UNIT TRIANGULAR FACTORIZATION OF THE MATRIX SYMPLECTIC GROUP*

PENGZHAN JIN†, YIFA TANG†, AND AIQING ZHU†

Abstract. In this work, we prove that any symplectic matrix can be factored into no more than 9 unit triangular symplectic matrices. This structure-preserving factorization of the symplectic matrices immediately reveals two well-known features that (i) the determinant of any symplectic matrix is one and (ii) the matrix symplectic group is path connected, as well as a new feature that (iii) all the unit triangular symplectic matrices form a set of generators of the matrix symplectic group. Furthermore, this factorization yields effective methods for the unconstrained parametrization of the matrix symplectic group as well as its structured subsets. The unconstrained parametrization enables us to apply faster and more efficient unconstrained optimization algorithms to the problems with symplectic constraints under certain circumstances.

1. Introduction. Denote the d-by-d identity matrix by I_d; let

$$J := \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix},$$

which is an orthogonal, skew-symmetric real matrix, so that $J^{-1} = J^T = -J$.

DEFINITION 1.1. *A matrix $H \in \mathbb{R}^{2d \times 2d}$ is called symplectic if $H^T J H = J$.*

We denote the collection of symplectic matrices by

$$SP = \{H \in \mathbb{R}^{2d \times 2d} | H^T J H = J\},$$

which forms a group, i.e., matrix symplectic group. The group SP is relevant from both the pure mathematical point of view [17] and the point of view of applications. For instance, symplectic matrices play an important role in classical mechanics and Hamiltonian dynamical systems [1, 2, 3]. The symplectic matrices also arise in the symplectic integrators which are the preferred methods for the numerical solutions of the differential equations appearing in these physical problems preserving the Hamiltonian structure [14, 19, 32]. Moreover, it appears in linear control [13, 29], optimal control [20], and the theory of parametric resonance [22]. The applications have motivated the development of computing certain specific factorizations of symplectic matrices [5].

---

†LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (jpz@lsec.cc.ac.cn, tyf@lsec.cc.ac.cn, zaq@lsec.cc.ac.cn).

There have been several modern factorizations of symplectic matrices. Mehrmann [28] provides the block LDU (lower-diagonal-upper) factorization of symplectic matrices which require the nonsingular left upper block. Dopico and Johnson [11, 12] extend the LDU factorization to adapt general symplectic matrices by symplectic interchanges and permutation matrices. There are also QR-like factorization [9], polar factorization [21], as well as symplectic SVD (singular value decomposition)-like factorization [36]. In addition, [16, 27] express the $2d$-by-$2d$ symplectic matrices by the products of at most $4d$ symplectic transvections.

As aforementioned, we briefly present these modern factorizations as follows.

- **Block LDU parametrization.** Before describing the block LDU parametrization, we first show the definition of symplectic interchange.

DEFINITION 1.2. *Let $1 \leq j \leq d$. The symplectic interchange matrix Π_j is the $2d$-by-$2d$ matrix obtained by interchanging the columns j and $j + d$ of the $2d$-by-$2d$ identity matrix and multiplying the jth column of the resulting matrix by -1. The symplectic interchange matrix $\widetilde{\Pi}_j$ is the $2d$-by-$2d$ matrix obtained by interchanging the columns j and $j + d$ of the $2d$-by-$2d$ identity matrix and multiplying the $j + d$th column of the resulting matrix by -1. Notice that $\Pi_j^T = \widetilde{\Pi}_j$.*

With the above definition, there is the block LDU parametrization.

THEOREM 1.3. *The set of $2d$-by-$2d$ symplectic matrices is*

$$SP = \left\{ Q \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & G^{-T} \end{bmatrix} \begin{bmatrix} I & E \\ 0 & I \end{bmatrix} \middle| G \in \mathbb{R}^{d \times d} \text{ nonsingular}, \right.$$

$$\left. C = C^T, \ E = E^T, \ Q \text{ a product of symplectic interchanges} \right\}.$$

The detailed information of the block LDU parametrization can be found in [12].

- **QR-like factorization.**

THEOREM 1.4. *The set of $2d$-by-$2d$ symplectic matrices is*

$$SP = \left\{ Q \begin{bmatrix} R & Z \\ 0 & R^{-T} \end{bmatrix} \middle| \begin{matrix} R \in \mathbb{R}^{d \times d} \text{ upper triangular} \\ Q \text{ symplectic orthogonal} \end{matrix} \right\}.$$

The detailed information of the QR-like factorization can be found in [9].

- **Polar factorization.**

THEOREM 1.5. *The set of $2d$-by-$2d$ symplectic matrices is*

$$SP = \left\{ QP \middle| \begin{matrix} P = P^T \text{ symplectic positive definite} \\ Q \text{ symplectic orthogonal} \end{matrix} \right\}.$$

The detailed information of the polar factorization can be found in [21].

- **SVD-like factorization.**

THEOREM 1.6. *The set of $2d$-by-$2d$ symplectic matrices is*

$$SP = \left\{ U \begin{bmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{bmatrix} V^T \middle| \begin{matrix} U, V \text{ symplectic orthogonal} \\ \Omega = diag(\omega_1, \ldots, \omega_d) \\ \omega_1 \geq \cdots \geq \omega_d \geq 1 \end{matrix} \right\}.$$

The detailed information of the SVD-like factorization can be found in [36].

- **Transvections factorization.** In text books on classical groups like [4, 33, 35] this result is mostly stated as a corollary of another basic fact, namely, that the symplectic group is generated by so-called symplectic transvections.

DEFINITION 1.7. *For $0 \neq u \in \mathbb{R}^{2d}$ and $0 \neq \beta \in \mathbb{R}$, the matrix*

$$G = I + \beta u u^T J \in \mathbb{R}^{2d \times 2d}$$

is symplectic, and G is called a symplectic transvection.

Any symplectic matrix can be factored as the product of many symplectic transvections.

THEOREM 1.8. *The set of $2d$-by-$2d$ symplectic matrices is*

$$SP = \left\{ G_1 G_2 \cdots G_m \middle| \begin{matrix} G_i \ a \ symplectic \ transvection \\ m \leq 4d \end{matrix} \right\}.$$

Although this factorization can freely parameterize the symplectic matrices, in numerical practice, the computed product of lots of symplectic matrices may be far from being symplectic when d is large. There are some works parameterizing the symplectic matrices as the finite product of certain elementary units [24, 30], however, which take effect on only special structured symplectic matrices. The detailed information of the transvections factorization can be found in [16, 27].

In this work, we will show a more elementary factorization of the symplectic matrices. Considering

(1.1)
$$\begin{bmatrix} I & S \\ 0 & I \end{bmatrix}, \quad \begin{bmatrix} I & 0 \\ S & I \end{bmatrix},$$

where S is symmetric, we will factor the symplectic matrices as the products of at most 9 *unit triangular symplectic matrices* shown above. In addition, we call the symplectic matrices like

(1.2)
$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}$$

as *diagonal symplectic matrices*. These two types of symplectic matrices are concerned as the basic units in our factorizations.

Since the symplectic matrix with nonsingular left upper block is easy to deal with, in classical factorizations, researchers have to exploit some intractable factors such as symplectic interchanges and permutation matrices to make a general symplectic matrix equipped with nonsingular left upper block. Different from the conventional methods, we find a family of unit triangular symplectic matrices which can transform a general symplectic matrix into one having nonsingular left upper block; it is the core of the theoretical results in this work. With this result, we are able to factor the symplectic matrix into three unit triangular symplectic matrices and one diagonal symplectic matrix, called *unit ULU (upper-lower-upper) factorization*. Furthermore, we decompose the diagonal symplectic matrix into several unit triangular symplectic matrices and subsequently obtain the final *unit triangular factorization*.

The unit triangular factorization immediately leads to many significant properties of the matrix symplectic group, such as (i) the determinant of symplectic matrix is one, (ii) the matrix symplectic group is path connected, and (iii) all the unit triangular

symplectic matrices form a set of generators of the matrix symplectic group. In addition, we will consider some structured subsets of the matrix symplectic group such as the symplectic matrices with nonsingular left upper block, the positive definite symplectic matrices, and the symplectic M-matrices, which can be factored into less than 9 unit triangular symplectic matrices due to their special structures.

A remarkable characteristic of the unit triangular factorization is that it provides an unconstrained parametrization for symplectic matrices. For the general symplectic matrices, the positive definite symplectic matrices and the singular symplectic matrices, we will explore their unconstrained parametrization, respectively. With the unconstrained parametrization, one may be able to apply faster and more efficient unconstrained optimization algorithms to the problems with symplectic constraints under certain circumstances, which will be discussed in detail later. It was in fact our work on optimization problem with symplectic constraints that led us to consider the unit triangular factorization of the symplectic matrices.

The remaining parts of this paper are organized as follows. Some basic properties that will be used later are given in section 2. Section 3 presents the detailed proofs of the factorizations and deduces several important inferences. In section 4, we provide the unconstrained parametrization for the matrix symplectic group and its structured subsets. Some conclusions will be given in the last section.

2. Preliminaries. At first, we introduce the notations that will be used as well as some relevant basic results for factorization of symplectic matrices. As shown in Definition 1.1, in this work we consider the real $2d$-by-$2d$ matrix symplectic group SP. Properties 2.1–2.3 are well-known and can be easily verified according to the definition of SP. We omit the proofs and the readers are referred to [10, 14] for more details. Since the features of symplectic matrices are not apparent in Definition 1.1, there is an equivalent condition of being symplectic.

PROPERTY 2.1. *If $H = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \in \mathbb{R}^{2d \times 2d}$ is a symplectic matrix and $A_i, B_i \in \mathbb{R}^{d \times d}$, then*
(i) $A_1^T A_2 = A_2^T A_1$,
(ii) $B_1^T B_2 = B_2^T B_1$,
(iii) $A_1^T B_2 - A_2^T B_1 = I$,
vice versa. Note that an $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \in \mathbb{R}^{2d \times d}$ satisfying the first item is called a symmetric pair.

Property 2.1 points out the detailed relationship among the four blocks in a symplectic matrix; it plays an important role in unpacking the features. Based on Property 2.1, one may check that a symplectic matrix being unit and triangular must be written in form (1.1). Besides the inner structure of symplectic matrix, there also hold several overall natures, for instance, SP remains closed under multiplication, transposition, and inversion.

PROPERTY 2.2. *If $H, G \in SP$, then*
(i) $HG \in SP$,
(ii) $H^T \in SP$,
(iii) $H^{-1} \in SP$.

With the closure under multiplication and inversion, we can easily check that SP immediately forms a group. In addition, the second item in Property 2.2 indicates that the equivalent condition in Property 2.1 can be written in another way:

$$(2.1) \quad \begin{cases} A_1 B_1^T = B_1 A_1^T, \\ A_2 B_2^T = B_2 A_2^T, \\ A_1 B_2^T - B_1 A_2^T = I. \end{cases}$$

In this work, we are aiming to seek elementary factorizations based on only (1.1) and (1.2), which are the most concise symplectic matrices. To figure out how the unit triangular matrices and the diagonal matrices are symplectic, it is shown in the following property.

PROPERTY 2.3. *If* $S, P, Q, A, B, C \in \mathbb{R}^{d \times d}$, *then*
(i) *The matrix* $\begin{bmatrix} I & S \\ 0 & I \end{bmatrix}$ *is symplectic if and only if* $S^T = S$,
(ii) *The matrix* $\begin{bmatrix} I & 0 \\ S & I \end{bmatrix}$ *is symplectic if and only if* $S^T = S$,
(iii) *The matrix* $\begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$ *is symplectic if and only if* $Q = P^{-T}$.

The mentioned two types of matrices, i.e., the unit triangular symplectic matrices and the diagonal symplectic matrices, have satisfactory properties and clear structures. In the later sections we will think of a way to do the unconstrained parametrization of symplectic matrices and its structured subsets; for this purpose, the unit triangular symplectic matrices (1.1) are more preferred, since each one is determined by $d(d+1)/2$ free parameters while the diagonal symplectic matrices (1.2) require the nonsingularity on P.

Although the determinant of symplectic matrix is obviously ± 1, it is not easy to confirm that the determinant is always 1. However, in some special cases, we can prove it without any difficulty. For instance, when the symplectic matrix has a nonsingular left upper block, we immediately derive such a LDU factorization whose factors are simple enough to obtain the determinants.

PROPERTY 2.4 (LDU Factorization). *If* $H = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \in \mathbb{R}^{2d \times 2d}$ *is a symplectic matrix and* $A_i, B_i \in \mathbb{R}^{d \times d}$; *moreover* A_1 *is nonsingular; then* H *has three unique factorizations,*

$$(2.2) \quad \begin{cases} H = \begin{bmatrix} P_1 & 0 \\ 0 & P_1^{-T} \end{bmatrix} \begin{bmatrix} I & 0 \\ S_1 & I \end{bmatrix} \begin{bmatrix} I & T_1 \\ 0 & I \end{bmatrix}, \\ S_1 = A_1^T A_2, \ T_1 = A_1^{-1} B_1, \ P_1 = A_1, \end{cases}$$

$$(2.3) \quad \begin{cases} H = \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} P_2 & 0 \\ 0 & P_2^{-T} \end{bmatrix} \begin{bmatrix} I & T_2 \\ 0 & I \end{bmatrix}, \\ S_2 = A_2 A_1^{-1}, \ T_2 = A_1^{-1} B_1, \ P_2 = A_1, \end{cases}$$

$$(2.4) \quad \begin{cases} H = \begin{bmatrix} I & 0 \\ S_3 & I \end{bmatrix} \begin{bmatrix} I & T_3 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_3 & 0 \\ 0 & P_3^{-T} \end{bmatrix}, \\ S_3 = A_2 A_1^{-1}, \ T_3 = B_1 A_1^T, \ P_3 = A_1, \end{cases}$$

where $S_1, S_2, S_3, T_1, T_2, T_3$ *are symmetric and* P_1, P_2, P_3 *are nonsingular.*

Proof. Here we show the proof of (2.3), and the others are the same. Assume that S_2, T_2, P_2 are defined as in (2.3); then

$$S_2^T = (A_2 A_1^{-1})^T = A_1^{-T} A_2^T A_1 A_1^{-1} = A_1^{-T} A_1^T A_2 A_1^{-1} = A_2 A_1^{-1} = S_2,$$
$$T_2^T = (A_1^{-1} B_1)^T = A_1^{-1} A_1 B_1^T A_1^{-T} = A_1^{-1} B_1 A_1^T A_1^{-T} = A_1^{-1} B_1 = T_2,$$

which mean S_2, T_2 are symmetric. Furthermore,

$$P_2^{-T} + S_2 P_2 T_2 = A_1^{-T} + A_2 A_1^{-1} A_1 A_1^{-1} B_1 = A_1^{-T}(I + A_1^T A_2 A_1^{-1} B_1)$$
$$= A_1^{-T}(I + A_2^T A_1 A_1^{-1} B_1)$$
$$= A_1^{-T} A_1^T B_2$$
$$= B_2;$$

therefore

$$\begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} P_2 & 0 \\ 0 & P_2^{-T} \end{bmatrix} \begin{bmatrix} I & T_2 \\ 0 & I \end{bmatrix} = \begin{bmatrix} P_2 & P_2 T_2 \\ S_2 P_2 & P_2^{-T} + S_2 P_2 T_2 \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} = H.$$

The uniqueness is easy to check, and we omit it.   □

Property 2.4 decomposes the symplectic matrices with nonsingular left upper block into two unit triangular symplectic matrices and one diagonal symplectic matrices, where the diagonal one can be placed at an arbitrary position. What we expect is to derive a similar result for the general symplectic matrices.

In order to complete our proofs, we also have to give the other properties for checking if some columns are a part of a symplectic matrix. The relevant properties can be found in [14, 18].

PROPERTY 2.5. *If $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \in \mathbb{R}^{2d \times k}, 1 \le k \le d$, A is full-rank, and $A_1^T A_2 = A_2^T A_1$, then there exists a $A' \in \mathbb{R}^{2d \times (d-k)}$ such that $\begin{bmatrix} A & A' \end{bmatrix}$ is a full-rank symmetric pair.*

Proof. It is shown as [14, p. 147, Theorem 3.27].   □

PROPERTY 2.6. *If $A \in \mathbb{R}^{2d \times d}$ is a full-rank symmetric pair, then there exists a $B \in \mathbb{R}^{2d \times d}$ such that $\begin{bmatrix} A & B \end{bmatrix} \in SP$.*

Proof. It is shown as [14, p. 147, Theorem 3.28].   □

At the end of this section, we provide a table including the main notations in this paper. See Table 1.

3. Unit triangular factorization. In this section, we will present the detailed proof of the unit triangular factorization. To describe the problem clearly, denote

$$\mathcal{L}_n = \left\{ \begin{bmatrix} I & 0/S_n \\ S_n/0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix} \middle| S_i \in \mathbb{R}^{d \times d}, S_i^T = S_i, i = 1, 2, \ldots, n \right\},$$

where the unit upper triangular symplectic matrices and the unit lower triangular symplectic matrices appear alternately. And it is clear that $\mathcal{L}_m \subset \mathcal{L}_n \subset SP$ for all integers $1 \le m \le n$. Now the main theorem is given as follows.

THEOREM 3.1 (unit triangular factorization). *$SP = \mathcal{L}_9$. Thus any symplectic matrix can be factored into no more than 9 unit triangular symplectic matrices.*

To prove this theorem, we first list some necessary auxiliary results. For convenience, we denote all the unimportant blocks in matrices by "$\star$" throughout this paper.

THEOREM 3.2. *For any symplectic matrix $H \in \mathbb{R}^{2d \times 2d}$, there exists a symmetric $S \in \mathbb{R}^{d \times d}$ such that, the factorization*

$$H = \begin{bmatrix} I & \lambda S \\ 0 & I \end{bmatrix} \begin{bmatrix} P_\lambda & \star \\ \star & \star \end{bmatrix}$$

TABLE 1
The main notations in this paper.

$\mathbb{R}^{2d \times 2d}$	All the symplectic matrices involved are in the real space $\mathbb{R}^{2d \times 2d}$.
J	$\begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix}$.
SP	The collection of all the $2d$-by-$2d$ symplecic matrices.
SPN	The collection of all the $2d$-by-$2d$ symplecic matrices with nonsingular left upper block.
SPP	The collection of all the $2d$-by-$2d$ symmetric positive definite symplecic matrices.
SPM	The collection of all the $2d$-by-$2d$ symplectic M-matrices. The definition of M-matrices is shown in Definition 3.19.
SPS	The collection of all the $2d$-by-$2d$ singular symplecic matrices. The definition of singular symplecic matrices is shown in Definition 4.2.
$\mathcal{L}_n$	$\{[\begin{smallmatrix} I & 0/S_n \\ S_n/0 & I \end{smallmatrix}] \cdots [\begin{smallmatrix} I & 0 \\ S_2 & I \end{smallmatrix}][\begin{smallmatrix} I & S_1 \\ 0 & I \end{smallmatrix}] \mid S_i \in \mathbb{R}^{d \times d}, S_i^T = S_i \}$. The unit upper triangular symplectic matrices and the unit lower triangular symplectic matrices appear alternately.
$\mathcal{L}_n^2$	$\{ LL^T \mid L \in \mathcal{L}_n \}$. A subset of SPP.
$\mathcal{L}_1^T$	$\{ L^T \mid L \in \mathcal{L}_1 \}$. The collection of all the unit lower triangular symplectic matrices.
SG	$\mathcal{L}_1 \cup \mathcal{L}_1^T$. A set of generators of the group SP.
Pa	The map for extracting the lower triangular parameters for symmetric matrices. $Pa(S) = (s_{11}, s_{21}, s_{22}, s_{31}, \ldots, s_{dd})$, where $S = (s_{ij}) \in \mathbb{R}^{d \times d}$, $S^T = S$.

holds with a nonsingular P_λ for all $\lambda \neq 0$. Hence any symplectic matrix can be decomposed into a unit upper triangular symplectic matrix and a symplectic matrix with nonsingular left upper block. Furthermore, if needed, S can be set to

$$S = -P \begin{bmatrix} O_r & 0 \\ 0 & I_{d-r} \end{bmatrix} P^T$$

when the left upper block A_1 of H with rank r is decomposed as

$$A_1 = P \begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix} Q,$$

where $P, Q \in \mathbb{R}^{d \times d}$ are nonsingular.

Proof. Denote $H = [\begin{smallmatrix} A_1 & B_1 \\ A_2 & B_2 \end{smallmatrix}] \in SP$ where $A_i, B_i \in \mathbb{R}^{d \times d}$, and assume that $rank(A_1) = r$. There exist nonsingular P, Q such that $P^{-1} A_1 Q^{-1} = [\begin{smallmatrix} I_r & 0 \\ 0 & O_{d-r} \end{smallmatrix}]$, and let $S = -P[\begin{smallmatrix} O_r & 0 \\ 0 & I_{d-r} \end{smallmatrix}]P^T$, $\lambda \neq 0$. Then

$$\begin{bmatrix} I & -\lambda P^{-1} S P^{-T} \\ 0 & I \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ 0 & P^T \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & Q^T \end{bmatrix}$$

$$= \begin{bmatrix} I & \lambda \begin{bmatrix} O_r & 0 \\ 0 & I_{d-r} \end{bmatrix} \\ 0 & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix} & \star \\ \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} & \star \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} I_r & 0 \\ \lambda C_3 & \lambda C_4 \end{bmatrix} & \star \\ \star & \star \end{bmatrix}$$

$$= \begin{bmatrix} D_\lambda & \star \\ \star & \star \end{bmatrix},$$

where $C_1 \in \mathbb{R}^{r \times r}, C_2 \in \mathbb{R}^{r \times (d-r)}, C_3 \in \mathbb{R}^{(d-r) \times r}, C_4 \in \mathbb{R}^{(d-r) \times (d-r)}$ as well as $D_\lambda = \begin{bmatrix} I_r & 0 \\ \lambda C_3 & \lambda C_4 \end{bmatrix}$; next we explain why D_λ is nonsingular. Since all the matrices involved above are symplectic, we know that $\begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix}$ and $\begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}$ satisfy

$$\begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix}^T \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}^T \begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix}$$

due to Property 2.1, which implies that $C_2 = 0$. Furthermore,

$$\begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \\ C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \\ C_1 & 0 \\ C_3 & C_4 \end{bmatrix}$$

is full-rank as the subcolumns of a symplectic matrix; thus C_4 is also full-rank, i.e., nonsingular. Hence $D_\lambda = \begin{bmatrix} I_r & 0 \\ \lambda C_3 & \lambda C_4 \end{bmatrix}$ is nonsingular.

Now we have

$$\begin{aligned}
H = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} &= \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & \lambda P^{-1} S P^{-T} \\ 0 & I \end{bmatrix} \begin{bmatrix} D_\lambda & \star \\ \star & \star \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q^{-T} \end{bmatrix} \\
&= \begin{bmatrix} P & \lambda S P^{-T} \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} D_\lambda & \star \\ \star & \star \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q^{-T} \end{bmatrix} \\
&= \begin{bmatrix} I & \lambda S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} D_\lambda & \star \\ \star & \star \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q^{-T} \end{bmatrix} \\
&= \begin{bmatrix} I & \lambda S \\ 0 & I \end{bmatrix} \begin{bmatrix} P D_\lambda Q & \star \\ \star & \star \end{bmatrix} \\
&= \begin{bmatrix} I & \lambda S \\ 0 & I \end{bmatrix} \begin{bmatrix} P_\lambda & \star \\ \star & \star \end{bmatrix},
\end{aligned}$$

where $P_\lambda = P D_\lambda Q$ is nonsingular for all $\lambda \neq 0$. $\qquad \square$

Remark 3.3. In the most cases, we are unconcerned about the λ in Theorem 3.2 and only focus on the factorization

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & \star \\ \star & \star \end{bmatrix},$$

where S is symmetric and P is nonsingular. It in fact yields a symplectic matrix with nonsingular left upper block which is much more easier to deal with.

To prove the determinant of any symplectic matrix is one, researchers make efforts to have the symplectic matrix equipped with nonsingular left upper block; however, all of their transformations require the intractable blocks like symplectic interchanges and permutation matrices. Theorem 3.2 is the core of the proof of Theorem 3.1; it transforms a symplectic matrix into one having nonsingular left upper block by one unit triangular symplectic matrix.

LEMMA 3.4 (unit ULU factorization). *For any symplectic matrix $H \in \mathbb{R}^{2d \times 2d}$, there exist symmetric $S, T, U \in \mathbb{R}^{d \times d}$ and a nonsingular $P \in \mathbb{R}^{d \times d}$ such that*

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & \star \\ \star & \star \end{bmatrix} = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ T & I \end{bmatrix} \begin{bmatrix} I & U \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}.$$

Furthermore, T, U, P are uniquely determined by H and S.

Proof. Theorem 3.2 shows that there exists the factorization

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & \star \\ \star & \star \end{bmatrix},$$

where S is symmetric and P is nonsingular. Moreover, by (2.4),

$$\begin{bmatrix} P & \star \\ \star & \star \end{bmatrix} = \begin{bmatrix} P & A \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ BP^{-1} & I \end{bmatrix} \begin{bmatrix} I & AP^T \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix};$$

hence

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ T & I \end{bmatrix} \begin{bmatrix} I & U \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}, \qquad \square$$

where $T = BP^{-1}$ and $U = AP^T$ are symmetric. In addition, when H and S are fixed, we are able to obtain P, A, B, C, T, U in sequence; thus the factorization is determined by H, S. Note again that all the matrices involved above are symplectic.

Remark 3.5. Notice that

$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & PSP^T \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}.$$

The diagonal symplectic matrices like $\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}$ can be moved to any position between a series of many $\begin{bmatrix} I & \star \\ 0 & I \end{bmatrix}$ and $\begin{bmatrix} I & 0 \\ \star & I \end{bmatrix}$. For instance,

$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & 0 \\ \star & I \end{bmatrix} \begin{bmatrix} I & \star \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ \star & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & \star \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} I & 0 \\ \star & I \end{bmatrix} \begin{bmatrix} I & \star \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix};$$

therefore we can adjust the position of diagonal symplectic matrices among the unit triangular symplectic matrices by moving, even merge them into one.

What we just did is applying the LDU factorization to Theorem 3.2 in the proof of Lemma 3.4. According to Remark 3.5, Lemma 3.4 has some other forms, like

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ \star & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & \star \\ 0 & I \end{bmatrix}$$

and

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & 0 \\ \star & I \end{bmatrix} \begin{bmatrix} I & \star \\ 0 & I \end{bmatrix}.$$

The most important fact is the number of unit triangular matrices rather than the position of diagonal matrix. Note that all the forms involved share the same P due to Property 2.4.

Actually, Theorem 3.2 and Lemma 3.4 provide an algorithm for unit ULU factorization; see Algorithm 3.1.

So far we have proved that any symplectic matrix can be factored into three unit triangular symplectic matrices together with one diagonal symplectic matrix, where the number three is optimal, since the product of two only yields a symplectic matrix with nonsingular left upper block. To obtain the final unit triangular factorization, we have to consider about how to decompose the diagonal symplectic matrices next.

Algorithm 3.1. Unit ULU factorization.

Input: $H \in SP$

Output: S, T, U, P such that $H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ T & I \end{bmatrix}\begin{bmatrix} I & U \\ 0 & I \end{bmatrix}\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix}$

   Given $H := \begin{bmatrix} A & \star \\ \star & \star \end{bmatrix} \in SP$

   Compute the factorization $A = P\begin{bmatrix} I_r & 0 \\ 0 & O_{d-r} \end{bmatrix}Q$, where P, Q are nonsingular

   Set $S := -P\begin{bmatrix} O_r & 0 \\ 0 & I_{d-r} \end{bmatrix}P^T$

   Set $\begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} := \begin{bmatrix} I & -S \\ 0 & I \end{bmatrix}H$

   Set $T := A_2 A_1^{-1}$

   Set $U := B_1 A_1^T$

   Set $P := A_1$

   return S, T, U, P

LEMMA 3.6. *For any symplectic matrix* $H = \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \in \mathbb{R}^{2d \times 2d}$, *where* $P \in \mathbb{R}^{d \times d}$ *is nonsingular, there exist symmetric* $S_1, S_2, \ldots, S_7 \in \mathbb{R}^{d \times d}$ *as well as symmetric* $T_1, T_2, \ldots, T_7 \in \mathbb{R}^{d \times d}$ *such that*

$$H = \begin{bmatrix} I & S_7 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ S_6 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix}\begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} I & 0 \\ T_7 & I \end{bmatrix}\begin{bmatrix} I & T_6 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & T_2 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ T_1 & I \end{bmatrix}.$$

Proof. At first, there exist two symmetric matrices $P_1, P_2 \in \mathbb{R}^{d \times d}$ such that $P = P_1 \cdot P_2$; then we have

$$H = \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} = \begin{bmatrix} P_1 & 0 \\ 0 & P_1^{-T} \end{bmatrix}\begin{bmatrix} P_2 & 0 \\ 0 & P_2^{-T} \end{bmatrix}.$$

Moreover,

$$\begin{bmatrix} P_1 & 0 \\ 0 & P_1^{-T} \end{bmatrix} = \begin{bmatrix} I & -P_1 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ P_1^{-1} - I & I \end{bmatrix}\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ P_1 - I & I \end{bmatrix},$$

$$\begin{bmatrix} P_2 & 0 \\ 0 & P_2^{-T} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -P_2^{-1} & I \end{bmatrix}\begin{bmatrix} I & P_2 - I \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ I & I \end{bmatrix}\begin{bmatrix} I & P_2^{-1} - I \\ 0 & I \end{bmatrix},$$

so that

$$H = \begin{bmatrix} I & -P_1 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ P_1^{-1} - I & I \end{bmatrix}\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ P_1 - I - P_2^{-1} & I \end{bmatrix}$$
$$\cdot \begin{bmatrix} I & P_2 - I \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ I & I \end{bmatrix}\begin{bmatrix} I & P_2^{-1} - I \\ 0 & I \end{bmatrix}.$$

Up to now, we have obtained the expressions of S_i exactly. For T_i, we just need to express H^T in the above way and return to H by transposition operator again. □

Remark 3.7. As shown in the proof of Lemma 3.6, some of the blocks can be required to be I_d for reducing the degree of freedom of factorization, such as S_2, S_5.

Now we are able to provide the proof of Theorem 3.1.

Proof. Let $H \in SP$. According to Lemma 3.4, we have

$$H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ T & I \end{bmatrix} \begin{bmatrix} I & U \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix},$$

where S, T, U are symmetric and P is nonsingular. Lemma 3.6 shows that

$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} = \begin{bmatrix} I & S_7 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_6 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix},$$

so that

$$\begin{aligned} H &= \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ T & I \end{bmatrix} \begin{bmatrix} I & U \\ 0 & I \end{bmatrix} \begin{bmatrix} I & S_7 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_6 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ T & I \end{bmatrix} \begin{bmatrix} I & U + S_7 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_6 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix}; \end{aligned}$$

thus H has been decomposed into 9 unit triangular symplectic matrices. □

Remark 3.8. This factorization starts with an upper triangular matrix in the right-hand side. In fact, if we decompose the H^T into the above 9 factors, subsequently we obtain an expression of H starting with a lower triangular matrix by transposition.

Theorem 3.1 shows an elegant expression of symplectic matrix, and it implies several important properties.

COROLLARY 3.9. *The determinant of any symplectic matrix is one.*

Proof. The determinant of any unit triangular symplectic matrix is one. □

There are many proofs for the determinant of symplectic matrix [8, 27]. Compared to the conventional proofs for the determinant of symplectic matrix, such as QR-like factorization, SVD-like factorization, and transvections factorization, Theorem 3.1 is the most symplectic-determinant-revealing and elementary.

COROLLARY 3.10. *The matrix symplectic group SP is path connected.*

Proof. Given two symplectic matrices $H, G \in SP$, we can express them as

$$\begin{aligned} H &= \begin{bmatrix} I & S_9 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix}, \\ G &= \begin{bmatrix} I & T_9 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ T_2 & I \end{bmatrix} \begin{bmatrix} I & T_1 \\ 0 & I \end{bmatrix}, \end{aligned}$$

where S_i, T_i are symmetric, due to Theorem 3.1. Define the continuous $\gamma : [0, 1] \to SP$ as

$$\gamma(t) = \begin{bmatrix} I & (1-t) \cdot S_9 + t \cdot T_9 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ (1-t) \cdot S_2 + t \cdot T_2 & I \end{bmatrix} \begin{bmatrix} I & (1-t) \cdot S_1 + t \cdot T_1 \\ 0 & I \end{bmatrix};$$

then we have $\gamma(0) = H$ and $\gamma(1) = G$; thus SP is path connected. □

COROLLARY 3.11. *Denote $SG = \mathcal{L}_1 \cup \mathcal{L}_1^T$ where $\mathcal{L}_1^T := \{H^T | H \in \mathcal{L}_1\}$; then SG is a set of generators of the group SP.*

Proof. It is obvious by Theorem 3.1. □

Remark 3.12. Although Corollary 3.9, 3.10 are well-known and can be found in many literatures such as [10, 14], here we provide the alternative and concise proofs which are deduced directly from the unit triangular factorization. It is noteworthy that Corollary 3.11 is indeed a new feature, while the existing theory of generators

of SP requires the factors of diagonal symplectic matrices (1.2). For example, [10, p. 49, Corollary 2.40] states that $\{J\} \cup \{$*unit lower triangular symplectic matrices*$\} \cup$ $\{$*diagonal symplectic matrices*$\}$ is a set of generators of SP. In fact, Corollary 3.11 indicates that $\{$*diagonal symplectic matrices*$\}$ can be removed from the above statement.

Up to now, we have studied several properties of the general symplectic matrices. Next we focus on the structured sets of symplectic matrices.

3.1. Symplectic matrices with nonsingular left upper block. For symplectic matrices with other structures, which are naturally subsets of SP, we can exploit their special structures for reducing 9 to a smaller number. Here we turn to the desirable symplectic matrices which have nonsingular left upper blocks. Denote

$$SPN = \left\{ H = \begin{bmatrix} A & \star \\ \star & \star \end{bmatrix} \in SP \middle| A \in \mathbb{R}^{d \times d} \ nonsingular \right\}.$$

We have proved that $SP = \mathcal{L}_9$; furthermore, it is easy to show that $SPN \subsetneq \mathcal{L}_8$.

THEOREM 3.13. $SPN \subsetneq \mathcal{L}_8$. *Thus any symplectic matrix with nonsingular left upper block can be factored into no more than 8 unit triangular symplectic matrices.*

Proof. With the LDU factorization and Lemma 3.6, we know

$$H = \begin{bmatrix} A & \star \\ \star & \star \end{bmatrix} = \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & A^{-T} \end{bmatrix} \begin{bmatrix} I & T \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} I & S_7 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 + T \\ 0 & I \end{bmatrix},$$

so that $SPN \subset \mathcal{L}_8$. Moreover, consider $J = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$; we have $J \in \mathcal{L}_3 \subset \mathcal{L}_8$ and $J \notin SPN$; hence $SPN \subsetneq \mathcal{L}_8$. $\square$

Until now, we have $SPN \subsetneq \mathcal{L}_8 \subset \mathcal{L}_9 = SP$; however, it is not sure whether $\mathcal{L}_8 = \mathcal{L}_9$ or $\mathcal{L}_8 \subsetneq \mathcal{L}_9$. Dopico and Johnson [12] provides a proof that SPN is a dense subset in the group of symplectic matrices, although it can be accomplished through general properties of algebraic manifolds. Here we present a simple proof again according to Theorem 3.2, and this result immediately shows that $\mathcal{L}_8$ is at least dense in $\mathcal{L}_9 = SP$.

COROLLARY 3.14. *SPN is dense in SP, i.e., $\overline{SPN} = SP$.*

Proof. Theorem 3.2 shows that, for any $H \in SP$, there exists a symmetric S such that

$$H = \begin{bmatrix} I & \lambda S \\ 0 & I \end{bmatrix} \begin{bmatrix} P_\lambda & \star \\ \star & \star \end{bmatrix}$$

holds with a nonsingular P_λ for all $\lambda \neq 0$. Therefore

$$H = \lim_{\lambda \to 0} \begin{bmatrix} I & -\lambda S \\ 0 & I \end{bmatrix} H = \lim_{\lambda \to 0} \begin{bmatrix} P_\lambda & \star \\ \star & \star \end{bmatrix} \in \overline{SPN},$$

which means $SP \subset \overline{SPN}$; consequently $SP = \overline{SPN}$. $\square$

COROLLARY 3.15. *$\mathcal{L}_8$ is dense in $\mathcal{L}_9$, i.e., $\overline{\mathcal{L}_8} = \mathcal{L}_9$.*

Proof. SPN is dense in SP. $\square$

3.2. Positive definite symplectic matrices. Denote the collection of symmetric positive definite symplectic matrices by

$$SPP = \{H \in SP | H \ symmetric \ positive \ definite\}.$$

THEOREM 3.16. $SPP \subsetneq \mathcal{L}_4$. *Thus any symmetric positive definite symplectic matrix can be factored into no more than 4 unit triangular symplectic matrices.*

Proof. If H is symmetric positive definite, then all principal submatrices of H are symmetric positive definite. Denote that $H = [\begin{smallmatrix} H_1 & H_2 \\ H_3 & H_4 \end{smallmatrix}] \in SPP$; here H_1 is positive definite, thus nonsingular. So that in the process of decomposition, instead of going through Theorem 3.2, we just apply the LDU factorization to H and get

$$H = \begin{bmatrix} I & 0 \\ H_3 H_1^{-1} & I \end{bmatrix} \begin{bmatrix} H_1 & 0 \\ 0 & H_1^{-T} \end{bmatrix} \begin{bmatrix} I & H_1^{-1} H_2 \\ 0 & I \end{bmatrix},$$

where $H_3 H_1^{-1}$ and $H_1^{-1} H_2$ are symmetric matrices. Moreover, because of the symmetry of H, $H_3 H_1^{-1} = (H_1^{-1} H_2)^T = H_1^{-1} H_2$. Therefore, every positive definite symplectic matrix can be written as

$$(3.1) \qquad H = \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & S \\ 0 & I \end{bmatrix}$$

with P symmetric positive definite. Since P is already symmetric here, we do not have to decompose it into the product of two symmetric matrices and just have the factorization

$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -P^{-1} & I \end{bmatrix} \begin{bmatrix} I & P - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \begin{bmatrix} I & P^{-1} - I \\ 0 & I \end{bmatrix};$$

as a consequence

$$\begin{aligned} H &= \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -P^{-1} & I \end{bmatrix} \begin{bmatrix} I & P - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \begin{bmatrix} I & P^{-1} - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ S - P^{-1} & I \end{bmatrix} \begin{bmatrix} I & P - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \begin{bmatrix} I & S + P^{-1} - I \\ 0 & I \end{bmatrix}; \end{aligned}$$

thus $H \in \mathcal{L}_4$.

On the other hand, $J = [\begin{smallmatrix} I & I \\ 0 & I \end{smallmatrix}][\begin{smallmatrix} I & 0 \\ -I & I \end{smallmatrix}][\begin{smallmatrix} I & I \\ 0 & I \end{smallmatrix}] \in \mathcal{L}_3 \subset \mathcal{L}_4$ while $J \notin SPP$; hence $SPP \subsetneq \mathcal{L}_4$. □

Remark 3.17. The above factorization starts with an upper triangular matrix in the right-hand side. Similar to the proof of Theorem 3.16 , we can use the LDU factorization in the case of H_4 invertible,

$$(3.2) \qquad H = \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & 0 \\ S & I \end{bmatrix},$$

to obtain another decomposition where the first one is a lower triangular matrix.

Going a step further, we can prove that $\mathcal{L}_4$ is optimal, i.e., $SPP \nsubseteq \mathcal{L}_3$.

THEOREM 3.18. $SPP \nsubseteq \mathcal{L}_3$.

Proof. Let $G = [\begin{smallmatrix} 2I & 0 \\ 0 & \frac{1}{2}I \end{smallmatrix}] \in SPP$. If there exist S_1, S_2, S_3 such that

$$G = \begin{bmatrix} I & S_3 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix} = \begin{bmatrix} I + S_3 S_2 & S_1 + S_3 S_2 S_1 + S_3 \\ S_2 & I + S_2 S_1 \end{bmatrix},$$

then we have $S_2 = 0$, $I = 2I$, which is a contradiction. □

3.3. Symplectic M-matrices. M-matrices occur very often in a wide variety of areas including finite difference methods for partial differential equations, economics, probability, and statistics [6]. There are many equivalent definitions of M-matrices; we adopt the following one.

DEFINITION 3.19. $A \in \mathbb{R}^{n \times n}$ *is an M-matrix if* $a_{ij} \leq 0$ *for* $i \neq j$ *and* $Re(\lambda) > 0$ *for every eigenvalue* λ *of* A.

The next result is a factorization of the symplectic M-matrices. It appears in [12] and is the key result on how symplectic M-matrices are decomposed into the product of several unit triangular symplectic matrices.

THEOREM 3.20. *The set of* $2d$-*by*-$2d$ *symplectic M-matrices is*

$$SPM = \left\{ \begin{bmatrix} I & 0 \\ H & I \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} I & K \\ 0 & I \end{bmatrix} \middle| \begin{array}{l} D \in \mathbb{R}^{d \times d} \ positive \ diagonal \\ H = H^T \leq 0 \\ K = K^T \leq 0 \\ HDK \ diagonal \end{array} \right\},$$

where the inequalities $H \leq 0$ *and* $K \leq 0$ *mean that* $h_{ij} \leq 0$ *and* $k_{ij} \leq 0$ *for all* i, j.

Here we treat D as an ordinary symmetric matrix; as in Theorem 3.16, any symplectic M-matrix can be decomposed as the product of four triangular blocks.

THEOREM 3.21. $SPM \subsetneq \mathcal{L}_4$. *Thus any symplectic M-matrix can be factored into no more than* 4 *unit triangular symplectic matrices.*

Proof. It is easy to show $SPM \subset \mathcal{L}_4$ based on Theorem 3.20. Moreover, $-I_{2d} \in \mathcal{L}_4$ while $-I_{2d} \notin SPM$; hence $SPM \subsetneq \mathcal{L}_4$. $\square$

THEOREM 3.22. $SPM \nsubseteq \mathcal{L}_3$.

Proof. It is similar to the proof of Theorem 3.18. $\square$

4. Unconstrained parametrization and optimization. In this section, we first consider the unconstrained parametrization of the symplectic matrices or its subsets. For a set Ω, we expect to find a map $\phi : \mathbb{R}^n \to \Omega$ which is smooth and surjective. Therefore one may be able to deal with some problems on unconstrained parameter space $\mathbb{R}^n$ instead of Ω by employing ϕ. In the last subsection, we turn to the topic about how the unconstrained parametrization impacts on the optimization problems with symplectic constraints. In order to represent the free parameters of a symmetric matrix, we define the map Pa for extracting the lower triangular parameters as $Pa(S) = (s_{11}, s_{21}, s_{22}, s_{31}, \ldots, s_{dd})$, where $S = (s_{ij}) \in \mathbb{R}^{d \times d}$, $S^T = S$.

4.1. Parametrization of general symplecic matrices. Based on Theorem 3.1, one can parametrize SP efficiently. Take symmetric $S_1, S_2, \ldots, S_9 \in \mathbb{R}^{d \times d}$; then

$$(4.1) \qquad H(Pa(S_1), \ldots, Pa(S_9)) = \begin{bmatrix} I & S_9 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix}$$

can represent any symplectic matrix when $Pa(S_1), \ldots, Pa(S_9)$ vary. Note that here $Pa(S_i)$ are free parameters without any constraint; thus H is a surjective map from $\mathbb{R}^{9d(d+1)/2}$ into SP. With (4.1), it is easy to write out the inverse element and the transpose of H, i.e.,

$$(4.2) \qquad H(Pa(S_1), \ldots, Pa(S_9))^{-1} = \begin{bmatrix} I & -S_1 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} I & 0 \\ -S_8 & I \end{bmatrix} \begin{bmatrix} I & -S_9 \\ 0 & I \end{bmatrix},$$

$$(4.3) \qquad H(Pa(S_1), \ldots, Pa(S_9))^T = \begin{bmatrix} I & 0 \\ S_1 & I \end{bmatrix} \cdots \begin{bmatrix} I & S_8 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_9 & I \end{bmatrix}.$$

The concise expressions of the inverse element and the transpose make it possible to construct more complex structured symplectic matrices.

4.2. Parametrization of positive definite symplectic matrices. By Theorem 3.16, we know that SPP is a subset of $\mathcal{L}_4$. Since $\mathcal{L}_n$ cannot guarantee the symmetry for its elements, a new collection of symplectic matrices is needed. The following theorem will show how to construct the unconstrained parametrization of SPP. Denote

$$\mathcal{L}_n^2 = \{LL^T | L \in \mathcal{L}_n\}.$$

By (4.2) and (4.3), we know L is invertible and L^T is still symplectic, so that LL^T is a positive definite symplectic matrix. Thus $\mathcal{L}_m^2 \subset \mathcal{L}_n^2 \subset SPP$ for all integers $1 \leq m \leq n$.

THEOREM 4.1. $SPP = \mathcal{L}_4^2$. *Thus any positive definite symplectic matrix can be represented as the product of a matrix which can be factored into no more than 4 unit triangular symplectic matrices and its transpose.*

Proof. According to (3.1) in the proof of Theorem 3.16, every positive definite symplectic matrix can be written as

$$H = \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I & S \\ 0 & I \end{bmatrix}$$

with P symmetric positive definite. Let $P^{1/2}$ be the unique positive definite square root of P. $P = P^{1/2}P^{1/2}$ implies that $P^{-1} = (P^{1/2})^{-1}(P^{1/2})^{-1}$; then

$$H = \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P^{1/2} & 0 \\ 0 & (P^{1/2})^{-1} \end{bmatrix} \begin{bmatrix} P^{1/2} & 0 \\ 0 & (P^{1/2})^{-1} \end{bmatrix} \begin{bmatrix} I & S \\ 0 & I \end{bmatrix}$$

$$= \left(\begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P^{1/2} & 0 \\ 0 & (P^{1/2})^{-1} \end{bmatrix} \right) \left(\begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P^{1/2} & 0 \\ 0 & (P^{1/2})^{-1} \end{bmatrix} \right)^T.$$

Just like what we did in Theorem 3.16, taking into account Lemma 3.6 and the symmetry of $P^{1/2}$, we have

$$\begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} P^{1/2} & 0 \\ 0 & (P^{1/2})^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -(P^{1/2})^{-1} & I \end{bmatrix} \begin{bmatrix} I & P^{1/2} - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \begin{bmatrix} I & (P^{1/2})^{-1} - I \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ S - (P^{1/2})^{-1} & I \end{bmatrix} \begin{bmatrix} I & P^{1/2} - I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \begin{bmatrix} I & (P^{1/2})^{-1} - I \\ 0 & I \end{bmatrix};$$

hence we have obtained the expression of L exactly. $\square$

Based on Theorem 4.1, one can represent SPP by parameterizing $\mathcal{L}_4$. Take symmetric $S_1, S_2, S_3, S_4 \in \mathbb{R}^{d \times d}$; then denote

$$H(Pa(S_1), \ldots, Pa(S_4)) = LL^T,$$

$$L = \begin{bmatrix} I & 0 \\ S_4 & I \end{bmatrix} \begin{bmatrix} I & S_3 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S_2 & I \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix}.$$

Here $Pa(S_i)$ are free parameters without any constraint; thus H is a surjective map from $\mathbb{R}^{2d(d+1)}$ into SPP.

4.3. Parametrization of singular symplectic matrices. There have been some studies on the structure of singular symplectic matrices [25, 26].

DEFINITION 4.2. *A symplectic matrix $H \in SP$ is singular if $det(H - I) = 0$.*

Denote all the singular symplectic matrices by

$$SPS = \{H \in SP | H \ singular\}.$$

A symplectic matrix H is singular when 1 is one of its eigenvalues, and there is a $0 \neq u \in \mathbb{R}^{2d}$ such that $Hu = u$ where u is called the fixed vector of H. Theorem 3.1 prompts us to consider the following question: If there are several linearly independent vectors $u_1, u_2, \ldots, u_n \in \mathbb{R}^{2d}, n \leq 2d$, how do we search for a symplectic matrix H such that $Hu_i = u_i$?

Let us consider a slightly simpler case. Denote

$$U = [u_1, u_2, \ldots, u_n] = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \in \mathbb{R}^{2d \times n}, \quad V_1, V_2 \in \mathbb{R}^{d \times n},$$

where U is a full-rank upper triangular matrix, and assume that $rank(V_1) = r_1$, $rank(V_2) = r_2$. Now let

$$\mathcal{V}_1 = \left\{ \begin{bmatrix} I & 0 \\ S & I \end{bmatrix} \middle| S = \begin{bmatrix} O_{r_1} & 0 \\ 0 & T \end{bmatrix}, \quad T \in \mathbb{R}^{(d-r_1) \times (d-r_1)} \ symmetric \right\},$$

$$\mathcal{V}_2 = \left\{ \begin{bmatrix} I & S \\ 0 & I \end{bmatrix} \middle| S = \begin{bmatrix} O_{r_2} & 0 \\ 0 & T \end{bmatrix}, \quad T \in \mathbb{R}^{(d-r_2) \times (d-r_2)} \ symmetric \right\},$$

$$\mathcal{V}_U = \{H_k \cdots H_2 H_1 | H_i \in \mathcal{V}_1 \cup \mathcal{V}_2, k \geq 1\}.$$

Note that

$$\begin{cases} r_1 = n, r_2 = 0 & when \ 1 \leq n < d, \\ r_1 = d, r_2 = n - d & when \ d \leq n \leq 2d. \end{cases}$$

Given any $H \in \mathcal{V}_U$, it is easy to verify that $HU = U$. For instance, if $H \in \mathcal{V}_1$, then

$$HU = \begin{bmatrix} I & 0 \\ \begin{bmatrix} O_{r_1} & 0 \\ 0 & T \end{bmatrix} & I \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ \begin{bmatrix} O_{r_1} & 0 \\ 0 & T \end{bmatrix} V_1 + V_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = U.$$

Note that the last $d - r_1$ rows of V_1 are all 0.

Denote the set of all the symplectic matrices H satisfying $HU = U$ by $\mathcal{W}_U$. It is easy to see that $\mathcal{W}_U$ and $\mathcal{V}_U$ are both groups; moreover, $\mathcal{V}_U$ is a subgroup of $\mathcal{W}_U$. Whether $\mathcal{V}_U = \mathcal{W}_U$ or $\mathcal{V}_U \subsetneq \mathcal{W}_U$ is an interesting problem to be concerned with. At least, in the case of $n = 0$, Theorem 3.1 immediately points out that $\mathcal{V}_U = \mathcal{W}_U$ since both of $\mathcal{V}_U$ and $\mathcal{W}_U$ equal to SP, while in the case of $n = 2d$, $\mathcal{V}_U$ and $\mathcal{W}_U$ degenerate both into $\{I_{2d}\}$. A theorem is proposed as follows.

THEOREM 4.3. $\mathcal{V}_U = \mathcal{W}_U$. *Here $U \in \mathbb{R}^{2d \times n}$ $(1 \leq n \leq 2d)$ is any full-rank upper triangular matrix. Furthermore,*

$$\begin{cases} \mathcal{V}_U = \{H_{10} \cdots H_2 H_1 | H_{2i-1} \in \mathcal{V}_2, H_{2i} \in \mathcal{V}_1\} & when \ 1 \leq n < d, \\ \mathcal{V}_U = \{H | H \in \mathcal{V}_2\} & when \ d \leq n \leq 2d. \end{cases}$$

Proof. 1. Case $d \leq n \leq 2d$. Denote

$$U = \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ 0 & \begin{bmatrix} \widetilde{U}_{22} \\ 0 \end{bmatrix} \end{bmatrix},$$

where $U_{11} \in \mathbb{R}^{d \times d}, U_{12} \in \mathbb{R}^{d \times (n-d)}, U_{22} \in \mathbb{R}^{d \times (n-d)}, \widetilde{U}_{22} \in \mathbb{R}^{(n-d) \times (n-d)}$. Since U is full-rank upper triangular, U_{11} and $\widetilde{U}_{22}$ are nonsingular upper triangular matrices. For any $H = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \in \mathcal{W}_U$,

$$HU = U \Rightarrow \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} A_1 U_{11} & A_1 U_{12} + B_1 U_{22} \\ A_2 U_{11} & A_2 U_{12} + B_2 U_{22} \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} \Rightarrow A_1 = I, \ A_2 = 0.$$

With $A_1 = I, A_2 = 0$ and the third item of Property 2.1, we find that $H = \begin{bmatrix} I & B_1 \\ 0 & I \end{bmatrix}$. Moreover,

$$A_1 U_{12} + B_1 U_{22} = U_{12} \Rightarrow B_1 U_{22} = 0 \Rightarrow \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \widetilde{U}_{22} \\ 0 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} B_{11} \widetilde{U}_{22} \\ B_{21} \widetilde{U}_{22} \end{bmatrix} = 0 \Rightarrow B_{11} = 0, \ B_{21} = 0,$$

and the symmetry of B_1 shows that $B_{12} = B_{21} = 0$. Hence

$$H = \begin{bmatrix} I & \begin{bmatrix} O_{n-d} & 0 \\ 0 & B_{22} \end{bmatrix} \\ 0 & I \end{bmatrix} \in \mathcal{V}_2 \subset \mathcal{V}_U,$$

which means $\mathcal{W}_U \subset \mathcal{V}_U$; consequently $\mathcal{V}_U = \mathcal{W}_U$.

2. Case $1 \leq n < d$. Denote

$$U = \begin{bmatrix} U_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ 0 \end{bmatrix}, \quad U_1 \in \mathbb{R}^{d \times n}, \widetilde{U}_1 \in \mathbb{R}^{n \times n},$$

where $\widetilde{U}_1$ is nonsingular upper triangular. For any $H = \begin{bmatrix} A & \star \\ B & \star \end{bmatrix} \in \mathcal{W}_U$,

$$HU = U \Rightarrow \begin{bmatrix} A & \star \\ B & \star \end{bmatrix} \begin{bmatrix} U_1 \\ 0 \end{bmatrix} = \begin{bmatrix} U_1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} & \star \\ \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} & \star \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \cdot \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \begin{bmatrix} A_{11} \widetilde{U}_1 \\ A_{21} \widetilde{U}_1 \end{bmatrix} \\ \begin{bmatrix} B_{11} \widetilde{U}_1 \\ B_{21} \widetilde{U}_1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \widetilde{U}_1 \\ 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

$$\Rightarrow A_{11} = I_n, \ A_{21} = 0, \ B_{11} = O_n, \ B_{21} = 0.$$

Property 2.1 leads to

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}^T \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} I_n & A_{12} \\ 0 & A_{22} \end{bmatrix}^T \begin{bmatrix} O_n & B_{12} \\ 0 & B_{22} \end{bmatrix} = \begin{bmatrix} O_n & B_{12} \\ 0 & B_{22} \end{bmatrix}^T \begin{bmatrix} I_n & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} O_n & B_{12} \\ 0 & A_{12}^T B_{12} + A_{22}^T B_{22} \end{bmatrix} = \begin{bmatrix} O_n & 0 \\ B_{12}^T & B_{12}^T A_{12} + B_{22}^T A_{22} \end{bmatrix}$$

$$\Rightarrow B_{12} = 0, \ A_{22}^T B_{22} = B_{22}^T A_{22}.$$

Now we have

$$H = \begin{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} & \star \\ \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} & \star \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} I_n & A_{12} \\ 0 & A_{22} \end{bmatrix} & \star \\ \begin{bmatrix} O_n & 0 \\ 0 & B_{22} \end{bmatrix} & \star \end{bmatrix}, \quad A_{22}^T B_{22} = B_{22}^T A_{22}.$$

Since $rank(\begin{bmatrix} A \\ B \end{bmatrix}) = d$, we know $rank(\begin{bmatrix} A_{22} \\ B_{22} \end{bmatrix}) = d - rank(\begin{bmatrix} I_n & A_{12} \end{bmatrix}) = d - n$, thus $G = \begin{bmatrix} A_{22} \\ B_{22} \end{bmatrix}$ is full-rank. Property 2.6 tells us that G is a part of a $2(d-n)$-by-$2(d-n)$ symplectic matrix $\widetilde{G} = \begin{bmatrix} A_{22} & \star \\ B_{22} & \star \end{bmatrix}$. Similar to the proof of Theorem 3.2, there exist a symmetric S and nonsingular P, Q such that

$$\begin{bmatrix} P & 0 \\ 0 & P^{-T} \end{bmatrix} \begin{bmatrix} I_{d-n} & 0 \\ S & I_{d-n} \end{bmatrix} \widetilde{G} \begin{bmatrix} Q & 0 \\ 0 & Q^{-T} \end{bmatrix} = \begin{bmatrix} G_{11} & \star \\ I_{d-n} & \star \end{bmatrix},$$

where G_{11} is symmetric. We have

$$\begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & -I_{d-n} \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} I & \begin{bmatrix} O_n & -A_{12}Q \\ -Q^T A_{12}^T & I_{d-n} - G_{11} \end{bmatrix} \\ 0 & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & P \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} I_n & 0 \\ 0 & P^{-T} \end{bmatrix} \end{bmatrix}$$

$$\cdot \begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & S \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & A_{12} \\ 0 & A_{22} \end{bmatrix} & \star \\ \begin{bmatrix} O_n & 0 \\ 0 & B_{22} \end{bmatrix} & \star \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & Q \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} I_n & 0 \\ 0 & Q^{-T} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & -I_{d-n} \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} I & \begin{bmatrix} O_n & -A_{12}Q \\ -Q^T A_{12}^T & I_{d-n} - G_{11} \end{bmatrix} \\ 0 & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & A_{12}Q \\ 0 & G_{11} \end{bmatrix} & \star \\ \begin{bmatrix} O_n & 0 \\ 0 & I_{d-n} \end{bmatrix} & \star \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & -I_{d-n} \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & I_{d-n} \end{bmatrix} & \star \\ \begin{bmatrix} O_n & 0 \\ 0 & I_{d-n} \end{bmatrix} & \star \end{bmatrix} = \begin{bmatrix} I & \star \\ 0 & \star \end{bmatrix} = \begin{bmatrix} I & T \\ 0 & I \end{bmatrix},$$

where T is symmetric; hence

$$H = \begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & -S \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & P^{-1} \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} I_n & 0 \\ 0 & P^T \end{bmatrix} \end{bmatrix} \begin{bmatrix} I & \begin{bmatrix} O_n & A_{12}Q \\ Q^T A_{12}^T & G_{11} - I_{d-n} \end{bmatrix} \\ 0 & I \end{bmatrix}$$

$$\cdot \begin{bmatrix} \begin{bmatrix} I & 0 \\ O_n & 0 \\ 0 & I_{d-n} \end{bmatrix} & 0 \\ & I \end{bmatrix} \begin{bmatrix} I & T \\ 0 & I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & Q^{-1} \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} I_n & 0 \\ 0 & Q^T \end{bmatrix} \end{bmatrix}.$$

Taking into account Remark 3.5 and Lemma 3.6, we can easily factor H into 10 unit triangular symplectic matrices from $\mathcal{V}_1$ and $\mathcal{V}_2$. Therefore $\mathcal{W}_U \subset \mathcal{V}_U$; consequently $\mathcal{W}_U = \mathcal{V}_U$. □

Now return to the topic about how to parameterize singular symplectic matrices. Theorem 4.3 gives a specific structure of singular symplectic matrices with many linearly independent vectors as their eigenvectors, especially when $U = e_1 = [1, 0, \ldots, 0]^T$, we obtain $\mathcal{W}_{e_1} = \mathcal{V}_{e_1}$ whose structure is clear to be freely parameterized.

COROLLARY 4.4. *The set of* $2d$-*by*-$2d$ *singular symplectic matrices is*

$$
SPS = \left\{ Q \begin{bmatrix} \begin{matrix} I & \\ \begin{bmatrix} 0 & 0 \\ 0 & S_{10} \end{bmatrix} \end{matrix} & \begin{matrix} 0 \\ \\ I \end{matrix} \end{bmatrix} \begin{bmatrix} I & S_9 \\ 0 & I \end{bmatrix} \cdots \begin{bmatrix} \begin{matrix} I & \\ \begin{bmatrix} 0 & 0 \\ 0 & S_2 \end{bmatrix} \end{matrix} & \begin{matrix} 0 \\ \\ I \end{matrix} \end{bmatrix} \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix} Q^{-1} \right.
$$

$$
\left. \left| \; S_{2i-1} \in \mathbb{R}^{d \times d} \; symmetric, \; S_{2i} \in \mathbb{R}^{(d-1) \times (d-1)} \; symmetric, \; Q \; symplectic \right. \right\}.
$$

One can express Q *as the product of* 9 *unit triangular symplectic matrices if needed.*

Proof. For any

$$
H = Q \begin{bmatrix} \begin{matrix} I & \\ \begin{bmatrix} 0 & 0 \\ 0 & S_{10} \end{bmatrix} \end{matrix} & \begin{matrix} 0 \\ \\ I \end{matrix} \end{bmatrix} \cdots \begin{bmatrix} I & S_1 \\ 0 & I \end{bmatrix} Q^{-1} \in SPS,
$$

let $u = Qe_1 \neq 0$. Then it is easy to check that $Hu = u$; hence H is singular.

For any singular $H \in SP$, assume that $Hu = u$, where $0 \neq u \in \mathbb{R}^{2d}$. From Property 2.5 and Property 2.6, we know that u is a part of a symplectic matrix $Q = [u \; \star]$. Therefore $Q^{-1}HQe_1 = Q^{-1}Hu = Q^{-1}u = e_1$, which means that $Q^{-1}HQ \in \mathcal{W}_{e_1} = \mathcal{V}_{e_1}$ due to Theorem 4.3. We immediately know that $H \in SPS$. □

4.4. Unconstrained optimization. An optimization problem with symplectic constraint is in the following form

$$
(4.4) \qquad \min_{X \in \mathbb{R}^{2d \times 2d}} f(X), \quad \text{such that (s.t.)} \; X^T J X = J.
$$

Because of the antisymmetry of $X^T J X$, the constraint of (4.4) is made up of $2d^2 - d$ equations. There have been many works on optimization on the real symplectic group [7, 15, 34], in which one performs optimization by considering the gradients along the manifold. In this work, the unit triangular factorization provides an approach to the symplectic optimization from a new perspective, i.e., optimizing in a higher dimensional unconstrained parameter space. By (4.1), (4.4) is equivalent to

$$
\min_{Pa(S_i) \in \mathbb{R}^{\frac{d(d+1)}{2}}} f(H(Pa(S_1), \ldots, Pa(S_9))),
$$

which is indeed an unconstrained optimization problem. In recent years, with the expansion of variables required by practical application, the limitations of existing constrained optimization methods are reflected in practice. Fortunately, the unconstrained parametrization of the symplectic matrices allows us to circumvent this limitation and apply faster and more efficient unconstrained optimization algorithms, such as the recently popular deep learning techniques, which are able to deal with the optimization problems with billions of unconstrained parameters. In fact, this

method has been utilized in our recent work [23], where we construct symplectic neural networks based on the factorization and achieve a great success. In such a case, the unit triangular factorization-based optimization can be implemented directly within the deep learning framework and performs well, while the traditional Riemannian-steepest-descent approach faces challenges.

5. Conclusions. The factorization theorems shown in this work overcome some defects of conventional factorizations. The works [9, 12, 21, 31, 36] provide several factorizations of the matrix symplectic group. However, all the factorizations require cells of symplectic interchanges, permutation matrices or symplectic-orthogonal matrices, which are not elementary enough hence hard to be freely parameterized. These works [16, 27] factor the $2d$-by-$2d$ symplectic matrices as the products of at most $4d$ symplectic transvections, which can freely parameterize the matrix symplectic group, but may be unstable with a large d in practice.

Theorem 3.1 proves that any symplectic matrix can be factored into no more than 9 unit triangular symplectic matrices. The core of the proof is Theorem 3.2, which writes a symplectic matrix as the product of a symplectic matrix with nonsingular left upper block and a unit triangular symplectic matrix. Subsequently, we derive some important corollaries by Theorem 3.1 and Theorem 3.2, such as (i) the determinant of symplectic matrix is one, (ii) the matrix symplectic group is path connected, (iii) all the unit triangular symplectic matrices form a set of generators of the matrix symplectic group. Furthermore, this factorization yields effective methods for unconstrained parametrization of the matrix symplectic group and its structured subsets. With the unconstrained parametrization, we are able to apply the unconstrained optimization algorithms to the problems with symplectic constraints.

Although we proved that any symplectic matrix can be factored into 9 unit triangular symplectic matrices, it is still unknown that whether 9 is the optimal number in this factorization. What we know regarding to the optimal number is that it is indeed between 4 and 9. This problem is left for future work.

REFERENCES

[1] R. ABRAHAM, J. E. MARSDEN, AND J. E. MARSDEN, *Foundations of Mechanics*, vol. 36, Benjamin/Cummings Publishing Company Reading, MA, 1978.

[2] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, vol. 60, Springer Science & Business Media, New York, 2013.

[3] V. I. ARNOLD, V. V. KOZLOV, AND A. I. NEISHTADT, *Mathematical Aspects of Classical and Celestial Mechanics*, vol. 3, Springer Science & Business Media, New York, 2007.

[4] E. ARTIN, *Geometric Algebra*, Courier Dover Publications, 2016.

[5] M. BENZI AND N. RAZOUK, *On the Iwasawa decomposition of a symplectic matrix*, Appl. Math. Lett., 20 (2007), pp. 260–265.

[6] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, vol. 9, SIAM, Philadelphia, 1994.

[7] P. BIRTEA, I. CAŞU, AND D. COMĂNESCU, *Optimization on the real symplectic group*, Monatshefte für Mathematik, 191 (2020), pp. 465–485.

[8] F. BÜNGER AND S. M. RUMP, *Yet more elementary proofs that the determinant of a symplectic matrix is* 1, Linear Algebra and Its Applications, 515 (2017), pp. 87–95.

[9] A. BUNSE-GERSTNER, *Matrix factorizations for symplectic QR-like methods*, Linear Algebra and Its Applications, 83 (1986), pp. 49–77.

[10] M. A. DE GOSSON, *Symplectic Geometry and Quantum Mechanics*, vol. 166, Springer Science & Business Media, New York, 2006.

[11] F. M. DOPICO AND C. R. JOHNSON, *Complementary bases in symplectic matrices and a proof that their determinant is one*, Linear Algebra and Its Applications, 419 (2006), pp. 772–778.

[12] F. M. Dopico and C. R. Johnson, *Parametrization of the matrix symplectic group and applications*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 650–673.

[13] H. Fassbender, *Symplectic Methods for the Symplectic Eigenproblem*, Springer Science & Business Media, New York, 2007.

[14] K. Feng and M. Qin, *Symplectic Geometric Algorithms for Hamiltonian Systems*, Springer, New York, 2010.

[15] S. Fiori, *Solving minimal-distance problems over the manifold of real-symplectic matrices*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 938–968.

[16] U. Flaschka, V. Mehrmann, and D. Zywietz, *An analysis of structure preserving numerical methods for symplectic eigenvalue problems*, Prod. Inform. Ind, 25 (1991), pp. 165–189.

[17] A. T. Fomenko, *Symplectic Geometry*, vol. 5, CRC Press, 1995.

[18] G. Freiling, V. Mehrmann, and H. Xu, *Existence, uniqueness, and parametrization of Lagrangian invariant subspaces*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 1045–1069.

[19] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-preserving Algorithms for Ordinary Differential Equations*, vol. 31, Springer Science & Business Media, New York, 2006.

[20] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*, vol. 16, SIAM, Philadelphia, 1999.

[21] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 143–159.

[22] V. A. Iakubovich and V. M. Starzhinskiĭ, *Linear Differential Equations with Periodic Coefficients*, vol. 2, Wiley, New York, 1975.

[23] P. Jin, Z. Zhang, A. Zhu, Y. Tang, and G. E. Karniadakis, *SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems*, Neural Networks, 132 (2020), pp. 166–179.

[24] J. Kronsbein and T. Welton, *Note on parametrization of the four-dimensional real symplectic group*, J. Lond. Math. Soc., 1 (1964), pp. 291–295.

[25] Y. Long, *Maslov-type index, degenerate critical points, and asymptotically linear Hamiltonian systems*, Science in China, Ser. A, 33 (1990), pp. 1409–1419.

[26] Y. Long, *The structure of the singular symplectic matrix set*, Science in China, Ser. A, 34 (1991), pp. 897–907.

[27] D. S. Mackey and N. Mackey, *On the Determinant of Symplectic Matrices*, Manchester Centre for Computational Mathematics, 2003.

[28] V. Mehrmann, *A symplectic orthogonal method for single input or single output discrete time optimal quadratic control problems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 221–247.

[29] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution*, vol. 163, Springer, New York, 1991.

[30] F. D. Murnaghan, *The parametrisation and element of volume of the unitary symplectic group*, Proc. Nat. Acad. Sci. USA, 39 (1953), p. 324.

[31] M. Sadkane and A. Salam, *A note on symplectic block reflectors*, Electron. Trans. Numer. Anal., 33 (2009), pp. 45–52.

[32] J.-M. Sanz-Serna and M.-P. Calvo, *Numerical Hamiltonian Problems*, Courier Dover Publications, 2018.

[33] D. E. Taylor, *The Geometry of the Classical Groups*, vol. 9, Heldermann Verlag, 1992.

[34] J. Wang, H. Sun, and S. Fiori, *A Riemannian-steepest-descent approach for optimization on the real symplectic group*, Math. Meth. Appl. Sci., 41 (2018), pp. 4273–4286.

[35] H. Weyl, *The Classical Groups: Their Invariants and Representations*, vol. 1, Princeton University Press, Princeton, NJ, 1946.

[36] H. Xu, *An SVD-like matrix decomposition and its applications*, Linear Algebra and Its Applications, 368 (2003), pp. 1–24.