Numerical Gradient Algorithms for Eigenvalue and Singular Value Calculations
— Source link ↗

John B. Moore, Robert Mahony, Uwe Helmke

Topics: Divide-and-conquer eigenvalue algorithm, Symmetric matrix, Singular value decomposition, Matrix decomposition and Singular solution

Related papers:

- Singular-Value Decomposition via Gradient and Self-Equivalent Flows

- Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems

- The projected gradient methods for least squares matrix approximations with spectral constraints

- Matrix computations

- The Geometry of Algorithms with Orthogonality Constraints

Share this paper:

View more about this paper here: https://typeset.io/papers/numerical-gradient-algorithms-for-eigenvalue-and-singular-445pxex2dg

NUMERICAL GRADIENT ALGORITHMS FOR EIGENVALUE AND SINGULAR VALUE CALCULATIONS*

J.B. MOORE[†], R.E. MAHONY[†], AND U. HELMKE[‡]

Abstract. Recent work has shown that the algebraic question of determining the eigenvalues, or singular values, of a matrix can be answered by solving certain continuous-time gradient flows on matrix manifolds. To obtain computational methods based on this theory, it is reasonable to develop algorithms that iteratively approximate the continuous-time flows. In this paper the authors propose two algorithms, based on a double Lie-bracket equation recently studied by Brockett, that appear to be suitable for implementation in parallel processing environments. The algorithms presented achieve, respectively, the eigenvalue decomposition of a symmetric matrix and the singular value decomposition of an arbitrary matrix. The algorithms have the same equilibria as the continuous-time flows on which they are based and inherit the exponential convergence of the continuous-time solutions.

Key words. eigenvalue decomposition, singular value decomposition, numerical gradient algorithm

AMS subject classifications. 15A18, 65F10

1. Introduction. A traditional algebraic approach to determining the eigenvalue and eigenvector structure of an arbitrary matrix is the QR-algorithm. In the early 1980s it was observed that the QR-algorithm is closely related to a continuous-time differential equation that has become known through study of the Toda lattice. Symes [13] and Deift, Nanda, and Tomei [6] showed that for tridiagonal real symmetric matrices, the QR-algorithm is a discrete-time sampling of the solution to a continuous-time differential equation. This result was generalised to full complex matrices by Chu [3], and Watkins and Elsner [14] provided further insight in the late 1980s.

Brockett [2] studied dynamic matrix flows generated by the double Lie-bracket equation

$$\dot{H} = [H, [H, N]], \qquad H(0) = H_0$$

for constant symmetric matrices N and H_0, and where we use the Lie-bracket notation $[X, Y] = XY - YX$. We call this differential equation the *double-bracket* equation, and we call solutions of this equation *double-bracket flows*. Similar matrix differential equations in the area of Physics were known and studied prior to the references given above. An example, is the Landau–Lifschitz–Gilbert equation of micromagnetics

$$\frac{d\hat{m}}{dt} = \frac{\gamma}{1 + \alpha^2}(\hat{m} \times \overline{H} - \alpha\hat{m} \times (\hat{m} \times \overline{H})) \quad |\hat{m}|^2 = 1,$$

as $\alpha \to \infty$ and $\gamma/\alpha \to k$, a constant. In this equation $\hat{m}, \overline{H} \in \mathbb{R}^3$ and the cross-product is equivalent to a Lie-bracket operation. The relevance of such equations

to traditional linear algebra problems, however, has only recently been studied and discretisations of such flows have not been investigated.

The double-bracket equation is not known to be a continuous-time version of any previously existing linear algebra algorithm; however, it exhibits exponential convergence to an equilibrium point on the manifold of self-equivalent symmetric matrices [2], [5], [9]. Brockett [2] was able to show that this flow could be used to diagonalise real symmetric matrices, and thus, to find their eigenvalues, sort lists, and even to solve linear programming problems. Part of the flexibility and theoretical appeal of the double-bracket equation follows from its dependence on the arbitrary matrix parameter N, which can be varied to control the transient behaviour of the differential equation.

In independent work by Driessel [7], Chu and Driessel [5], Smith [12] and Helmke and Moore [8], a similar gradient flow approach is developed for the task of computing the singular values of a general nonsymmetric, nonsquare matrix. The differential equation obtained in these approaches is almost identical to the double-bracket equation. In [8], it is shown that these flows can also be derived as special cases of the double-bracket equation for a nonsymmetric matrix, suitably augmented to be symmetric.

With the theoretical aspects of these differential equations becoming known, and with applications in the area of balanced realizations [10], [11] along with the more traditional matrix eigenvalue problems, there remains the question of efficiently computing their solutions. No explicit solutions to the differential equations have been obtained and a direct numerical estimate of their integral solutions seems unlikely to be an efficient computational algorithm. Iterative algorithms that approximate the continuous-time flows, however, seem more likely to yield useful numerical methods. Furthermore, discretisations of such isospectral matrix flows are of general theoretical interest in the field of numerical linear algebra. For example, the algorithms proposed in this paper involve adjustable parameters, such as step-size selection schemes and a matrix parameter N, which are not present in traditional algorithms such as the QR-algorithm or the Jacobi method.

In this paper, we propose a new algorithm termed the *Lie-bracket algorithm*, for computing the eigenvalues of an arbitrary symmetric matrix

$$H_{k+1} = e^{-\alpha_k [H_k, N]} H_k e^{\alpha_k [H_k, N]}.$$

For suitably small α_k, termed *time-steps*, the algorithm is an approximation of the solution to the continuous time double-bracket equation. Thus, the algorithm represents an approach to developing new recursive algorithms based on approximating suitable continuous-time flows. We show that for suitable choices of time-steps, the Lie-bracket algorithm inherits the same equilibria as the double-bracket flow. Furthermore, exponential convergence of the algorithm is shown. This paper presents only theoretical results on the Lie-bracket algorithm and does not attempt to compare its performance to that of existing methods for calculating the eigenvalues of a matrix.

Continuous-time gradient flows that compute the singular values of arbitrary nonsymmetric matrices, such as those covered in [5], [8], [9], [12], have a similar form to the double-bracket equation on which the Lie-bracket algorithm was based. We use this similarity to generate a new scheme for computing the singular values of a general matrix termed the *singular value algorithm*. The natural equivalence between the Lie-bracket algorithm and the singular value algorithm is demonstrated and exponential convergence results follow almost directly.

Associated with the main algorithms presented for the computation of the eigenvalues or singular values of matrices are algorithms that compute the full eigenspace decompositions of given matrices. These algorithms are closely related to the Lie-bracket algorithm and also display exponential convergence.

The paper is divided into eight sections including the Introduction and an Appendix. In §2 of this paper, we consider the Lie-bracket algorithm and prove a proposition that ensures the algorithm converges to a fixed point. Section 3 deals with choosing step-size selection schemes and proposes two valid deterministic functions for defining the time-steps. Considering the particular step-size selection schemes presented in §3 we return to the question of stability in §4 and show that the Lie-bracket algorithm has a unique exponentially attractive fixed point, though several of the technical proofs are deferred to the Appendix. This completes the discussion for the symmetric case and §5 considers the nonsymmetric case and the singular value decomposition. Section 6 presents associated algorithms that compute the eigenspace decompositions of given initial conditions. A number of computational issues are briefly mentioned in §7, while §8 provides a conclusion.

2. The Lie-bracket algorithm. In this section, we begin by introducing the least squares potential that underpins the recent gradient flow results and then we describe the double Lie-bracket equation first derived by Brockett [2]. The Lie-bracket recursion is introduced and conditions are given that guarantee convergence of the algorithm.

Let N and H be real symmetric matrices and consider the potential function

$$
\begin{aligned}
\psi(H) &:= ||H - N||^2 \\
&= ||H||^2 + ||N||^2 - 2\mathrm{tr}(NH),
\end{aligned}
\tag{1}
$$

where the norm used is the Frobenius norm $||X||^2 := \mathrm{tr}(X^T X) = \sum x_{ij}^2$, with x_{ij} the elements of X. Note that $\psi(H)$ measures the least squares difference between the elements of H and the elements of N. Let $M(H_0)$ be the set of orthogonally similar matrices, generated by some symmetric initial condition $H_0 = H_0^T \in \mathbb{R}^{n \times n}$. Then

$$
M(H_0) = \{U^T H_0 U \mid U \in O(n)\},
\tag{2}
$$

where $O(n)$ denotes the group of all $n \times n$ real orthogonal matrices. It is shown in [9, p. 48] that $M(H_0)$ is a smooth compact Riemannian manifold with explicit forms given for its tangent space and Riemannian metric. Furthermore, in [1], [5] the gradient of $\psi(H)$, with the respect to the normal Riemannian metric on $M(H_0)$ [9, p. 50], is shown to be $\nabla\psi(H) = -[H, [H, N]]$. Consider the gradient flow given by the solution of

$$
\begin{aligned}
\dot{H} &= -\nabla\psi(H) \\
&= [H, [H, N]], \quad \text{with } H(0) = H_0,
\end{aligned}
\tag{3}
$$

which we call the *double-bracket flow* [2], [5]. Thus, the double-bracket flow is a gradient flow that acts to decrease or minimise the least squares potential ψ on the manifold $M(H_0)$. Note that from (1), this is equivalent to increasing or maximising $\mathrm{tr}(NH)$. We refer to the matrix H_0 as the *initial condition* and the matrix N as the *target matrix.*

The *Lie-bracket algorithm* proposed in this paper is

$$
H_{k+1} = e^{-\alpha_k [H_k, N]} H_k e^{\alpha_k [H_k, N]}
\tag{4}
$$

for arbitrary symmetric $n \times n$ matrices H_0 and N and some suitably small scalars α_k termed *time-steps*. To motivate the Lie-bracket algorithm, consider the curve $H_{k+1}(t) = e^{-t[H_k, N]} H_k e^{t[H_k, N]}$. Thus, $H_{k+1}(0) = H_k$ and $H_{k+1} = H_{k+1}(\alpha_k)$, the $(k+1)$th iteration of (4). Observe that

$$\frac{d}{dt}(e^{-t[H_k, N]} H_k e^{t[H_k, N]})\Big|_{t=0} = [H_k, [H_k, N]],$$

and thus, $e^{-t[H_k, N]} H_k e^{t[H_k, N]}$ is a first approximation of the double-bracket flow at $H_k \in M(H_0)$. It follows that for small α_k, the solution to (3) evaluated at time $t = \alpha_k$ with $H(0) = H_k$ is approximately $H_{k+1} = H_{k+1}(\alpha_k)$.

It is easily seen from above that stationary points of (3) are fixed points of (4). In general, (4) may have more fixed points than just the stationary points of (3), however, Proposition 2.1 shows that this is not the case for a suitable choice of time-step α_k. We use the term *equilibrium point* to mean a fixed point of the algorithm that is also a stationary point of (3).

To implement (4) it is necessary to specify the time-steps α_k. We do this by considering functions $\alpha_N : M(H_0) \to \mathbb{R}_+$ and setting $\alpha_k := \alpha_N(H_k)$. We refer to the function α_N as the *step-size selection scheme*. We require that the step-size selection scheme satisfies the following condition.

CONDITION 2.1. *Let $\alpha_N : M(H_0) \to \mathbb{R}_+$ be a step-size selection scheme for the Lie-bracket algorithm on $M(H_0)$. Then α_N is well defined and continuous on all of $M(H_0)$, except possibly those points $H \in M(H_0)$ where $HN = NH$. Furthermore, there exist real numbers $B, \gamma > 0$, such that $B > \alpha_N(H) \geq \gamma$ for all $H \in M(H_0)$ where α_N is well defined.*

Remark 2.1. We find that the variable step-size selection scheme proposed in this paper, which provides the best simulation results, is discontinuous at all the points $H \in M(H_0)$, such that $[H, N] = 0$.

Remark 2.2. Note that the definition of a step-size selection scheme depends implicitly on the matrix parameter N. Indeed, α_N can be thought of as a function in two matrix variables N and H.

CONDITION 2.2. *Let N be a diagonal $n \times n$ matrix with distinct diagonal entries $\mu_1 > \mu_2 > \ldots > \mu_n$.*

Remark 2.3. This condition on N, along with Condition 2.1 on the step-size selection scheme, is chosen to ensure that the Lie-bracket algorithm converges to a diagonal matrix from which the eigenvalues of H_0 can be directly determined.

Let $\lambda_1 > \lambda_2 > \ldots > \lambda_r$ be the eigenvalues of H_0 with associated algebraic multiplicities $n_1, \ldots, n_r$ satisfying $\sum_{i=1}^{r} n_i = n$. Note that as H_0 is symmetric, the eigenvalues of H_0 are all real. Thus, the diagonalisation of H_0 is

$$(5) \qquad \Lambda := \begin{bmatrix} \lambda_1 I_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r I_{n_r} \end{bmatrix},$$

where I_{n_i} is the $n_i \times n_i$ identity matrix. For generic initial conditions and a target matrix N that satisfies Condition 2.2, the continuous-time equation (3) converges exponentially fast to Λ [2], [9]. Thus, the eigenvalues of H_0 are the diagonal entries of the limiting value of the infinite time solution to (3). The Lie-bracket algorithm behaves similarly to (3) for small α_k and, given a suitable step-size selection scheme, should converge to the same equilibrium as the continuous-time equation.

PROPOSITION 2.1.   *Let H_0 and N be $n \times n$ real symmetric matrices where N satisfies Condition 2.2. Let $\psi(H)$ be given by (1) and let $\alpha_N : M(H_0) \to \mathbb{R}_+$ be a step-size selection scheme that satisfies Condition 2.1. For $H_k \in M(H_0)$, let $\alpha_k = \alpha_N(H_k)$ and define*

(6)
$$\triangle \psi(H_k, \alpha_k) := \psi(H_{k+1}) - \psi(H_k),$$

where H_{k+1} is given by (4). Suppose

(7)
$$\triangle \psi(H_k, \alpha_k) < 0 \quad when \ [H_k, N] \neq 0.$$

Then (a) The iterative equation (4) defines an isospectral (eigenvalue preserving) recursion on the manifold $M(H_0)$.

(b) The fixed points of (4) are characterised by matrices $H \in M(H_0)$ satisfying

(8)
$$[H, N] = 0.$$

(c) Every solution H_k, for $k = 1, 2, \ldots$, of (4), converges as $k \to \infty$, to some $H_\infty \in M(H_0)$ where $[H_\infty, N] = 0$.

Proof. To prove part (a), note that the Lie-bracket $[H, N]^T = -[H, N]$ is skew-symmetric. As the exponential of a skew-symmetric matrix is orthogonal, (4) is an orthogonal conjugation of H_k and hence is isospectral.

For part (b) note that if $[H_k, N] = 0$, then by direct substitution into (4) we see $H_{k+1} = H_k$ and thus, $H_{k+l} = H_k$ for $l \geq 1$, and H_k is a fixed point of (4). Conversely if $[H_k, N] \neq 0$, then from (7), $\triangle \psi(H_k, \alpha_k) \neq 0$, and thus $H_{k+1} \neq H_k$. By inspection, points satisfying (8) are stationary points of (3), and indeed are known to be the only stationary points of (3) [9, pg. 50]. Thus, the fixed points of (4) are equilibrium points in the sense that they are all stationary points of (3). To prove part (c) we need the following lemma.

LEMMA 2.2.   *Let N satisfy Condition 2.2 and α_N satisfy Condition 2.1 such that the Lie-bracket algorithm satisfies (7). The Lie-bracket algorithm (4) has exactly $n!/\prod_{i=1}^r (n_i!)$ distinct equilibrium points in $M(H_0)$. These equilibrium points are characterised by the matrices $\pi^T \Lambda \pi$, where π is an $n \times n$ permutation matrix, a rearrangement of the rows of the identity matrix, and Λ is given by (5).*

Proof. Note that part (b) of Proposition 2.1 characterises equilibrium points of (4) as $H \in M(H_0)$ such that $[H, N] = 0$. Evaluating this condition componentwise for $H = \{h_{ij}\}$ gives

$$h_{ij}(\mu_j - \mu_i) = 0,$$

and hence by Condition 2.2, $h_{ij} = 0$ for $i \neq j$. Using the fact that (4) is isospectral, it follows that equilibrium points are diagonal matrices that have the same eigenvalues as H_0. Such matrices are distinct and can be written in the form $\pi^T \Lambda \pi$ for π an $n \times n$ permutation matrix. A simple counting argument yields the number of matrices that satisfy this condition to be $n!/\prod_{i=1}^r (n_i!)$. $\square$

Consider for a fixed initial condition H_0, the sequence H_k generated by the Lie-bracket algorithm. Observe that condition (7) implies that $\psi(H_k)$ is strictly monotonic decreasing for all k where $[H_k, N] \neq 0$. Also, since ψ is a continuous function on the compact set $M(H_0)$, then ψ is bounded from below and $\psi(H_k)$ will converge to some nonnegative value ψ_∞. As $\psi(H_k) \to \psi_\infty$ then $\triangle \psi(H_k, \alpha_k) \to 0$.

For an arbitrary positive number ϵ, define the open set $D_\epsilon \subset M(H_0)$, consisting of all points of $M(H_0)$, within an ϵ neighbourhood of some equilibrium point of (4).

The set $M(H_0) - D_\epsilon$ is a closed, compact subset of $M(H_0)$ on which the matrix function $H \mapsto [H, N]$ does not vanish. As a consequence, the difference function (6) is continuous and strictly negative on $M(H_0) - D_\epsilon$, and thus can be over bounded by some strictly negative number $\delta_1 < 0$. Moreover, as $\triangle\psi(H_k, \alpha_k) \to 0$, then there exists a $K = K(\delta_1)$ such that for all $k > K$ then $0 \geq \triangle\psi(H_k, \alpha_k) > \delta_1$. This ensures that $H_k \in D_\epsilon$ for all $k > K$. In other words, H_k is converging to some subset of possible equilibrium points.

Imposing the upper bound B on the step-size selection scheme α_N, Condition 2.2, it follows that $\alpha_N(H_k)[H_k, N] \to 0$ as $k \to \infty$. Thus, $e^{\alpha_N(H_k)[H_k,N]} \to I$, the identity matrix, and hence, $e^{-\alpha_N(H_k)[H_k,N]}H_k e^{\alpha_N(H_k)[H_k,N]} \to H_k$ as $k \to \infty$. As a consequence $||H_{k+1} - H_k|| \to 0$ for $k \to \infty$ and this combined with the distinct nature of the fixed points, Lemma 2.2, and the partial convergence already shown, completes the proof. $\square$

Remark 2.4. In Condition 2.2 it was required that N have distinct diagonal entries. If this condition is not satisfied, the equilibrium condition $[H, N] = 0$ may no longer force H to be diagonal, and thus, though the algorithm will converge, it is unlikely to converge to a diagonal matrix.

3. Step-size selection. The Lie-bracket algorithm (4) requires a suitable step-size selection scheme before it can be implemented. To generate such a scheme, we use the potential (1) as a measure of the convergence of (4) at each iteration. Thus, we aim to choose each time-step to maximise the absolute change in potential $| \triangle \psi|$ of (6), such that $\triangle\psi < 0$. Optimal time-steps can be determined at each step of the iteration by completing a line search to maximise the absolute change in potential as the time-step is increased. Such an approach, however, involves high computational overheads and we aim rather to obtain a step-size selection scheme in the form of a scalar equation depending on known values.

Using the Taylor expansion, we express $\triangle\psi(H_k, \tau)$ for a general time-step τ, as a linear term plus a higher order error term. By estimating the error term we obtain a mathematically simple function $\triangle\psi_U(H_k, \tau)$, which is an upper bound to $\triangle\psi(H_k, \tau)$ for all τ. Then, choosing a suitable time-step α_k based on minimising $\triangle\psi_U$, we guarantee that the actual change in potential, $\triangle\psi(H_k, \alpha_k) \leq \triangle\psi_U(H_k, \alpha_k) < 0$, satisfies (7). Due to the simple nature of the function $\triangle\psi_U$, there is an explicit form for the time-step α_k depending only on H_k and N. We begin by deriving an expression for the error term.

LEMMA 3.1. *For the kth step of the recursion (4) the change in potential* $\triangle\psi(H_k, \tau)$ *of (6), for a time-step τ is*

$$(9) \qquad \triangle\psi(H_k, \tau) = -2\tau||[H_k, N]||^2 - 2\tau^2 \text{tr}(N\mathcal{R}_2(\tau))$$

with

$$(10) \qquad \mathcal{R}_2(\tau) := \int_0^1 (1-s)H''_{k+1}(s\tau)ds,$$

where $H''_{k+1}(\tau)$ is the second derivative of $H_{k+1}(\tau)$ with respect to τ.

Proof. Let $H_{k+1}(\tau)$ be the $(k+1)$th recursive estimate for an arbitrary time-step τ. Thus $H_{k+1}(\tau) = e^{-\tau[H_k,N]}H_k e^{\tau[H_k,N]}$. It is easy to verify that the first and second time derivatives of H_{k+1} are exactly

$$H'_{k+1}(\tau) = [H_{k+1}(\tau), [H_k, N]],$$
$$H''_{k+1}(\tau) = [[H_{k+1}(\tau), [H_k, N]], [H_k, N]].$$

Applying Taylor's theorem, then

(11)
$$H_{k+1}(\tau) = H_{k+1}(0) + \tau\frac{d}{d\tau}H_{k+1}(0) + \tau^2\int_0^1 (1-s)H''_{k+1}(s\tau)ds,$$
$$= H_k + \tau[H_k,[H_k,N]] + \tau^2\mathcal{R}_2(\tau).$$

Consider the change in the potential $\psi(H)$ between the points H_k and $H_{k+1}(\tau)$,

(12)
$$\begin{aligned}
\triangle\psi(H_k,\tau) &= \psi(H_{k+1}(\tau)) - \psi(H_k) \\
&= -2\mathrm{tr}(N(H_{k+1}(\tau) - H_k)) \\
&= -2\mathrm{tr}(N(\tau[H_k,[H_k,N]] + \tau^2\mathcal{R}_2(\tau))) \\
&= -2\tau||[H_k,N]||^2 - 2\tau^2\mathrm{tr}(N\mathcal{R}_2(\tau)). \qquad \square
\end{aligned}$$

Note that for $\tau = 0$, then $\triangle\psi(H_k,0) = 0$ and also that

$$\frac{d}{d\tau}\triangle\psi(H_k,\tau)\bigg|_{\tau=0} = -2||[H_k,N]||^2.$$

Thus, for sufficiently small τ the error term $\tau^2\mathrm{tr}(N\mathcal{R}_2(\tau))$ becomes negligible and $\triangle\psi(H_k,\tau)$ is strictly negative. Let $\alpha_{\mathrm{opt}} > 0$ be the first time for which

$$\frac{d}{d\tau}\triangle\psi(H_k,\tau)\bigg|_{\tau=\alpha_{\mathrm{opt}}} = 0,$$

then $\triangle\psi(H_k,\alpha_{\mathrm{opt}}) < \triangle\psi(H_k,\tau) < 0$ for all strictly positive $\tau < \alpha_{\mathrm{opt}}$. It is not possible, however, to estimate α_{opt} directly from (12) due to the transcendental nature of the error term $\mathcal{R}_2(\tau)$. By considering two separate estimates of the error term, we obtain two step-size selection schemes $\alpha_k \leq \alpha_{\mathrm{opt}}$. The first and constant step-size selection scheme follows from a loose bound of the error, whereas the second variable step-size selection scheme follows from a more sophisticated argument and results in faster convergence of (4).

LEMMA 3.2 (Constant step-size selection scheme). *The constant time-step*

(13)
$$\alpha_N^c = \frac{1}{4||H_0|| \cdot ||N||}$$

satisfies Condition 2.1. Furthermore, the Lie-bracket algorithm, equipped with the step-size selection scheme α_N^c, satisfies (7).

Proof. Recall that for the Frobenius norm $|\mathrm{tr}(XY)| \leq ||X|| \cdot ||Y||$. Then

(14)
$$\begin{aligned}
\triangle\psi(H_k,\tau) &\leq -2\tau||[H_k,N]||^2 + 2\tau^2|\mathrm{tr}(N\mathcal{R}_2(\tau))| \\
&\leq -2\tau||[H_k,N]||^2 + 2\tau^2||N|| \cdot ||\mathcal{R}_2(\tau)|| \\
&\leq -2\tau||[H_k,N]||^2 + 2\tau^2||N|| \cdot \\
&\qquad \int_0^1 (1-s)||[[H_{k+1}(s\tau),[H_k,N]],[H_k,N]]||ds \\
&\leq -2\tau||[H_k,N]||^2 + 4\tau^2||N|| \cdot ||H_0|| \cdot ||[H_k,N]||^2 \\
&=: \triangle\psi_U(H_k,\tau).
\end{aligned}$$

Thus $\triangle\psi_U(H_k,\tau)$ is an upper bound for $\triangle\psi(H_k,\tau)$ and has the property that for sufficiently small τ, it is strictly negative; see Fig. 1. Due to the quadratic form of
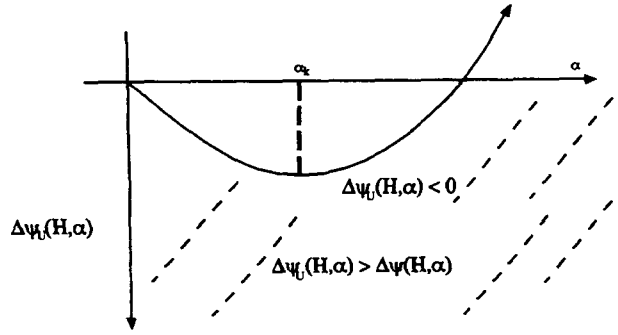
FIG. 1. *The upper bound on $\triangle\psi(H_k,\alpha)$ viz $\triangle\psi_U(H_k,\alpha)$.*

$\triangle\psi_U(H_k,\tau)$ in τ, it is immediately clear that $\alpha_k^c = \alpha_N^c(H_k) = 1/(4||H_0||||N||)$ of (13) is the minimum of (14). $\square$

A direct norm bound of the integral error term is not likely to be a tight estimate of the error and the function $\triangle\psi_U$ is a fairly crude bound for $\triangle\psi$. The following more sophisticated estimate results in a step-size selection scheme that causes the Lie-bracket algorithm to converge an order of magnitude faster.

LEMMA 3.3 (An improved bound for $\triangle\psi(H_k,\tau)$). *Note the difference function* $\triangle\psi(H_k,\tau)$ *can be over bounded by*

$$
\begin{aligned}
\triangle\psi(H_k,\tau) &\leq -2\tau||[H_k,N]||^2 \\
(15) \qquad &+ \frac{||H_0|| \cdot ||[N,[H_k,N]]||}{||[H_k,N]||} \left(e^{2\tau||[H_k,N]||} - 1 - 2\tau||[H_k,N]||\right) \\
&=: \triangle\psi_U^*(H_k,\tau).
\end{aligned}
$$

Proof. Consider the Taylor series expansion of the matrix exponential

$$
e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots.
$$

It is easily verified that

$$
\begin{aligned}
(16) \qquad e^A B e^{-A} &= B + [A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \cdots \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} ad_A^i B.
\end{aligned}
$$

Here $ad_A^i B = ad_A (ad_A^{i-1} B)$, $ad_A^0 B = B$, where $ad_A : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ is the linear map $X \mapsto AX - XA$. Substituting $-\tau[H_k,N]$ and H_k for A and B in (16) and comparing with (11), gives

$$
\tau^2 \mathcal{R}_2(\tau) = \sum_{j=2}^{\infty} \frac{1}{j!} ad_{-\tau[H_k,N]}^j (H_k).
$$

Considering $|\text{tr}(N\mathcal{R}_2(\tau))|$ and using the readily established identity $\text{tr}(N ad_{-A}^j B) =$

$\operatorname{tr}((ad_A^j N)B)$ gives

$$\left| \tau^2 \operatorname{tr}(N\mathcal{R}_2(\tau)) \right| = \left| \sum_{j=2}^{\infty} \frac{1}{j!} \operatorname{tr}\left(ad_{\tau[H_k,N]}^j (N) H_k \right) \right|$$

$$\leq \sum_{j=2}^{\infty} \frac{1}{j!} \|ad_{\tau[H_k,N]}^j (N)\| \cdot \|H_0\|$$

$$\leq \sum_{j=2}^{\infty} \frac{1}{j!} (2\|\tau[H_k,N]\|)^{j-1} \|ad_{\tau[H_k,N]}(N)\| \cdot \|H_0\|$$

$$= \frac{\|H_0\| \cdot \|ad_{\tau[H_k,N]}(N)\|}{2\tau\|[H_k,N]\|} \sum_{j=2}^{\infty} \frac{1}{j!} (2\tau\|[H_k,N]\|)^j$$

$$= \frac{\|H_0\| \cdot \|[N,[H_k,N]]\|}{2\|[H_k,N]\|} \left(e^{2\tau\|[H_k,N]\|} - 1 - 2\tau\|[H_k,N]\| \right).$$

Thus combining this with the first line of (14) gives (15). □

The variable step-size selection scheme is derived from this estimate of the error term in the same manner the constant step-size selection scheme was derived in Lemma 3.2.

LEMMA 3.4 (Variable step-size selection scheme). *The step-size selection scheme* $\alpha_N^* : M(H_0) \to \mathbb{R}_+$

$$(17) \qquad \alpha_N^*(H) = \frac{1}{2\|[H,N]\|} \log \left(\frac{\|[H,N]\|^2}{\|H_0\| \, \|[N,[H,N]]\|} + 1 \right),$$

where all norms are Frobenius norms, satisfies Condition 2.1. Furthermore, the Lie-bracket algorithm, equipped with the step-size selection scheme α_N^*, *satisfies* (7).

Proof. We first show that α_N^* satisfies the requirements of Condition 2.1. As the Frobenius norm is a continuous function, then α_N^* is well defined and continuous at all points $H \in M(H_0)$ such that $[H,N] \neq 0$. Note that when $[H,N] = 0$, then α_N^* is not well defined. To show that there exists a positive constant γ, such that $\alpha_N^*(H) > \gamma$, consider the following lower bound,

$$\alpha_N^L := \frac{1}{2\|[H_k,N]\|} \log \left(\frac{\|[H_k,N]\|}{2\|[H_0\| \, \|N\|} + 1 \right)$$

$$(18) \qquad \leq \frac{1}{2\|[H_k,N]\|} \log \left(\frac{\|[H_k,N]\|^2}{2\|H_0\| \, \|N\| \, \|[H_k,N]\|} + 1 \right)$$

$$\leq \frac{1}{2\|[H_k,N]\|} \log \left(\frac{\|[H_k,N]\|^2}{\|H_0\| \, \|[N,[H_k,N]]\|} + 1 \right),$$

which is just α_N^*. Using L'Hôpital's rule it can be seen that the limit of α_N^L at an equilibrium point, $H \in M(H_0)$ such that $[H,N] = 0$, is $1/(4\|H_0\| \cdot \|N\|)$. Including these points in the definition of α_N^L, gives that α_N^L is a continuous, strictly positive, well-defined function for all $H \in M(H_0)$. Thus, as $M(H_0)$ is compact, there exists a real number $\gamma > 0$ such that

$$\alpha_N^* \geq \alpha_N^L \geq \gamma > 0$$

on $M(H_0) - \{H_\infty \mid [H_\infty, N] = 0\}$.

To show that there exists a real number $B > 0$, such that $\alpha_N^*(H) < B$, $H \in M(H_0)$, set $[H, N] = X = \{x_{ij}\}$. For N given by Condition 2.2, then $\|[N, X]\| = \sum_{i=j}(\mu_i - \mu_j)^2 x_{ij}^2$, where $x_{ii} = 0$ as $[H, N]$ is skew-symmetric. Observe that

$$\|X\|/\|[N, X]\| = \frac{\sum_{i=j} x_{ij}^2}{\sum_{i=j}(\mu_i - \mu_j)^2 x_{ij}^2}$$

$$\leq \max_{i=j}(\mu_i - \mu_j)^{-2} =: b$$

for all choices of $X = -X^T$. It follows that

$$\alpha_N^*(H) = \frac{1}{2\|X\|} \log\left(\frac{\|X\|^2}{\|H_0\| \, \|[N, X]\|} + 1\right)$$

$$\leq \frac{1}{2\|X\|} \log\left(\frac{\|X\| b}{\|H_0\|} + 1\right)$$

$$\leq \frac{b}{2\|H_0\|} =: B$$

since $\log(x + 1) \leq x$ for $x > 0$.

Finally, for a matrix $H_k \in M(H_0)$, $[H_k, N] \neq 0$, the time-step $\alpha_N^*(H_k) = \alpha_k^* > 0$ minimises (15), and from Lemma 3.3 it follows that $0 \geq \triangle\psi_U^*(H_k, \tau) \geq \triangle\psi(H_k, \tau)$. Thus, the Lie-bracket algorithm, equipped with the step-size selection scheme α_N^*, satisfies (7) and the proof is complete.   □

4. Stability analysis.

In this section we study the stability of equilibria of the Lie-bracket algorithm (4). It is shown that for generic initial conditions and any step-size selection scheme that satisfies Condition 2.1 and (7), the solution H_k of the Lie-bracket algorithm converges to the unique equilibrium point Λ given by (5). Furthermore, we derive local exponential bounds on the rate of convergence. To improve the readability of the paper the proofs of a number of the more technical results have been deferred to an appendix. We begin by showing that Λ is the unique locally asymptotically stable equilibrium point of (4).

LEMMA 4.1. *Let N satisfy Condition (2.2) and α_N be some selection scheme that satisfies Condition 2.1 and (7). The Lie-bracket algorithm (4) has a unique locally asymptotically stable equilibrium point Λ given by (5). All other equilibrium points of (4) are unstable.*

Proof. It is known that Λ is the unique local and global minimum of the potential function ψ on $M(H_0)$ [9]. By assumptions on N and α_N, $\psi(H_k)$ is monotonically decreasing. Thus the domain of attraction of Λ contains an open neighbourhood of Λ, and hence, Λ is a locally asymptotically stable equilibrium point of (4).

All other equilibrium points H_∞ are either saddle points or maxima of ψ [9]. Thus for any neighbourhood D of some equilibrium point $H_\infty \neq \Lambda$, there exists some $H_0 \in D$ such that $\psi(H_0) < \psi(H_\infty)$. It follows that the solution to the Lie-bracket algorithm, with initial condition H_0, will not converge to H_∞ and thus H_∞ is unstable.   □

Lemma 4.1 is sufficient to conclude that for generic initial conditions the Lie-bracket algorithm will converge to the unique matrix Λ. It is difficult to characterise the set of initial conditions for which the algorithm converges to some unstable equilibrium point $H_\infty \neq \Lambda$. For the continuous-time double-bracket flow, however, it is

known that the unstable basins of attraction of such points are of zero measure in $M(H_0)$ [9].

LEMMA 4.2. *Let N satisfy Condition 2.2. Let $d \in \mathbb{R}_+$ be a constant such that $0 < d < 1/2||H_0||_2||N||_2$ and consider the constant step-size selection scheme, $\alpha_N^d :$ $M(H_0) \rightarrow \mathbb{R}_+$,*

$$\alpha_N^d(H) = d.$$

The Lie-bracket algorithm (4) equipped with the step-size selection scheme α_N^d has a unique locally exponentially asymptotically stable equilibrium point Λ given by (5).

Proof. Since α_N^d is a constant function, the time-step $\alpha_k^d = \alpha_N^d(H_k) = d$ is constant. Thus, the map

$$H_k \mapsto e^{-d[H_k,N]}H_k e^{d[H_k,N]}$$

is a differentiable map on all $M(H_0)$, and we may consider the linearisation of this map at the equilibrium point Λ given by (5). The linearisation of this recursion expressed in terms of $\Xi_k \in T_\Lambda M(H_0)$ (the tangent space of the equilibrium point Λ) is

$$(19) \qquad \Xi_{k+1} = \Xi_k - d[(\Xi_k N - N\Xi_k)\Lambda - \Lambda(\Xi_k N - N\Xi_k)].$$

Thus for the elements of Ξ_k, we have

$$(20) \qquad (\xi_{ij})_{k+1} = [1 - d(\lambda_i - \lambda_j)(\mu_i - \mu_j)](\xi_{ij})_k \quad \text{for } i,j = 1,\dots,n.$$

The tangent space $T_\Lambda M(H_0)$ at Λ consists of those matrices $\Xi = [\Lambda, \Omega]$ where $\Omega \in$ Skew(n), the class of skew-symmetric matrices [9, p. 53]. Thus, the matrices Ξ are parameterised by their components ξ_{ij}, where $i < j$, and $\lambda_i \neq \lambda_j$. This is a linearly independent parameterisation of $T_\Lambda M(H_0)$ and the eigenvalues of the linearisation (19) can be read directly from (20) as $1 - d(\lambda_{\pi(i)} - \lambda_{\pi(j)})(\mu_i - \mu_j)$, for $i < j$ and $\lambda_i \neq \lambda_j$. Since $\lambda_i \geq \lambda_j$ when $i > j$, then if $d < 1/2||H_0||_2||N||_2$ it follows that

$$|1 - d(\lambda_i - \lambda_j)(\mu_i - \mu_j)| < 1$$

for all $i < j$ with $\lambda_i \neq \lambda_j$. Classical stability theory gives that Λ is a locally exponentially asymptotically stable equilibrium point of the recursion (4) with an exponential rate of convergence of $\max_{i<j,\lambda_i=\lambda_j}\{d(\lambda_i - \lambda_j)(\mu_i - \mu_j)\}$. □

Remark 4.1. As $||N||_2||H_0||_2 < 2||N||||H_0||$, the constant step-size selection scheme α_N^c is an example of such a selection scheme where $c = 1/(4||H_0|| \cdot ||N||)$.

Remark 4.2. Let $\alpha_N : M(H_0) \rightarrow \mathbb{R}_+$ be a step-size selection scheme that satisfies Condition 2.1 and (7) and is also continuous on all $M(H_0)$. Let Λ be the locally asymptotically stable equilibrium point given by (5). Set $\alpha_\infty = \alpha_N(\Lambda)$ and observe that the linearisation of the Lie-bracket algorithm will be of the form (19) with d replaced by α_∞. Recall that the α_N^L scheme defined in (18) is continuous with limit $\alpha_N^L(H_\infty) = 1/(4||H_0|| \cdot ||N||)$. Thus, Λ is an exponentially asymptotically stable equilibrium point for the Lie-bracket recursion equipped with the step-size selection scheme α_N^L.

To show that the Lie-bracket algorithm is exponentially stable at Λ for the α_N^* step-size selection scheme is technically difficult due to the discontinuous nature of α_N^* at equilibrium points. The proof of the following proposition is given in the Appendix.

PROPOSITION 4.3. *Let N satisfy assumption (2.2) and α_N^* be the step-size selection scheme given by Lemma 3.4. The iterative algorithm (4), has a unique exponentially attractive equilibrium point Λ given by (5).*
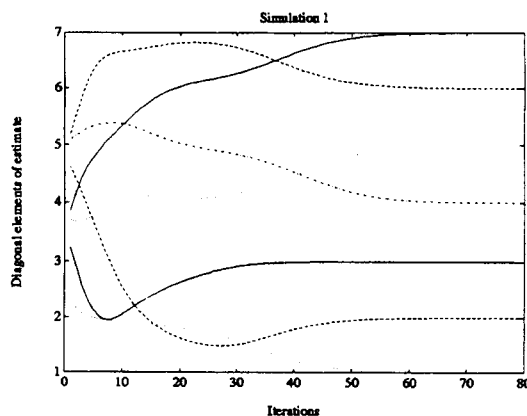
FIG. 2. *A plot of the diagonal elements h_{ii} of each iteration H_k of the Lie-bracket algorithm run on a 7×7 initial condition H_0 with eigenvalues $(1, \ldots, 7)$. The target matrix N was chosen to be $\mathrm{diag}(1, \ldots, 7)$.*
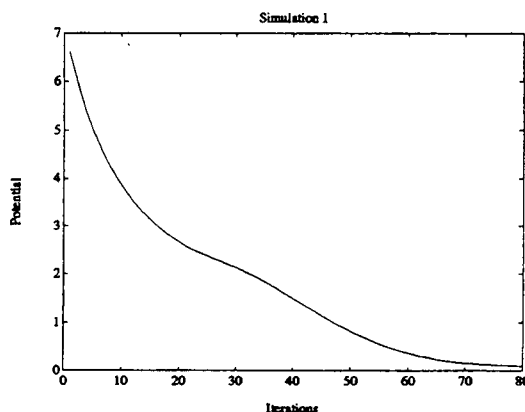

FIG. 3. *The potential $\psi(H_k) = \|H_k - N\|^2$ for the Lie-bracket recursion.*

To give an indication of the behaviour of the Lie-bracket algorithm, two plots of a simulation have been included as Figs. 2 and 3. The simulation was run on a random 7×7 symmetric initial value matrix with eigenvalues $1, \ldots, 7$. The target matrix N is chosen as $\mathrm{diag}(1, \ldots, 7)$ and as a consequence the minimum potential is $\psi_\infty = 0$. Figure 2 is a plot of the diagonal entries of the recursive estimate H_k. The off-diagonal entries converge to zero as the diagonal entries converge to the eigenvalues of H_k. Figure 3 is a plot of the potential $\|H_k - N\|^2$ verses the iteration k. This plot clearly shows the monotonic decreasing nature of the potential at each step of the algorithm.

We summarise the results of §§2–4 in Theorem 4.4.

THEOREM 4.4. *Let $H_0 = H_0^T$ be a real symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Let $N \in \mathbb{R}^{n \times n}$ satisfy Condition 2.2 and let α_N be either the constant step-size selection (13) or the variable step-size selection (17). The Lie-bracket*

recursion

$$H_{k+1} = e^{-\alpha_k [H_k, N]} H_k e^{\alpha_k [H_k, N]},$$
$$\alpha_k = \alpha_N(H_k),$$

with initial condition H_0, has the following properties:

(i) *The recursion is isospectral.*

(ii) *If H_k is a solution of the Lie-bracket algorithm, then $\psi(H_k) = ||H_k - N||^2$ is strictly monotonically decreasing for every $k \in \mathbb{N}$, where $[H_k, N] \neq 0$.*

(iii) *Fixed points of the recursive equation are characterised by matrices $H \in M(H_0)$ such that*

$$[H, N] = 0.$$

(iv) *Fixed points of the recursion are exactly the stationary points of the double-bracket equation. These points are termed* equilibrium *points.*

(v) *Let H_k, $k = 1, 2, \ldots$, be a solution to the Lie-bracket algorithm, then H_k converges to a matrix $H_\infty \in M(H_0)$, $[H_\infty, N] = 0$, an equilibrium point of the recursion.*

(vi) *All equilibrium points of the Lie-bracket algorithm are strictly unstable except $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, which is locally exponentially asymptotically stable.*

5. Singular value computations. In this section we consider discretisations of continuous-time flows to compute the singular values of an arbitrary matrix.

A singular value decomposition of a matrix $H_0 \in \mathbb{R}^{m \times n}$, $m \geq n$ is a matrix decomposition

$$(21) \qquad\qquad\qquad H_0 = V^T \Sigma U,$$

where $V \in O(m)$, $U \in O(n)$ and

$$(22) \qquad\qquad \Sigma = \begin{pmatrix} \sigma_1 I_{n_1} & \cdots & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & \sigma_r I_{n_r} \\ & 0_{(m-n) \times n} & & \end{pmatrix}.$$

Here $\sigma_1 > \sigma_2 > \cdots > \sigma_r \geq 0$ are the distinct singular values of H_0 occurring with multiplicities $n_1, \ldots, n_r$, such that $\sum_{i=1}^r n_i = n$. By convention the singular values of a matrix are chosen to be nonnegative. It should be noted that although such a decomposition always exists and Σ is unique, there is no unique choice of orthogonal matrices V and U. The approach we take is to define an algorithm that converges to Σ and thus computes the singular values of H_0 without directly generating the orthogonal decomposition.

Let $S(H_0)$ be the set of all orthogonally equivalent matrices to H_0,

$$(23) \qquad S(H_0) = \{ V^T H_0 U \in \mathbb{R}^{m \times n} \mid V \in O(m), \ U \in O(n) \}.$$

It is shown in [9, p. 89] that $S(H_0)$ is a smooth compact Riemannian manifold with explicit forms given for its tangent space and Riemannian metric. Following [4], [5], [8], [9], and [12] we consider the task of calculating the singular values of a matrix H_0 by minimising the least squares cost function $\psi : S(H_0) \to \mathbb{R}_+$, $\psi(H) = ||H - N||^2$. It is shown in [8] and [9] that ψ achieves a unique local and global minimum at the

point $\Sigma \in S(H_0)$. Moreover, in [8], [9], and [12] the explicit form for the gradient $\nabla\psi$ is calculated. The gradient flow is

$$
\begin{aligned}
\dot{H} &= -\nabla\psi(H) \\
&= H\{H,N\} - \{H^T, N^T\}H,
\end{aligned}
\tag{24}
$$

with $H(0) = H_0$ the initial condition. Here we have used a generalisation of the Lie-bracket $\{X, Y\} := X^T Y - Y^T X = -\{X, Y\}^T$.

To accomplish the task of computing the singular values of a matrix we require N to satisfy the following.

CONDITION 5.1. *Let N be an $m \times n$ matrix*

$$
N = \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \\ & 0_{(m-n)\times n} & \end{bmatrix},
$$

where $\mu_1 > \mu_2 > \cdots > \mu_n > 0$ are strictly positive, distinct real numbers.

For generic initial conditions and a target matrix N that satisfies Condition 5.1, it is known that (24) converges exponentially fast to $\Sigma \in S(H_0)$ [8], [12]. A recursive version of this flow follows from an analogous argument to that used in the derivation of the Lie-bracket algorithm. For H_0 and N constant $m \times n$ matrices, the *singular value algorithm* proposed is

$$
H_{k+1} = e^{-\alpha_k \{H_k^T, N^T\}} H_k e^{\alpha_k \{H_k, N\}}.
\tag{25}
$$

The singular value algorithm and the Lie-bracket algorithm are closely linked as is shown in the following lemma.

LEMMA 5.1. *Let H_0, N be $m \times n$ matrices. For any $H \in \mathbb{R}^{m \times n}$ define a map $H \mapsto \widehat{H} \in \mathbb{R}^{(m+n)\times(m+n)}$, where*

$$
\widehat{H} = \begin{pmatrix} 0_{m\times m} & H \\ H^T & 0_{n\times n} \end{pmatrix}.
\tag{26}
$$

For any sequence of real numbers α_k, $k = 1, \ldots, \infty$ the iterations

$$
H_{k+1} = e^{-\alpha_k \{H_k^T, N^T\}} H_k e^{\alpha_k \{H_k, N\}}
\tag{27}
$$

with initial condition H_0 and

$$
\widehat{H}_{k+1} = e^{-\alpha_k [\widehat{H}_k, \widehat{N}]} \widehat{H}_k e^{\alpha_k [\widehat{H}_k, \widehat{N}]}
\tag{28}
$$

with initial condition $\widehat{H}_0$ are equivalent.

Proof. Consider the iterative solution to (28) and evaluate the multiplication in the block form of (26). This gives two equivalent iterative solutions, one the transpose of the other, both of which are equivalent to the iterative solution to (27).   □

Remark 5.1. Note that $\widehat{H}_0$ and $\widehat{N}$ are symmetric $(m+n) \times (m+n)$ matrices and that, as a result, the iteration (28) is just the Lie-bracket algorithm.

Remark 5.2. The equivalence given by Lemma 5.1 is complete in every way. In particular, H_∞ is an equilibrium point of (27) if and only if $\widehat{H}_\infty$ is an equilibrium point of (28). Similarly, $H_k \to H_\infty$ if and only if $\widehat{H}_k \to \widehat{H}_\infty$ as $k \to \infty$.

This leads us directly to consider step-size selection schemes for the singular value algorithm induced by selection schemes that we have already considered for the Lie-bracket algorithm. Indeed if $\alpha_{\widehat{N}} : M(\widehat{H}_0) \to \mathbb{R}_+$ is a step-size selection scheme for (4) on $M(\widehat{H}_0)$, and $H_k \in S(H_0)$, then we can define a time-step α_k for the singular value algorithm by

$$(29) \qquad\qquad\qquad \alpha_k = \alpha_{\widehat{N}}(\widehat{H}_k).$$

Thus, if (28) equipped with a step-size selection scheme $\alpha_{\widehat{N}}$ satisfies Condition 2.1 and (7), then from Lemma 5.1, (27) will satisfy similar conditions. For simplicity, we deal only with the step-size selection schemes induced by the constant step-size selection (13) and the variable step-size selection (17). Thus we may state the main convergence theorem for the singular value algorithm.

THEOREM 5.2. *Let H_0, N be $m \times n$ matrices where $m \geq n$ and N satisfies Condition 5.1. Let $\alpha_{\widehat{N}} : M(\widehat{H}_0) \to \mathbb{R}_+$ be either the constant step-size selection (13), or the variable step-size selection (17). The singular value algorithm*

$$H_{k+1} = e^{-\alpha_k \{H_k^T, N^T\}} H_k e^{\alpha_k \{H_k, N\}},$$
$$\alpha_k = \alpha_{\widehat{N}}(\widehat{H}_k),$$

with initial condition H_0, has the following properties:

(i) *The singular value algorithm is a self-equivalent (singular value preserving) recursion on the manifold $S(H_0)$.*

(ii) *If H_k is a solution of the singular value algorithm, then $\psi(H_k) = ||H_k - N||^2$ is strictly monotonically decreasing for every $k \in \mathbb{N}$, where $\{H_k, N\} \neq 0$ and $\{H_k^T, N^T\} \neq 0$.*

(iii) *Fixed points of the recursive equation are characterised by matrices $H \in S(H_0)$ such that*

$$(30) \qquad\qquad \{H_k, N\} = 0 \quad and \quad \{H_k^T, N^T\} = 0.$$

Fixed points of the recursion are exactly the stationary points of the singular value gradient flow (24) and are termed equilibrium *points.*

(iv) *Let H_k, $k = 1, 2, \ldots$, be a solution to the singular value algorithm, then H_k converges to a matrix $H_\infty \in S(H_0)$, an equilibrium point of the recursion.*

(v) *All equilibrium points of the Lie-bracket algorithm are strictly unstable except Σ given by (22), which is locally exponentially asymptotically stable.*

Proof. To prove part (i), note that the generalised Lie-bracket $\{X, Y\} = -\{X, Y\}^T$ is skew-symmetric and thus (25) is an orthogonal conjugation and preserves the singular values of H_k. Also note that the potential $\psi(H_k) = \frac{1}{2}\psi(\widehat{H}_k)$. Moreover, Lemma 5.1 shows that the sequence $\widehat{H}_k$ is a solution to the Lie-bracket algorithm and thus from Proposition 2.1, $\frac{1}{2}\psi(\widehat{H}_k)$ must be monotonically decreasing for all $k \in \mathbb{N}$ such that $[\widehat{H}_k, \widehat{N}] \neq 0$, which is equivalent to (30). This proves part (ii) and part (iii) follows by noting that if $\{H_k^T, N^T\} = 0$ and $\{H_k, N\} = 0$, then $H_{k+l} = H_k$ for $l = 1, 2, \ldots$, and H_k is a fixed point of (25). Moreover, since $\psi(H_k)$ is strictly monotonic decreasing for all $\{H_k, N\} \neq 0$ and $\{H_k^T, N^T\} \neq 0$, then these points can be the only fixed points. It is known that these are the only stationary points of (24) [8], [9], [12].

To prove (iv), we need the following characterisation of equilibria of the singular value algorithm.

LEMMA 5.3. *Let N satisfy Condition* 5.1 *and $\alpha_{\widehat{N}}$ be either the constant step-size selection* (13) *or the variable step-size selection* (17). *The singular value algorithm* (25) *equipped with time-steps $\alpha_k = \alpha_{\widehat{N}}(\widehat{H}_k)$ has exactly $2^n n!/\prod_{i=1}^r (n_i!)$ distinct equilibrium points in $S(H_0)$. Furthermore, these equilibrium points are characterised by the matrices*

$$\begin{pmatrix} \pi^T & 0_{n \times (m-n)} \\ 0_{(m-n) \times n} & 0_{(m-n) \times (m-n)} \end{pmatrix} \Sigma S \pi,$$

where π is an $n \times n$ permutation matrix and $S = \mathrm{diag}(\pm 1, \ldots, \pm 1)$ a sign matrix.

Proof. Equilibrium points of (25) are characterised by the two conditions (30). For $H = (h_{ij})$, $\{H, N\} = 0$ is equivalent to

$$\mu_j h_{ji} - \mu_i h_{ij} = 0 \quad \text{for } i = 1, \ldots, n, \qquad j = 1, \ldots, n.$$

Similarly, the condition $\{H^T, N^T\} = 0$ is equivalent to

$$\mu_j h_{ij} - \mu_i h_{ji} = 0 \quad \text{for } i = 1, \ldots, n, \qquad j = 1, \ldots, n,$$
$$h_{ij} \mu_j = 0 \quad \text{for } i = n+1, \ldots, m, \qquad j = 1, \ldots, n.$$

By manipulating the relationships, and using the distinct, positive nature of the μ_i, it is easily shown that $h_{ij} = 0$ for $i \neq j$. Using the fact that (25) is self equivalent, the only possible matrices of this form that have the same singular values as H_0 are characterised as above. A simple counting argument shows that the number of distinct equilibrium points is $2^n n!/\prod_{i=1}^r (n_i!)$. □

The proof of Theorem 5.2 part (iv) is now directly analogous to the proof of Proposition 2.1 part (c). It remains only to prove Theorem 5.2 part (v), which involves the stability analysis of the equilibrium points characterised by (30). It is not possible to directly apply the results obtained in §4 to the Lie-bracket recursion $\widehat{H}_k$, since the $\widehat{N}$ does not satisfy Condition 2.2. However, for the constant step-size selection scheme induced by (13), and using analogous arguments to those used in Lemmas 4.1 and 4.2, it follows that Σ is the unique locally exponentially attractive equilibrium point of the singular value algorithm. Thus, for the constant step-size selection scheme, $\widehat{\Sigma}$ is the unique exponentially attractive equilibrium point of the Lie-bracket algorithm on $M(\widehat{H}_0)$, and now the argument from Proposition 4.3 applies directly and $\widehat{\Sigma}$ is exponentially attractive for the variable step-size selection scheme (17). This completes the proof. □

Remark 5.3. Theorem 5.2 holds true for any time-steps $\alpha_k = \alpha_{\widehat{N}}(\widehat{H}_k)$ induced by a step-size selection scheme, $\alpha_{\widehat{N}}$, that satisfies Condition 2.1, such that Theorem 4.4 holds.

Remark 5.4. It is possible that for nongeneric initial conditions, the singular value algorithm may converge to a diagonal matrix with the singular values ordered in a different manner to Σ. However, all simulations run have converged exponentially fast to the unique matrix Σ, and thus it is likely that the attractive basins of the unstable equilibrium points have zero measure. Note that for the continuous-time flows, it is known that the attractive basins of the unstable equilibrium points have zero measure in $S(H_0)$ [9].

6. Associated orthogonal algorithms. In the previous sections we have proposed the Lie-bracket and the singular value algorithms that calculate the eigenvalues and singular values, respectively, of given initial conditions. Associated with these recursions are orthogonal recursions that compute the eigenvectors or singular vectors

of given initial conditions and provide a full spectral decomposition. To simplify the subsequent analysis we impose a genericity condition on the initial condition H_0.

CONDITION 6.1. *If $H_0 = H_0^T \in \mathbb{R}^{n \times n}$ is a real symmetric matrix then assume that H_0 has distinct eigenvalues $\lambda_1 > \cdots > \lambda_n$. If $H_0 \in \mathbb{R}^{m \times n}$, where $m \geq n$, then assume that H_0 has distinct singular values $\sigma_1 > \cdots > \sigma_n > 0$.*

For a sequence of positive real numbers α_k for $k = 1, 2, \ldots$ the *associated orthogonal Lie-bracket algorithm* is

$$(31) \qquad U_{k+1} = U_k e^{\alpha_k [U_k^T H_0 U_k, N]}, \qquad U_0 \in O(n),$$

where $H_0 = H_0^T \in \mathbb{R}^{n \times n}$ is symmetric. For an arbitrary initial condition $H_0 \in \mathbb{R}^{m \times n}$ the *associated orthogonal singular value algorithm* is

$$(32) \qquad \begin{aligned} V_{k+1} &= V_k e^{\alpha_k \{U_k^T H_0^T V_k, N^T\}}, \qquad V_0 \in O(m) \\ U_{k+1} &= U_k e^{\alpha_k \{V_k^T H_0 U_k, N\}}, \qquad U_0 \in O(n). \end{aligned}$$

Note that in each case the exponents of the exponential terms are skew-symmetric and thus the recursions will remain orthogonal.

Let $H_0 = H_0^T \in \mathbb{R}^{n \times n}$ and consider the map $g : O(n) \to M(H_0)$, $U \mapsto U^T H_0 U$, which is a smooth surjection. If U_k is a solution to (31) it follows that

$$g(U_{k+1}) = e^{-\alpha_k [g(U_k), N]} g(U_k) e^{\alpha_k [g(U_k), N]},$$

which generates the Lie-bracket algorithm (4). Thus, g maps the associated orthogonal Lie-bracket algorithm with initial condition U_0 to the Lie-bracket algorithm with initial condition $U_0^T H_0 U_0$ on $M(U_0^T H_0 U_0) = M(H_0)$.

Remark 6.1. Consider the potential function $\phi : O(n) \to \mathbb{R}_+$, $\phi(U) = ||U^T H_0 U - N||^2$ on the set of orthogonal $n \times n$ matrices. Using the standard induced Riemannian metric from $\mathbb{R}^{n \times n}$ on $O(n)$, the *associated orthogonal gradient flow* is [2], [3], [5], [9]

$$\dot{U} = -\nabla \phi(U) = U[U^T H_0 U, N].$$

THEOREM 6.1. *Let $H_0 = H_0^T$ be a real symmetric $n \times n$ matrix that satisfies Condition 6.1. Let $N \in \mathbb{R}^{n \times n}$ satisfy Condition 2.2, and let α_N be either the constant step-size selection (13) or the variable step-size selection (17). The recursion*

$$\begin{aligned} U_{k+1} &= U_k e^{\alpha_k [U_k^T H_0 U_k, N]}, \qquad U_0 \in O(n), \\ \alpha_k &= \alpha_N(H_k) \end{aligned}$$

referred to as the associated orthogonal Lie-bracket algorithm *has the following properties:*

(i) *A solution U_k, $k = 1, 2, \ldots$, to the associated orthogonal Lie-bracket algorithm remains orthogonal.*

(ii) *Let $\phi(U) = ||U^T H_0 U - N||^2$ be a map from $O(n)$ to the set of nonnegative reals $\mathbb{R}_+$. Let U_k, $k = 1, 2, \ldots$, be a solution to the associated orthogonal Lie-bracket algorithm. Then $\phi(U_k)$ is strictly monotonically decreasing for every $k \in \mathbb{N}$ where $[U_k^T H_0 U_k, N] \neq 0$.*

(iii) *Fixed points of the algorithm are characterised by matrices $U \in O(n)$ such that*

$$[U^T H_0 U, N] = 0.$$

There are exactly $2^n n!$ distinct fixed points.

(iv) *Let U_k, $k = 1, 2, \ldots$, be a solution to the associated orthogonal Lie-bracket algorithm, then U_k converges to an orthogonal matrix U_∞, a fixed point of the algorithm.*

(v) *All fixed points of the associated orthogonal Lie-bracket algorithm are strictly unstable except those 2^n points $U_* \in O(n)$ such that*

$$U_*^T H_0 U_* = \Lambda,$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Such points U_* are locally exponentially asymptotically stable and $H_0 = U_* \Lambda U_*^T$ is an eigenspace decomposition of H_0.

Proof. Part (i) follows directly from the orthogonal nature of $e^{\alpha_k [U_k^T H_0 U_k, N]}$. Note that in part (ii) the definition of ϕ can be expressed in terms of the map $g(U) = U^T H_0 U$ from $O(n)$ to $M(H_0)$ and the Lie-bracket potential $\psi(H) = \|H - N\|^2$ of (1), i.e.,

$$\phi(U_k) = \psi(g(U_k)).$$

Observe that $g(U_0) = U_0^T H_0 U_0$ and thus $g(U_k)$ is the solution of the Lie-bracket algorithm with initial condition $U_0^T H_0 U_0$. As the step-size selection scheme α_N is either (13) or (17), then $g(U_k)$ satisfies (7). This ensures that part (ii) holds.

If U_k is a fixed point of the associated orthogonal Lie-bracket algorithm with initial condition $U_0^T H_0 U_0$, then $g(U_k)$ is a fixed point of the Lie-bracket algorithm. Thus, from Proposition 2.1, $[g(U_k), N] = [U_k^T H_0 U_k, N] = 0$. Moreover, if $[U_k^T H_0 U_k, N] = 0$ for some given $k \in \mathbb{N}$, then by inspection $U_{k+l} = U_k$ for $l = 1, 2, \ldots$, and U_k is a fixed point of the associated orthogonal Lie-bracket algorithm. From Lemma 2.2 it follows that if U is a fixed point of the algorithm then $U^T H_0 U = \pi^T \Lambda \pi$ for some permutation matrix π. By inspection any orthogonal matrix $W = SU\pi^T$, where S is a sign matrix $S = \mathrm{diag}(\pm 1, \ldots, \pm 1)$, is also a fixed point of the recursion, and indeed, any two fixed points are related in this manner. A simple counting argument shows that there are exactly $2^n n!$ distinct matrices of this form.

To prove (iv), note that since $g(U_k)$ is a solution to the Lie-bracket algorithm, it converges to a limit point $H_\infty \in M(H_0)$, $[H_\infty, N] = 0$ (Proposition 2.1). Thus U_k must converge to the preimage set of H_∞ via the map g. Condition 6.1 ensures that a set generated by the preimage of H_∞ is a finite distinct set, any two elements U_∞^1 and U_∞^2 of which are related by $U_\infty^1 = U_\infty^2 S$, $S = \mathrm{diag}(\pm 1, \ldots, \pm 1)$. Convergence to a particular element of this preimage follows since $\alpha_k [U_k^T H_0 U_k, N] \to 0$ as in Proposition 2.1.

To prove part (v), observe that the dimension of $O(n)$ is the same as the dimension of $M(H_0)$ due to genericity Condition 6.1. Thus g is locally a diffeomorphism on $O(n)$ that forms an exact equivalence between the Lie-bracket algorithm and the associated orthogonal Lie-bracket algorithm. Restricting g to a local region, the stability structure of equilibria are preserved under the map g^{-1}. Thus, all fixed points of the associated orthogonal Lie-bracket algorithm are locally unstable except those that map via g to the unique locally asymptotically stable equilibrium of the Lie-bracket recursion. Observe that due to the monotonicity of $\phi(U_k)$ a locally unstable equilibrium is also globally unstable.   □

THEOREM 6.2. *Let $H_0 \in \mathbb{R}^{m \times n}$ where $m \geq n$ satisfies Condition 6.1. Let $N \in \mathbb{R}^{m \times n}$ satisfy Condition 5.1. Let the time-step α_k be given by*

$$\alpha_k = \alpha_{\widehat{N}}(\widehat{H}),$$

where $\alpha_{\widehat{N}}$ is either the constant step-size selection (13) *or the variable step-size selection scheme* (17), *on $M(\widehat{H}_0)$. The recursion*

$$V_{k+1} = V_k e^{\alpha_k \{U_k^T H_0^T V_k, N^T\}}, \qquad V_0 \in O(m),$$
$$U_{k+1} = U_k e^{\alpha_k \{V_k^T H_0 U_k, N\}}, \qquad U_0 \in O(n),$$

referred to as the associated orthogonal singular value algorithm, *has the following properties*:

(i) *Let (V_k, U_k) be a solution to the associated orthogonal singular value algorithm, then both V_k and U_k remain orthogonal.*

(ii) *Let $\phi(V, U) = \|V^T H_0 U - N\|^2$ be a map from $O(m) \times O(n)$ to the set of nonnegative reals $\mathbb{R}_+$, then $\phi(V_k, U_k)$ is strictly monotonically decreasing for every $k \in \mathbb{N}$ where $\{V_k^T H_0 U_k, N\} \neq 0$ and $\{U_k^T H_0^T V_k, N^T\} \neq 0$. Moreover, fixed points of the algorithm are characterised by matrix pairs $(V, U) \in O(m) \times O(n)$ such that*

$$\{V^T H_0 U, N\} = 0 \quad and \quad \{U^T H_0^T V, N^T\} = 0.$$

(iii) *Let (V_k, U_k), $k = 1, 2, \ldots$, be a solution to the associated orthogonal singular value algorithm, then (V_k, U_k) converges to a pair of orthogonal matrices (V_∞, U_∞), a fixed point of the algorithm.*

(iv) *All fixed points of the associated orthogonal singular value algorithm are strictly unstable except those points $(V_*, U_*) \in O(m) \times O(n)$ such that*

$$V_*^T H_0 U_* = \Sigma,$$

where $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n}$. Each such point (V_*, U_*) is locally exponentially asymptotically stable and $H_0 = V_*^T \Sigma U_*$ is a singular value decomposition of H_0.

Proof. The proof of this theorem is analogous to the proof of Theorem 6.1. □

7. Computational considerations. There are several issues involved in the implementation of the Lie-bracket algorithm as a numerical tool that have not been dealt with in the body of this paper. Design and implementation of efficient code has not been considered and would depend heavily on the nature of the hardware on which such a recursion would be run. As each iteration requires the calculation of a time-step, an exponential and a $k + 1$ estimate, it is likely that it would be best to consider applications in parallel processing environments. Certainly in a standard computational environment the exponential calculation would limit the possible areas of useful application of the algorithms proposed.

It is also possible to consider approximations of the Lie-bracket algorithm that have good computational properties. For example, consider a (1,1) Padé approximation to the matrix exponential

$$e^{\alpha_k [H_k, N]} \approx \frac{2I + \alpha_k [H_k, N]}{2I - \alpha_k [H_k, N]}.$$

Such an approach has the advantage that, as $[H_k, N]$ is skew-symmetric, the Padé approximation will be orthogonal and will preserve the isospectral nature of the Lie-bracket algorithm. Similarly, an (n, n) Padé approximation of the exponential for any n will also be orthogonal. There are difficulties involved in obtaining direct step-size selection schemes based on the Padé approximate Lie-bracket algorithms. To guarantee that the potential ψ is monotonic decreasing for such schemes, direct

estimates of time-step must be chosen prohibitively small. A good heuristic choice of a step-size selection scheme, however, can be made based on the selection schemes given in this paper and simulations indicate that such an approach is viable.

Another approach is to take just the linear term from the Taylor expansion of $H_{k+1}(\alpha_k)$,

$$H_{k+1} \approx H_k + \alpha_k[H_k, [H_k, N]],$$

as an algorithm on $\mathbb{R}^{n \times n}$. An algorithm such as this is similar in form to approximating the curves generated by the Lie-bracket algorithm by straight lines. The approximation will not retain the isospectral nature of the Lie-bracket recursion; however, it is computationally cheap. Furthermore, when the curvature of the manifold $M(H_0)$ is small, then it can be imagined that the linear algorithm would be a good approximation to the Lie-bracket algorithm.

8. Conclusion. In this paper we have proposed two algorithms which, along with their associated orthogonal algorithms, calculate respectively, the eigenvalue decomposition of a symmetric matrix and the singular value decomposition of a general matrix. Moreover, we have presented two suitable step-size selection schemes which ensure that, for generic initial conditions, the algorithms proposed will converge exponentially fast to an asymptotically attractive fixed point.

In future work we hope to improve the theoretical understanding of the step-size selection schemes necessary for the Lie-bracket algorithm as well as to investigate a number of related applications of the double-bracket flow and its discretisation.

9. Appendix. The following discussion is a proof of Proposition 4.3.

Proof. By Lemma 4.1, Λ is the unique locally asymptotically stable equilibrium point and it remains to show that Λ is exponentially attractive. Note that direct linearisation techniques do not apply as the recursion will not necessarily be differentiable at the equilibrium Λ. To proceed we set $c = 1/(4||H_0|| \cdot ||N||)$, the constant time-step, and show that the Lie-bracket algorithm converges faster using the variable step-size selection scheme than it does with the constant time-step c. The proof is divided into a number of lemmas.

LEMMA 9.1. *Let $0 < \beta < \min(1, c)$, where $c = 1/(4||H_0|| \cdot ||N||)$. Then there exists a real number δ_1 such that for $H_k \in M(H_0)$ and $||[H_k, N]|| < \delta_1$, then*

$$(33) \qquad\qquad 0 > \triangle\psi(H_k, \beta) \geq -3\beta||[H_k, N]||^2.$$

Proof. Consider the error term $\tau^2 tr(N\mathcal{R}_2(\tau))$ defined in Lemma 3.1 and recall the estimation argument for Lemma 3.3. Employing a similar argument for $\tau = \beta$ gives

$$|\beta^2 tr(N\mathcal{R}_2(\beta))| \leq ||H_0|| \cdot ||N|| \cdot \left(e^{2\beta||[H_k, N]||} - 1 - 2\beta||[H_k, N]||\right).$$

Thus, combining this with (9) it follows that

$$
\begin{aligned}
\triangle\psi(H_k, \beta) &\geq -2\beta||[H_k, N]||^2 - 2\beta^2|tr(N\mathcal{R}_2(\beta))| \\
(34) \qquad &\geq -2\beta||[H_k, N]||^2 - 2||H_0|| \cdot ||N|| \cdot \\
&\qquad \left(e^{2\beta||[H_k, N]||} - 1 - 2\beta||[H_k, N]||\right).
\end{aligned}
$$

It is well known that

$$2(e^y - 1 - y) \sim y^2 \quad \text{for } y \to 0^+,$$

where "$\sim$" indicates that two functions are asymptotically equal. This is equivalent to saying that for any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$, such that for all y, where $\delta(\epsilon) > y > 0$, then $1 - \epsilon < 2(e^y - 1 - y)/y^2 < 1 + \epsilon$. Thus, choosing $\epsilon = \frac{1}{4}$, it follows that for $\delta(\frac{1}{4}) > y > 0$ then $2(e^y - 1 - y) < 2y^2$. Recall that we are restricting $\beta < 1$, and thus, there exists some real number $\delta_1 > 0$ such that if $||[H_k, N]|| < \delta_1$, then $2\beta||[H_k, N]|| < \delta(\frac{1}{4})$, and hence $2(e^{2\beta||[H_k, N]||} - 1 - 2\beta||[H_k, N]||) < 4\beta^2||[H_k, N]||^2$. Substituting this into (34) gives

$$(35) \qquad \triangle\psi(H_k, \beta) \geq -2\beta||[H_k, N]||^2 - 4\beta^2||H_0|| \cdot ||N|| \cdot ||[H_k, N]||^2.$$

By additionally requiring that $\beta < c = 1/(4||H_0|| \cdot ||N||)$ the lemma is proved. □

LEMMA 9.2. *Let α_N^* be the step-size selection scheme given by Lemma 3.4, and let $\gamma \in \mathbb{R}_+$, such that $\alpha_N^*(H_k) > \gamma > 0$ for all $[H_k, N] \neq 0$. Define $\overline{\gamma} := \min\{\gamma, c\}$ and choose $\beta \in \mathbb{R}_+$ such that*

$$0 < \beta < \min\left\{1, c, \frac{2}{3}(\overline{\gamma} - 2||H_0|| \cdot ||N||\overline{\gamma}^2)\right\}.$$

Then there exists a real number $\delta_2 > 0$ such that for any $H_k \in M(H_0)$ with $||[H_k, N]|| < \delta_2$

$$(36) \qquad -3\beta||[H_k, N]||^2 > \triangle\psi_U^*(H_k, \alpha_k^*).$$

Proof. Recall that α_k^* was chosen as the first critical point of the function $\triangle\psi_U^*(H_k, \tau)$. Thus $\triangle\psi_U^*(H_k, \tau)$ is monotonic decreasing on the interval $(0, \alpha_k^*)$. The lower bound $\overline{\gamma} < \gamma$, for α_N^*, must be less than α_k^*, and thus $\triangle\psi_U^*(H_k, \overline{\gamma}) > \triangle\psi_U^*(H_k, \alpha_k^*)$. Substituting $\overline{\gamma}$ into the definition of $\triangle\psi_U^*$ gives

$$\begin{aligned}\triangle\psi_U^*(H_k, \overline{\gamma}) &= -2\overline{\gamma}||[H_k, N]||^2 \\ &\quad + \frac{||H_0|| \cdot ||[N, [H_k, N]]||}{||[H_k, N]||}\left(e^{2\overline{\gamma}||[H_k, N]||} - 1 - 2\overline{\gamma}||[H_k, N]||\right), \\ &\leq -2\overline{\gamma}||[H_k, N]||^2 \\ &\quad + 2||H_0|| \cdot ||N||\left(e^{2\overline{\gamma}||[H_k, N]||} - 1 - 2\overline{\gamma}||[H_k, N]||\right).\end{aligned}$$

As shown in Lemma 9.1, there exists $\delta_2 > 0$, such that for any $H_k \in M(H_0)$, where $||[H_k, N]|| < \delta_2$, then $2\left(e^{2\gamma||[H_k, N]||} - 1 - 2\gamma||[H_k, N]||\right) < 4\gamma^2||[H_k, N]||^2$. Using this with the above inequality gives

$$\triangle\psi_U^*(H_k, \overline{\gamma}) \leq 2||[H_k, N]||^2\left(2||H_0|| \cdot ||N||\overline{\gamma}^2 - \overline{\gamma}\right).$$

Note that since $\overline{\gamma} < c$, then the right-hand side of the last inequality is strictly negative. Now as

$$0 < \beta < \frac{2}{3}\left(\overline{\gamma} - 2||H_0|| \cdot ||N||\overline{\gamma}^2\right),$$

then $-3\beta||[H_k, N]||^2 > 2||[H_k, N]||^2\left(2||H_0|| \cdot ||N||\overline{\gamma}^2 - \overline{\gamma}\right)$ and the result follows. □

The proof of Proposition 4.3 now follows by choosing

$$(37) \qquad \beta = \min\begin{cases} 1, \\ c, \\ \frac{2}{3}\left(\overline{\gamma} - 2||H_0|| \cdot ||N||\overline{\gamma}^2\right), \end{cases}$$

where $\overline{\gamma} = \min(\gamma, c)$. Thus, from Lemmas 9.1 and 9.2, choose δ_1 and δ_2 such that the results hold and set $\delta = \frac{1}{2} \min\{\delta_1, \delta_2\}$. Hence, combining the inequalities (33) and (36) gives

$$(38) \qquad\qquad \triangle\psi(H_k, \alpha_k^*) < \triangle\psi(H_k, \beta) < 0$$

for all $H_k \in M(H_0)$ with $\|[H_k, N]\| < \delta$.

Let D_δ be some open set around Λ such that $\|[H_k, N]\| < \delta$. Note that $\beta \le c$, and thus from Lemma 4.2 the Lie-bracket algorithm equipped with $\alpha_N^\beta = \beta$ as a step-size selection scheme is exponentially stable. Finally, note that within D_δ, and due to (38), $\psi(H_{k+1}(\alpha_k^*))$ will always decrease faster than $\psi(H_{k+1}(\beta))$, regardless of H_k. Since Λ is exponentially attractive for the Lie-bracket algorithm equipped with the selection scheme α_N^β, it follows that Λ must also be exponentially attractive for the same recursion equipped with the selection scheme α_N^*. $\square$

Acknowledgments. The authors would like to thank Kenneth Driessel and Wei-Yong Yan for many useful comments. Also the authors thank an anonymous referee for mentioning the connection of the double-bracket flow to micromagnetics as well as a number of useful comments.

REFERENCES

[1] A. M. BLOCH, R. W. BROCKETT, AND T. RATIU, *A new formulation of the generalised Toda lattice equations and their fixed point analysis via the momentum map*, Bull. Amer. Math. Soc., 23 (1990), pp. 477–485.

[2] R. W. BROCKETT, *Dynamical systems that sort lists, diagonalise matrices and solve linear programming problems*, Linear Algebra Appl., 146 (1991), pp. 79–91; see, also, Proc. IEEE Conf. Decisions and Control, 1988, pp. 799–803.

[3] M. T. CHU, *The generalized Toda flow, the QR-algorithm and the center manifold theory*, SIAM J. Discrete Math., 5 (1984), pp. 187–201.

[4] ———, *A differential equation approach to the singular value decomposition of bidiagonal matrices*, Linear Algebra Appl., 80 (1986), pp. 71–80.

[5] M. T. CHU AND K. R. DRIESSEL, *The projected gradient method for least squares matrix approximations with spectral constraints*, SIAM J. Numer. Anal., 27 (1990), pp. 1050–1060.

[6] P. DEIFT, T. NANDA, AND C. TOMEI, *Ordinary differential equations for the symmetric eigenvalue problem*, SIAM J. Numer. Anal., 20 (1983), pp. 1–22.

[7] K. R. DRIESSEL, *On isospectral gradient flows–solving matrix eigenproblems using differential equations*, in Inverse Problems, J. R. Cannon and U. Hornung, eds., Birkhauser-Verlag, 1986, pp. 69–90.

[8] U. HELMKE AND J. B. MOORE, *Singular value decomposition via gradient flows*, Systems Control Lett., 14 (1990), pp. 369–377.

[9] ———, *Optimization and dynamical systems*, in Communications and Control Engineering, Springer-Verlag, London, 1994.

[10] J. IMAE, J. PERKINS, AND J. MOORE, *Towards time varying balanced realisation via Riccati equations*, Math. Control Signals Systems, 5 (1992), pp. 313–326.

[11] J. E. PERKINS, U. HELMKE, AND J. B. MOORE, *Balanced realizations via gradient flow techniques*, Systems Control Lett., 14 (1990), pp. 369–380.

[12] S. T. SMITH, *Dynamical systems that perform the singular value decomposition*, Systems Control Lett., 16 (1991), pp. 319–327.

[13] W. W. SYMES, *The QR algorithm and scattering for the finite nonperiodic Toda lattice*, Phys. D4, (1982), pp. 275–280.

[14] D. WATKINS AND L. ELSNER, *Self-equivalent flows associated with the singular value decomposition*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 244–258.