrices. J Math Anal Appl 225:249–264
43. Lohmiller W, Slotine JJE (1998) On contraction analysis for non-linear systems. Automatica 34:683–696
44. Mallet-Paret J, Smith HL (1990) The Poincaré-Bendixson theorem for monotone cyclic feedback systems. J Dyn Differ Equ 2(4):367–421
45. Manchester IR, Slotine J-JE (2014) Combination properties of weakly contracting systems. arXiv:1408.5174
46. Margaliot M, Sontag ED (2019) Compact attractors of an antithetic integral feedback system have a simple structure. https://www.biorxiv.org/content/10.1101/868000v1
47. Margaliot M, Sontag ED (2019) Revisiting totally positive differential systems: a tutorial and new results. Automatica 101:1–14
48. Mesbahi M, Egerstedt M (2010) Graph-theoretic methods in multiagent networks. Princeton University Press, Princeton
49. Muldowney J, Samuylova E (2010) Dimension problems for nonhomogeneous differential equations. Can Appl Math Q 18(1):93–105
50. Muldowney JS (1990) Compound matrices and ordinary differential equations. Rocky Mt J Math 20(4):857–872
51. Ofir R, Ovseevich A, Margaliot M (2022) A sufficient condition for k-contraction in Lurie systems. arXiv:2211.05696
53. Ofir R, Margaliot M, Levron Y, Slotine J-J (2021) Serial interconnections of 1-contracting and 2-contracting systems. In: Proceedings of 60th IEEE conference on decision and control, pp 3906–3911
54. Ofir R, Margaliot M, Levron Y, Slotine J-J (2022) A sufficient condition for k-contraction of the series connection of two systems. IEEE Trans Autom Control 67(9):4994–5001
55. Oliva WM, Kuhl NM, Magalhaes LT (1993) Diffeomorphisms of \( \mathbb{R}^n \) with oscillatory Jacobians. Publ Mat 37(2):255–269
56. Pates R, Bergeling C, Rantzer A (2019) On the optimal control of relaxation systems. In: Proceedings of 58th IEEE conference on decision and control, pp 6068–6073
57. Pinkus A (2010) Totally positive matrices. Cambridge University Press, Cambridge
58. Pinkus A (1996) Spectral properties of totally positive kernels and matrices. In: Gasca M, Micchelli CA (eds) Total positivity and its applications. Springer Netherlands, Dordrecht, pp 477–511
59. Rantzer A (2015) Scalable control of positive systems. Eur J Control 24:72–80
60. Rantzer A, Valcher ME (2021) Scalable control of positive systems. Annu Rev Control Robot Auton Syst 4(1):319–341
61. Sanchez LA (2009) Cones of rank 2 and the Poincaré-Bendixson property for a new class of monotone systems. J Differ Equ 246(5):1978–1990
62. Schwarz B (1970) Totally positive differential systems. Pac J Math 32(1):203–229
63. Smillie J (1984) Competitive and cooperative tridiagonal systems of differential equations. SIAM J Math Anal 15:530–534
64. Smith HL (1990) A discrete Lyapunov function for a class of linear differential equations. Pac J Math 144(2):345–360
65. Smith HL (1991) Periodic tridiagonal competitive and cooperative systems of differential equations. SIAM J Math Anal 22(4):1102–1109
66. Smith HL (1995) Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems, Ser. Mathematical Surveys and Monographs, vol 41. Amer. Math. Soc., Providence
67. Smith RA (1986) Some applications of Hausdorff dimension inequalities for ordinary differential equations. Proc R Soc Edinb Sect A Math 104(3–4):235–259
68. Sontag ED (2007) Monotone and near-monotone biochemical networks. Syst Synth Biol 1:59–87
69. Sontag E (2020) Input-to-state stability. In: Baillieul J, Samad T (eds) Encyclopedia of systems and control. Springer, pp 1–9
70. Ström T (1975) On logarithmic norms. SIAM J Numer Anal 12(5):741–753
71. Tam T-Y. Additive compound matrices and representation of \( \mathfrak{gl}(n) \). https://www.semanticscholar.org/paper/Additive-Compound-Matrices-and-Representation-of-Gl-Tam/cf3e201dabe1f3e373a483f09d401dc443e781c
72. Tanaka T, Langbort C (2011) The bounded real lemma for internally positive systems and H-infinity structured static state feedback. IEEE Trans Autom Control 56(9):2218–2223
73. Thomas R (1999) Deterministic chaos seen in terms of feedback circuits: analysis, synthesis, “labyrinth chaos”. Int J Bifurc Chaos 9(10):1889–1905
74. Vidyasagar M (1978) Nonlinear systems analysis. Prentice Hall, Englewood Cliffs
75. Wang R, Jing Z, Chen L (2005) Modelling periodic oscillation in gene regulatory networks by cyclic feedback systems. Bull Math Biol 67(2):339–367
76. Wang W, Slotine JJ (2005) On partial contraction analysis for coupled nonlinear oscillators. Biol Cybern 92:38–53
77. Weiss E, Margaliot M (2021) A generalization of linear positive systems with applications to nonlinear systems: invariant sets and the Poincaré-Bendixson property. Automatica 123:109358
78. Weiss E, Margaliot M (2021) Is my system of ODEs k-cooperative? IEEE Control Syst Lett 5(1):73–78
79. Willems JC (1976) Realization of systems with internal passivity and symmetry constraints. J Frankl Inst 301(6):605–621
80. Wimmer HK (2009) Diagonal stability of matrices with cyclic structure and the secant condition. Syst Control Lett 58(5):309–313
81. Wu C, Grüne L, Kriecherbauer T, Margaliot M (2021) Behavior of totally positive differential systems near a periodic solution. In: Proceedings of 60th IEEE conference on decision and control, pp 3160–3165
82. Wu C, Kanevskiy I, Margaliot M (2022) k-contraction: theory and applications. Automatica 136:110048
83. Wu C, Margaliot M (2022) Diagonal stability of discrete-time $k$-positive linear systems with applications to nonlinear systems. IEEE Trans Autom Control 67(8):4308–4313
84. Wu C, Pines R, Margaliot M, Slotine J-J (2022) Generalization of the multiplicative and additive compounds of square matrices and contraction theory in the Hausdorff dimension. IEEE Trans Autom Control 67(9):4629–4644

Publisher’s Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.