Noname manuscript No.
(will be inserted by the editor)

Sheng-Jhih Wu · Chii-Ruey Hwang · Moody T. Chu

Attaining the Optimal Gaussian Diffusion Acceleration

Received: date / Accepted: date

Abstract Sampling from probability distributions in high dimensional spaces is generally impractical. Diffusion processes with invariant equilibrium distributions can be used as a means to generate approximations. An important task in such an endeavor is to design an equilibrium-preserving drift to accelerate the convergence. Starting from a reversible diffusion, it is desirable to depart for non-reversible dynamics via a perturbed drift so that the convergence rate is maximized with the common equilibrium. In the Gaussian diffusion acceleration, this problem can be cast as perturbing the inverse of a given covariance matrix by skew-symmetric matrices so that all resulting eigenvalues have identical real part. This paper describes two approaches to obtain the optimal rate of Gaussian diffusion. The asymptotical approach works universally for arbitrary Ornstein-Uhlenbeck processes, whereas the direct approach can be implemented as a fast divide-and-conquer algorithm. A comparison with recently proposed Lelièvre-Nier-Pavliotis algorithm is made.

This research was supported in part by the National Science Foundation under grant DMS-1014666.

Sheng-Jhih Wu
Institute of Mathematics, Academia Sinica, 6F, Astronomy-Mathematics Building, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan.
E-mail: sjw@math.sinica.edu.tw

Chii-Ruey Hwang
Institute of Mathematics, Academia Sinica, 6F, Astronomy-Mathematics Building, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan.
E-mail: crhwang@sinica.edu.tw

Moody T. Chu
Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205, USA.
E-mail: chu@math.ncsu.edu

Keywords Diffusion acceleration · Convergence to equilibrium · Ornstein-Uhlenbeck process · t-circulant matrices · Inverse eigenvalue problem · Recursive algorithm

1 Introduction

1.1 Motivation

When dealing with applications in areas that involve uncertainties, it often is necessary to collect samples subject to some probability distributions. Taking samples directly from probability distributions in high-dimensional spaces or large finite sets is usually impractical, especially when the underlying density functions are not completely known. One possible way to circumvent this difficulty is to approximate the underlying distributions via diffusion processes whose equilibrium distributions are the desired distributions. For practical reasons, it is further preferred that among all possible diffusions we can realize the equilibrium quickly. Accelerating the rate of convergence therefore is an important issue.

How good the approximation and acceleration are depends on the employed diffusion processes (or more generally, the Markov processes/chains) and the measurement criteria. In a series of research [15,16], some comparison criteria to gauge the convergence rate of diffusions have been proposed. The criteria for evaluation include the spectral gap of the infinitesimal generator, the convergence rate of transition probability to equilibrium in the variational norm, etc.. As for the discrete case, when the sample space is a large finite set, one often resorts to Markov chain Monte Carlo. In a similar spirit, there are other criteria developed in the literature for assessing the performance of Monte Carlo algorithms, such as geometric ergodicity, asymptotic variance, the second largest eigenvalue in absolute value, and so on.

The use of diffusions for sampling has useful applications in image analysis. The issue of the rate of convergence to equilibrium for various families of Markov processes/chains has been an extensively studied subject. Related works may be found in [2,3,4,5,12,14,19,20,23,25,26,27,28], and the many references contained therein. Many of the theoretical results can be regarded as conceptual algorithms having potential to advance new methods for Monte Carlo simulation with better performance. In statistical physics, considerable progress has been achieved for the problem of trend to equilibrium for various classes of diffusion equations. We refer the reader to [21,31] and many references therein.

Of particular interest is the scheme where the original reversible system is perturbed by introducing proper antisymmetric dynamics so that the convergence is accelerated. As for discrete case, given a reversible Markov chain with equilibrium distribution π, i.e., given a transition matrix P that is symmetric with respect to the inner product weighted by π, one seeks for a suitable skew-symmetric transition matrix Q added to P such that the resulting Markov chain generated by the transition matrix $P+Q$ to which π is invariant and possesses convergence rate as fast as possible, say, a smaller

asymptotic variance. See for example [13,6,7]. In a similar spirit, for continuous case, one's task is to add a proper antisymmetric perturbation drift to the given reversible diffusion such that the resulting non-reversible diffusion maintains the same equilibrium π and realizes π as fast as possible in terms of some prescribed criterion. For reference, see for example [12,15,17,16,26].

1.2 Principle of acceleration

We now briefly introduce the notion of non-reversible acceleration for diffusion processes described in [15,16]. Let $U : \mathbb{R}^n \to \mathbb{R}$ be a smooth potential satisfying the condition that $\int_{\mathbb{R}^n} e^{-U(x)}\, dx < \infty$. Suppose that the probability distribution π that we are interested in sampling from has the density

$$\rho(x) = \frac{1}{Z} e^{-U(x)}, \tag{1}$$

where the normalizing constant $Z = \int_{\mathbb{R}^n} e^{-U(x)}\, dx$ is the partition function.

Let $W(t)$ be a standard Brownian motion in $\mathbb{R}^n$. A commonly used reversible diffusion for approximating $\pi(x)$ is of the form

$$dX(t) = -\nabla U(X(t))dt + \sqrt{2}dW(t), \qquad t > 0, \quad X(0) = x_0, \tag{2}$$

which is known as a gradient diffusion. Let $p_t(x)$ be the probability density function of the solution $X(t)$ to (2) at time t. Then $p_t(x)$ satisfies the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \nabla \cdot (\nabla U p) + \Delta p. \tag{3}$$

A discussion concerning the exponential convergence of the density function p_t to the equilibrium π through the spectral gap of the Fokker-Planck equation (3) appearing in a Poincaré inequality can be found in [20]. Instead we discuss the convergence through the infinitesimal generator of the diffusion herein. Performing a change of variable to the unknown of the Fokker-Planck equation (3) via $p_t(x) = q_t(x)\rho(x)$, it can be shown that the function $q_t(x)$ is a solution of the backward Kolmogorov equation

$$\frac{\partial q}{\partial t} = -\nabla U \cdot \nabla q + \Delta q. \tag{4}$$

Let $L^2(\pi)$ denote the L^2 space weighted by π. With a suitable domain, the differential operator appearing on the right-hand side of (4) is the infinitesimal generator of the gradient diffusion (2). More precisely,

$$\mathcal{L}_0 f = -\nabla U \cdot \nabla f + \Delta f, \tag{5}$$

where $f \in \mathcal{D}(\mathcal{L}_0)$, the domain of $\mathcal{L}_0$, which is a subset of $L^2(\pi)$. The subscript 0 is used to denote that no additional drift is added to the diffusion (2). It is standard that the operator $\mathcal{L}_0$ is self-adjoint on $L^2(\pi)$. This is equivalent to the fact that the gradient diffusion (2) is reversible.

Since the gradient diffusion is one of the diffusions with the desired equilibrium, a natural extension is to consider a family of diffusions with the

common desired equilibrium. One would then seek for a diffusion that possesses a better rate of convergence within the family measured by certain prescribed criterion. In particular, one would be interested in searching the best diffusion (if it exists) within the family. Since the gradient diffusion is the only reversible diffusion, it is amount to finding a better (best) non-reversible dynamics while keeping the equilibrium.

We now are more specific about the non-reversible modification. So long as π remains as the common equilibrium, consider a family of diffusions

$$dX(t) = -\nabla U(X(t))dt + C(X(t))dt + \sqrt{2}dW(t), \qquad t > 0, \quad X(0) = x_0, \tag{6}$$

where the vector field C satisfies $\nabla\cdot(Ce^{-U}) = 0$, i.e., C is weighted divergence-free with respect to π. Indeed, in order to preserve π, the new drift is necessarily of the additive form [15, Proposition 2.2]. The infinitesimal generator of the diffusion (6) is

$$\mathcal{L}_c = -\nabla U \cdot \nabla + C \cdot \nabla + \Delta. \tag{7}$$

That is, $\mathcal{L}_c f = (-\nabla U + C) \cdot \nabla f + \Delta f = \mathcal{L} f + C \cdot \nabla f$. The $L^2(\pi)$-adjoint operator is

$$\mathcal{L}_c^* f = -\nabla \cdot ((-\nabla U + C)f) + \Delta f. \tag{8}$$

It is not hard to see that $\mathcal{L}_c$ is not self-adjoint if C is not zero. Hence, the diffusion (6) is non-reversible. The requirement that C is weighted divergence-free comes from the need for preserving π as the equilibrium. It is known that for a diffusion process if there is no explosion in the diffusion and if $\mathcal{L}^*\mu = 0$ for some measure μ, then μ is an invariant measure. In our setting, $\mathcal{L}_c^* e^{-U}$ has to be zero to keep the equilibrium π. Therefore, since $\mathcal{L}_0^* e^{-U} = 0$, from the relation between $\mathcal{L}_0^*$ and $\mathcal{L}_c^*$, one can easily check that the condition that $\nabla \cdot (Ce^{-U}) = 0$ (with no explosion in (6)) guarantees the equilibrium π for the diffusion (6).

We now describe the spectral gap as the comparison criterion. Given a Markov process with stationary distribution π and generator $\mathcal{L}$ on $L^2(\pi)$, let $\{P(t)\}_{t\geq 0}$ denote the associate strongly continuous Markovian contraction semigroup. Let $p(t, x, y)$ be the transition probability staring at x and taking y at time t. Then $P(t)$ acting on suitable function f evaluated at x is given by

$$(P(t)f)(x) = E_x[f(X(t))] = \int p(t, x, y) f(y)\, dy. \tag{9}$$

It is standard that $(P(t)f)(x)$ converges to $\pi(f) := \int f(x)\, \pi(dx)$ as t goes to infinity for any starting point x. Suppose that $P(t)$ converges exponentially on $L^2(\pi)$. That is, suppose that there is some constant $d > 0$ and $\lambda < 0$ such that for all $f \in L^2(\pi)$,

$$\int [(P(t)f)(x) - \pi(f)]^2\, \pi(dx) \leq d\, ||f - \pi(f)||^2 e^{2\lambda t} \quad \text{for all } t \geq 0. \tag{10}$$

It is clear that the constant function $\mathbf{1} \in \mathcal{D}(\mathcal{L})$ is an eigenfunction of $\mathcal{L}$ with eigenvalue 0. In fact, we can ignore the zero eigenvalue by restricting $\mathcal{D}(\mathcal{L})$ to

the orthogonal complement space of constant functions $\mathcal{H} := \{f \in \mathcal{D}(\mathcal{L}) \subset L^2(\pi) \,|\, \pi(f) = 0\}$. Under this restriction, one can obtain from (10) that

$$||P(t)|| \leq b\, e^{\lambda t} \quad \text{for all } t \geq 0, \tag{11}$$

for some constant $b > 0$. This implies that the spectral radius of $P(1)$ is a candidate for measuring the rate of convergence to equilibrium. Moreover, since $||P(t)||$ decays faster than a polynomial rate, the weak spectral mapping theorem holds. That is, $\sigma(P(t)) = \overline{exp(t\sigma(\mathcal{L}))}$, where $\sigma(\cdot)$ denotes the spectrum. This leads to the spectral gap of $\mathcal{L}_c$ defined in [16]:

$$\lambda(C) := sup\{\text{real part of } u \,|\, u \in \sigma(\mathcal{L}_c) \text{ and } u \neq 0\} \tag{12}$$

since it can gauge the closeness of the distribution of $X(t)$ and the equilibrium π at large time t. It is shown in [16, Theorem 1] that $\lambda(C) \leq \lambda(0)$. This means that any non-reversible diffusion (6) is at least as good as the reversible diffusion (2). Moreover, it is shown that the equality holds only in a rare situation. Consequently, the antisymmetrically perturbed dynamics (6) has a better rate of convergence . In other words, adding an antisymmetric operator $C \cdot \nabla$, where C is weighted divergence-free, accelerates the convergence to equilibrium.

1.3 Preliminaries for Gaussian case

In [15] the scheme of acceleration mainly focuses on Gaussian case. More precisely, the resulting drift term $-\nabla U(X(t)) + C(X(t))$ of the diffusion process is linear. That is, the extra drift C is linear and the potential U is quadratic. Starting from the gradient diffusion (2), consider the case when

$$U(x) = \frac{1}{2}\langle -Dx, x\rangle, \tag{13}$$

where D is a negative-definite real matrix. Obviously, $\nabla U(x) = -Dx$. Then it is the Ornstein-Uhlenbeck process with drift coefficient matrix D and diffusion coefficient matrix being identity multiplied by a constant:

$$dX(t) = DX(t)dt + \sqrt{2}dW(t), \qquad t > 0, \quad X(0) = x_0. \tag{14}$$

It is reversible with Gaussian equilibrium π and density $\rho(x) = \frac{1}{Z}exp(\frac{1}{2}\langle Dx, x\rangle)$. Alternatively, one may consider to depart from the reversible dynamics by imposing a different linear drift so long as π remains as its equilibrium. This leads to the following Ornstein-Uhlenbeck process

$$dX(t) = BX(t)dt + \sqrt{2}dW(t), \qquad t > 0, \quad X(0) = x_0, \tag{15}$$

where B is a stability matrix. Then it is shown in [15, Theorem 3.1] that the process (15) has π as its equilibrium if and only if $B = D+C$ for some matrix C satisfying the relation $\langle Cx, Dx\rangle = 0$ for all $x \in \mathbb{R}^n$. It can be argued easily that C can be written as $C = SD$ with S being some skew-symmetric matrix.

For a given matrix $A \in \mathbb{R}^{n\times n}$, let $\Re\lambda_M(A)$ and $\Re\lambda_m(A)$ denote the maximum and minimum real parts of eigenvalues of A, respectively. Then it can be argued that the convergence rate of the covariance of $X(t)$ is given by $\exp(2\Re\lambda_M(B)t)$ [15, Theorem 3.2]. If this quantity is used as the criterion to qualify a better approximation, we naturally would prefer a smaller $\Re\lambda_M(B)$ for faster convergence. It is not difficult to establish the inequalities [15, Theorem 3.3]

$$\Re\lambda_m(D) \le \Re\lambda_m(B) \le \Re\lambda_M(B) \le \Re\lambda_M(D), \tag{16}$$

showing that, among all drifts with equilibrium π, the diffusion with drift D has the worst rate. In contrast, improvement can be made with drift $B = D+SD$ by any skew-symmetric S. In fact, the perturbed diffusion by adding any drift C orthogonal to D is generally at least as good as the original diffusion and is guaranteed to improve the rate if and only if eigenvalues of D are not identical.

An important result established earlier in [15] is that

$$\inf\{\Re\lambda_M(D+SD)\,|\,S \text{ is skew-symmetric and } SD \perp D\} = \frac{\mathrm{trace}(D)}{n}. \tag{17}$$

Since $\Re\lambda_M(D+SD)$ measures the convergence rate under the perturbation S, the infimum over all possible skew-symmetrically perturbation, namely, the ratio $\frac{\mathrm{trace}(D)}{n}$, is referred to as the *optimal rate* of convergence. In addition, an earlier work has been made in [15, Theorem 4.1] to construct a special matrix $\hat{C}$, depending on D, such that

$$\inf_{\alpha>0}\{\Re\lambda_M(D+\alpha\hat{C})\} = \frac{\mathrm{trace}(D)}{n}. \tag{18}$$

A similar idea has also been exploited in [12] to study the limiting behavior of the spectral gap of growing divergence-free perturbations to the Laplace-Beltrami operator on a compact Riemannian manifold.

1.4 Outline of the present paper

The present paper is mainly motivated by [15] where our goal is to search for the skew-symmetric perturbation that gives rise to the optimal rate of Gaussian acceleration. In fact, such a problem can be cast as an inverse eigenvalue problem as discussed below. It is worth mentioning that the spectrum of the Ornstein-Uhlenbeck operator in the weighted $L^p(\mu)$ spaces, where $1 \le p < \infty$ and μ is an invariant measure, has been comprehensively studied in [24]. It is proven there that the spectrum of the infinitesimal generator of the Ornstein-Uhlenbeck process (15),

$$\mathcal{L}_{ou} = Bx\cdot\nabla + \Delta, \tag{19}$$

on $L^p(\pi)$, where $1 < p < \infty$, is fully determined by the eigenvalues of the drift coefficient matrix B. More specifically,

$$\sigma(\mathcal{L}_{ou}) = \left\{ \sum_{i=1}^{r} n_i\lambda_i \,\middle|\, n_i \in \mathbb{N} \right\}, \tag{20}$$

where $\lambda_1, \cdots, \lambda_r$ are the (distinct) eigenvalues of B. Consequently, the spectral gap of $\mathcal{L}_{ou}$ is $\Re\lambda_M(B)$. This shows that the criterion employed in [15] is exactly the spectral gap, a natural comparison criterion for general diffusion processes. If one considers the spectral gap as the measurement in the family of general diffusions with common desired equilibrium, then the problem of optimal perturbation for fast convergence in the Gaussian case can be cast as an inverse eigenvalue problem and can be tackled by matrix analysis. Motivated by these efforts, we study the following two problems in this paper.

Problem 1. Find a unified strategy for constructing a skew-symmetric matrix S, independent of D, so that the optimal rate is achieved asymptotically in the sense that the limiting behavior

$$\lim_{\alpha\to\infty} \Re\lambda_M(D + \alpha SD) = \frac{\mathrm{trace}(D)}{n}. \tag{21}$$

Problem 2. For each given D, and a skew-symmetric matrix S, dependent on D, so that the optimal rate is attained in the sense that

$$\Re\lambda_M(D + SD) = \frac{\mathrm{trace}(D)}{n}. \tag{22}$$

Obviously, the first problem differs from the second problem in two aspects. First, the skew-symmetric matrix S for Problem 1 is to be universal for all negative-definite matrices D, whereas the skew-symmetric S for Problem 2 varies as D. Second, Problem 1 concerns an asymptotic convergence to the optimal rate, whereas Problems 2 concerns a direct hit for the optimal rate.

In a certain sense, the parameter α in the first problem can be thought of as the magnitude of perturbation. Unlike the conventional notion where the perturbation is expected to be minuscule, there are practical reasons that the perturbed spectral information as the magnitude of perturbation approaches to infinity is of interest. It is tempting to think that a stronger perturbation should bring forth a better rate of convergence, but counterexamples exist [15], showing that the rate may slow down for general S as α grows. Our first contribution in this paper is to characterize a special class of skew-symmetric matrices whose perturbation causes not just $\Re\lambda_M$, but the real part of *all* eigenvalues, to converge to the optimal rate. This answers more than just Problem 1.

Instead of the asymptotic approach, the attainability of the optimal rate by a certain skew-symmetric matrix, as is desired in Problem 2, was probably first mentioned in [15]. It is of practical importance to actually construct such a perturbation matrix. The very fact that such an optimal irreversible, distribution-preserving perturbation of reversible diffusions does exist has been formally proved in [20] by the argument which can be turned into an algorithm. The drawback, in our opinion, is the need of repeated Gram-Schmidt orthogonalization which is prone to numerical instability. In contrast, our second contribution in this paper is to convert an induction argument used in [15] into a fast divide-and-conquer algorithm which constructs such a matrix recursively and effectively without the concern of stability as that in [20]. This approach fully answers Problem 2.

2 Asymptotic Approach to the Optimum

We begin our discussion by studying the asymptotic behavior of the perturbation in the form $D + \alpha SD$ when α goes to infinity. We are particularly interested in the behavior of $\Re\lambda_M(D+\alpha SD)$ as it represents the convergence rate of the Gaussian diffusion (15) to its equilibrium.

Performing an orthogonal similarity transformation if needed, we may assume without loss of generality that D is a diagonal matrix to begin with. Since D is negative definite, we may further carry out the similarity transformation

$$(-D)^{1/2}(D + SD)(-D)^{-1/2} = D - (-D)^{1/2}S(-D)^{1/2}, \tag{23}$$

where the matrix

$$K := (-D)^{1/2}S(-D)^{1/2} \tag{24}$$

remains to be skew-symmetric. Thus it suffices to consider perturbations of the form $D - \alpha K$, which effectively simplifies the analysis below.

2.1 A special case

In the 2-dimensional case, any skew-symmetric matrix K will guarantee the convergence of $\Re\lambda_M(D-\alpha K)$ to the optimal rate. It is also trivial to construct a matrix K to attain the optimal rate. For higher dimensional cases, we begin with a specific skew-symmetric matrix defined by

$$K_0 := \begin{bmatrix} 0 & 1 & \cdots & 1 \\ -1 & \ddots & & \vdots \\ \vdots & & \ddots & 1 \\ -1 & \cdots & -1 & 0 \end{bmatrix}, \tag{25}$$

whose properties eventually motivate us to consider more general cases. We shall show that this constant matrix K_0 defined above works universally for all D. Our analysis relies on the fact that we can characterize the spectral information of K_0 completely in the following lemma.

Lemma 1 *Suppose that $(\hat{\lambda}, \hat{\mathbf{x}})$ is an eigenpair of the matrix K_0 with $\|\hat{\mathbf{x}}\|_2 = 1$. Then*

1. *There are n eigenpairs whose eigenvalues are pure imaginary and whose eigenvectors form a unitary matrix.*
2. *The kth eigenvalues $\hat{\lambda}_k$ is of the form $\hat{\lambda}_k = \frac{1+c_k}{1-c_k}$, where $c_k = e^{(\pi(1+2k)i)/n}$ is an n-th root of negative unity.*
3. *If $\hat{\mathbf{x}} = [\hat{x}_1, \cdots, \hat{x}_n]^\top$, then $\hat{x}_j\bar{\hat{x}}_j = 1/n$ for all $1 \le j \le n$.*

Proof The first result is true for any skew-symmetric matrices. Only the last two assertions relevant to K_0 need be proved. From the equation $K_0\hat{\mathbf{x}} = \lambda\hat{\mathbf{x}}$, we can induce easily the system of equations

$$\begin{cases} \hat{x}_2 + \hat{x}_1 = \hat{\lambda}(\hat{x}_1 - \hat{x}_2), \\ \hat{x}_3 + \hat{x}_2 = \hat{\lambda}(\hat{x}_2 - \hat{x}_3), \\ \vdots \\ \hat{x}_n + \hat{x}_{n-1} = \hat{\lambda}(\hat{x}_{n-1} - \hat{x}_n). \end{cases}$$

It follows that

$$\begin{cases} \hat{x}_2 = \frac{\hat{\lambda}-1}{\hat{\lambda}+1}\hat{x}_1 := c\hat{x}_1, \\ \hat{x}_3 = \frac{\hat{\lambda}-1}{\hat{\lambda}+1}\hat{x}_2 = c^2\hat{x}_1, \\ \vdots \\ \hat{x}_n = \frac{\hat{\lambda}-1}{\hat{\lambda}+1}\hat{x}_{n-1} = c^{n-1}\hat{x}_1. \end{cases}$$

Using the facts that $(c + c^2 + c^3 + \cdots + c^{n-1})\hat{x}_1 = \hat{\lambda}\hat{x}_1$ and $\hat{\lambda} = \frac{1+c}{1-c}$, we see that $c^n = -1$ and hence $|\hat{x}_1| = |\hat{x}_2| = \cdots = |\hat{x}_n|$. Since $\|\hat{\mathbf{x}}\|_2 = 1$, we must have $|\hat{x}_k|^2 = 1/n$ for all $1 \le j \le n$.

□

The eigenvalues and eigenvectors of $D - \alpha K_0$ are continuous functions of α when α is large enough [1,18]. Denote this fact by

$$(D - \alpha K_0)\mathbf{x}(\alpha) = \lambda(\alpha)\mathbf{x}(\alpha). \tag{26}$$

Suppose that $(\lambda(\alpha), \mathbf{x}(\alpha))$ is an eigenpair of the matrix $D-\alpha K_0$ with $\|\mathbf{x}(\alpha)\|_2 = 1$. Then

$$\lambda(\alpha) = \langle (D - \alpha K_0)\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle = \langle D\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle - \alpha\langle K_0\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle. \tag{27}$$

Observe that the first term on the right-hand side of (27) is real and negative, while the second term is purely imaginary because $\langle K_0\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle = \langle \mathbf{x}(\alpha), K_0^\top\mathbf{x}(\alpha)\rangle = \langle \mathbf{x}(\alpha), -K_0\mathbf{x}(\alpha)\rangle = -\overline{\langle K_0\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle}$. We know therefore

$$\Re\lambda(\alpha) = \langle D\mathbf{x}(\alpha), \mathbf{x}(\alpha)\rangle. \tag{28}$$

Trivially, $D - \alpha K_0$ and $D/\alpha - K_0$ have the same eigenvectors. In the meantime, by continuity [1,18], eigenvectors of $D/\alpha - K_0$ converge to those of $-K_0$ when α goes to infinity[1]. Suppose that $\mathbf{x}(\alpha) \to \hat{\mathbf{x}}$, where $-K_0\hat{\mathbf{x}} = \hat{\lambda}\hat{\mathbf{x}}$. As α goes to infinity, by (28), $\Re\lambda(\alpha)$ converges to $\langle D\hat{\mathbf{x}}, \hat{\mathbf{x}}\rangle$. By using Lemma 1 and that D being diagonal, we find that

$$\langle D\hat{\mathbf{x}}, \hat{\mathbf{x}}\rangle = \langle D, \hat{\mathbf{x}}\hat{\mathbf{x}}^*\rangle = \frac{\mathrm{trace}(D)}{n}. \tag{29}$$

Since the above argument is applicable to every eigenpair $(\lambda(\alpha), \mathbf{x}(\alpha))$, we have established the following result which is new.

[1] Complex eigenvalues cannot be ordered. The continuous dependence of one particular eigenvalue and its associated eigenvector among other eigenpairs therefore has to be carefully discerned.

Theorem 1 *In the Ornstein-Uhlenbeck process (15), suppose that the skew-symmetric matrix K_0 is added to a negative diagonal matrix D such that the drift term is $B = D - \alpha K_0$. Then the optimal rate of convergence can be reach asymptotically via any of the eigenvalue curves, that is,* $\lim_{\alpha\to\infty} \Re\lambda_j(D - \alpha K_0) = \frac{\mathrm{trace}(D)}{n}$ *for every* $1 \le j \le n$.

Example 1. Demonstrated in Figure 1 is a numerical example of Theorem 1 for the matrix

$$D = \mathrm{diag}\left\{-3.6517, -2.4430, -2.8926, -1.1864, -2.2942\right\}.$$

All eigenvalues are drawn. Of particular interest is that in the initial state, curves of $\Re\lambda(\alpha)$ and $\Im\lambda(\alpha)$ might merge or bifurcate, indicating eigenvalues are evolving from distinct reals to complex conjugates or vice versa. Since complex numbers cannot be ordered, if all eigenvalues are generated simultaneously, we cannot tell which is which as a continuous function of α. Change of colors in any particular curve in Figure 1 indicates the different order that eigenvalue is computed by the MATLAB algorithm. After piecing together, continuity is preserved. It should be noted, however, that at certain points the eigenvalues are not differentiable.

In this particular example, we see that the function $\Re\lambda_M(D - \alpha K_0)$, namely, the top curve, is monotone decreasing. This is not always the case and counterexamples can easily be found. When α is sufficiently large, nevertheless, eigenvalues of $D - \alpha K_0$ for any D should behave proportionally (indeed, α times) to those of $-K_0$, which effectively ends the stage of merging or bifurcating. Convergence of real parts of all eigenvalue values of $D - \alpha K_0$ is evident.

Upon a close examination of the argument above, we actually have the following generalization which not only solves Problem 1, but also is valid for any real diagonal matrix D without definiteness restrictions. A more general class of matrices, including K_0, can be used for the purpose of asymptotic perturbation.

Theorem 2 *Given an arbitrary diagonal matrix D, suppose that K is a skew-symmetric matrix with the properties that*

1. *Its eigenvalues are pairwise distinct;*
2. *Each normalized eigenvector $\hat{\mathbf{x}} = [\hat{x}_1, \cdots, \hat{x}_n]^\top$ is such that $\hat{x}_j\bar{\hat{x}}_j = 1/n$ for all $1 \le j \le n$.*

Then $\lim_{\alpha\to\infty} \Re\lambda_j(D - \alpha K) = \frac{\mathrm{trace}(D)}{n}$ *for every* $1 \le j \le n$.

Proof The first property ensures that the eigenvector $\mathbf{x}(\alpha)$ of the pencil $D - \alpha K$ is continuous for α large enough [1,18], whence $\mathbf{x}(\alpha) \to \hat{\mathbf{x}}$ for some eigenvector $\hat{\mathbf{x}}$ of K. By the fact that D is diagonal, the second property guarantees that $\Re\lambda_j(D - \alpha K)$ converges to the optimal rate $\frac{\mathrm{trace}(D)}{n}$ for all $1 \le j \le n$.

□

The question now is what matrices can possess spectral properties described above. We characterize one special class of matrices below.

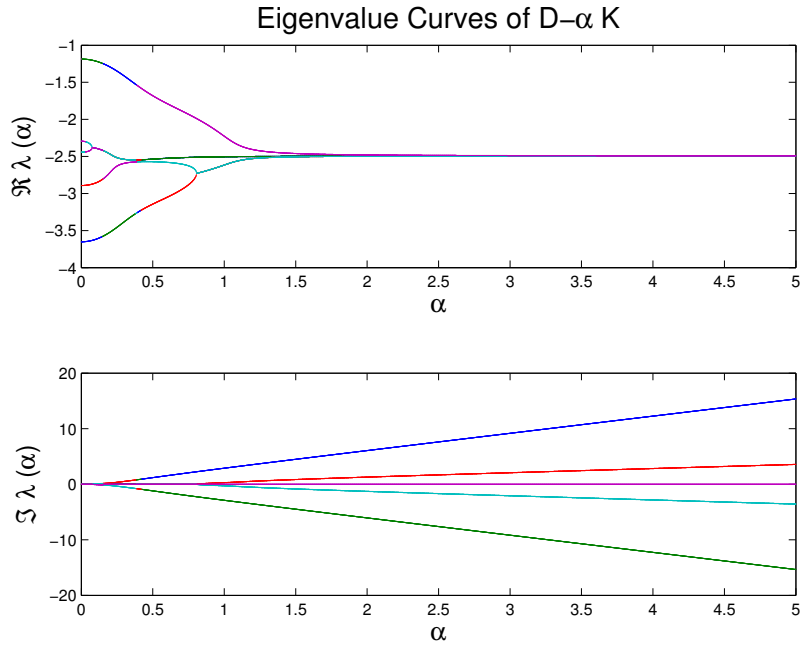


Fig. 1 Real and imaginary parts of the eigenvalue curves of $D - \alpha K$.

2.2 t-circulant matrices

Given an arbitrary row vector $\mathbf{c} := [c_0, \ldots, c_{n-1}]$ and a constant t, consider the so-called t-circulant matrix [10]

$$C(\mathbf{c}; t) = \begin{bmatrix} c_0 & c_1 & c_2 & \ldots & c_{n-1} \\ tc_{n-1} & c_0 & c_1 & & c_{n-2} \\ tc_{n-2} & tc_{n-1} & c_0 & & c_{n-3} \\ \vdots & \vdots & & & \vdots \\ tc_1 & tc_2 & tc_3 & \ldots & c_0 \end{bmatrix}. \tag{30}$$

The special case $C(\mathbf{c}; 1)$ is precisely a conventional circulant matrix that has been extensively studied in the literature [11]. Define the characteristic polynomial

$$p_{\mathbf{c}}(x) = \sum_{k=0}^{n-1} c_k x^k. \tag{31}$$

Let Π_t denote the specific matrix

$$\Pi_t := \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & & 1 \\ t & 0 & & \ldots & 0 \end{bmatrix}. \tag{32}$$

It is easy to see that

$$C(\mathbf{c};t) = \sum_{k=0}^{n-1} c_k \Pi_t^k = p_{\mathbf{c}}(\Pi_t). \tag{33}$$

Because of this representation, many important properties of t-circulant matrices are similar to those of circulant matrices. For instance, t-circulant matrices are closed under multiplication and commutative under multiplication. Indeed, the matrix Π_t plays a fundamental role [11].

Let ω denote the primitive n-th root of unity

$$\omega := \exp(\frac{2\pi\imath}{n}). \tag{34}$$

Define the diagonal matrix

$$\Omega := \mathrm{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}), \tag{35}$$

and let $F = [f_{ij}]$ denote the so-called discrete Fourier matrix where

$$f_{ij} := \frac{1}{\sqrt{n}} \omega^{(i-1)(j-1)}, \quad 1 \le i,\, j \le n. \tag{36}$$

It is well known that F is a unitary matrix and

$$\Pi_1 = F\Omega F^*. \tag{37}$$

It follows from (33) that

$$C(\mathbf{c};1) = F p_{\mathbf{c}}(\Omega) F^*. \tag{38}$$

is the spectral decomposition of the circulant matrix $C(\mathbf{c};1)$. In a similar way, let λ denote the primitive n-th root of t

$$\lambda := |t|^{1/n} \exp(\frac{(\theta + 2\pi)\imath}{n}), \tag{39}$$

where θ is the principal argument of t, and define

$$\Lambda := \mathrm{diag}(1, \lambda, \lambda^2, \ldots, \lambda^{n-1}). \tag{40}$$

It can be shown that

$$\Pi_t = (\Lambda F)(\lambda\Omega)(\Lambda F)^{-1} \tag{41}$$

and, thus,

$$C(\mathbf{c};t) = (\Lambda F) p_{\mathbf{c}}(\lambda\Omega)(\Lambda F)^{-1}. \tag{42}$$

Each column vector of ΛF has length

$$\mu(t) := \sqrt{\frac{t^2 - 1}{n(|t|^{2/n} - 1)}}$$

which converges to 1 as t approaches ± 1.

Note that $f_{ij}\overline{f}_{ij} = \frac{1}{n}$ for all $1 \leq i, j \leq n$. Thus, for any $\mathbf{c}$, the circulant matrix $C(\mathbf{c}; 1)$ seems to be a good candidate since its eigenvectors satisfy the second property of Theorem 2. For our application, however, we also need skew-symmetry. By (30), we need $c_0 = 0$ and $c_i = -tc_{n-i}$, $i = 1, \ldots, n-1$. It follows that we have two possible choices

$$c_i = \begin{cases} -c_{n-i}, \text{ if } t = 1, \\ \ \ c_{n-i}, \text{ if } t = -1, \end{cases} \tag{43}$$

provided that the condition of distinct eigenvalues of the resulting $C(\mathbf{c}; \pm 1)$ is also met.

In fact, when n is even, the case $t = 1$ does always have a zero eigenvalue with multiplicity at least two and, therefore, should not be used in the acceleration scheme. On the other hand, for *generic* $\mathbf{c}$ satisfying $c_i = c_{n-i}$, the eigenvalues $p_{\mathbf{c}}(\lambda\omega^j)$, $j = 1, \ldots, n$, of the (-1)-circulant skew-symmetric matrices $C(\mathbf{c}; -1)$ should be distinct. The matrix K_0 defined in (25) corresponds to the special case with $t = -1$ and $\mathbf{c} = [0, 1, 1, \ldots 1]$. In all, this class of skew-symmetric, (-1)-circulant matrices with distinct eigenvalues provides an additional choice of perturbation matrices usable for the Gaussian diffusion acceleration.

3 Attainability of the Optimum

Thus far we have been using a fixed skew-symmetric, (-1)-circulant matrix K to introduce the perturbation $D - \alpha K$. The resulting diffusion process is shown to converge to its equilibrium at a rate which is asymptotically optimal as α goes to infinity. In this section, we take a different perspective and want to achieve the optimal rate of acceleration by one single additive perturbation.

The existence of a single skew-symmetric matrix $\mathcal{K}$ such that the additive perturbation $D - \mathcal{K}$ attains the optimal rate was probably first noticed in [15]. This very fact was formally proved in a recent paper [20]. For completion, we briefly review the theory in both approaches. Our focus in the subsequent discussion is more on the computational aspect.

One common feature in both approaches is to construct a particular orthogonal matrix, though in different context. To contrast the dissimilarity, we term the construction based on [15] the H approach and that in [20] the Q approach, respectively. We outline the two approaches, respectively, followed by a comparison.

3.1 The Q approach

The following theorem represents a condensed version of a much more detailed discussion in [20]. We only highlight a few points that are essential for both the theory of existence and the construction of the optimal perturbation [20, Proposition 2, Lemma 2].

Theorem 3 *Given a symmetric and negative-definite matrix D, let $\mathcal{K}$ denote a skew-symmetric matrix. Then the following two conditions are equivalent:*

1. *The matrix $D - \mathcal{K}$ is diagonalizable with*

$$\Re\lambda(D - \mathcal{K}) = \frac{\text{trace}(D)}{n}$$

for every eigenvalue of $D - \mathcal{K}$.

2. *There exists a real, symmetric, and positive definite matrix Q such that*

$$\mathcal{K}Q - Q\mathcal{K} = QD + DQ - \frac{2\text{trace}(D)}{n}Q. \tag{44}$$

3. *Suppose that $\{(\lambda_k, \psi_i)\}_{k=1}^n$ are eigenpairs of Q and that $\psi_1, \ldots, \psi_n$ forms an orthonormal basis of $\mathbb{R}^n$. Then*

$$\langle \psi_j, D\psi_k \rangle = \begin{cases} \frac{\text{trace}(D)}{n}, & \text{if } j = k, \\ \frac{\lambda_k - \lambda_j}{\lambda_k + \lambda_j} \langle \psi_j, \mathcal{K}\psi_k \rangle, & \text{if } j \neq k. \end{cases} \tag{45}$$

The third condition in Theorem 3 can be converted into an algorithm which we refer to as the Q approach because of its dependence on the symmetric and positive definite matrix Q. The procedure proposed in [20] involves two steps. The first step is to construct an orthonormal basis $\{\psi_1, \ldots, \psi_n\}$ satisfying the first equation in (45). This construction is summarized in Algorithm 1. Noticeable in the algorithm is the requirement of repeated orthogonalization. Once the orthonormal basis $\{\psi_1, \ldots, \psi_n\}$ is obtained, the second step is to choose distinct and positive eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ and define the skew-symmetric matrix $\mathcal{K}$ according to the second equation in (45).

There are a few potential drawbacks of the algorithm proposed in [20]. First, the Gram-Schmidt orthogonalization process proposed in the third step in the loop of Algorithm 1 is prone to be numerical unstable, especially when the given matrix D is ill-conditioned. This part of calculation could be replaced by the QR decomposition. The nature of progressive orthogonalization only among $\{\widehat{\psi}, \psi_{k+1}, \ldots, \psi_n\}$, however, can still cause loss of orthogonality in the final $\{\psi_1, \ldots, \psi_n\}$. Second, as the re-orthogonalization is needed throughout the loop, the computation overhead is quite high. In contrast, we next discuss the H approach which employs the divide-and-conquer technique that is similar to the fast Fourier transform calculation.

3.2 The H approach

The following result [15, Lemma 4.1] is the theoretical basis of the H approach.

Theorem 4 *Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric with respect to a prescribed inner product $[\cdot, \cdot]$ in $\mathbb{R}^n$. Let m denote the greatest integer less than or equal to the fraction $\frac{n}{2}$. Then there exists an orthogonal decomposition with respect to $[\cdot, \cdot]$,*

$$\mathbb{R}^n = H_0 \oplus H_1 \oplus \cdots \oplus H_m$$

Algorithm 1 Constructing orthonormal basis (implicit Q) for optimal non-reversible perturbation

Require: An arbitrary orthonormal basis $\{\psi_1, \ldots, \psi_n\}$ of $\mathbb{R}^n$
Ensure: An orthonormal basis satisfying the first equation in (45)

for $k = 1, \cdots, n-1$ **do**

1. Permute the subsequence $\{\psi_k, \ldots, \psi_n\}$, if necessary, so that

$$\langle \psi_k, D\psi_k \rangle = \min_{j=k,\ldots,n} \langle \psi_j, D\psi_j \rangle < \frac{\mathrm{trace}(D)}{n}$$

and

$$\langle \psi_{k+1}, D\psi_{k+1} \rangle = \max_{j=k,\ldots,n} \langle \psi_j, D\psi_j \rangle > \frac{\mathrm{trace}(D)}{n}.$$

2. Compute the scalar s such that

$$\widehat{\psi} := \cos(s)\psi_k + \sin(s)\psi_{k+1}$$

satisfies

$$\langle \widehat{\psi}, D\widehat{\psi} \rangle = \frac{\mathrm{trace}(D)}{n}.$$

3. Employ the Gram-Schmidt process to orthogonalize the vectors $\{\widehat{\psi}, \psi_{k+1}, \ldots, \psi_n\}$ to $\{\widehat{\psi}, \widehat{\psi}_{k+1}, \ldots, \widehat{\psi}_n\}$.

end for

such that:

1. *For $1 \le k \le m$, the subspace H_k is of dimension two. Also,* $\mathrm{trace}(\mathscr{P}_k A) = \frac{2}{n}\mathrm{trace}(A)$*, where $\mathscr{P}_k$ denotes the orthogonal projection onto H_k;*
2. *If n is odd, then H_0 is of dimension one and* $\mathrm{trace}(\mathscr{P}_0 A) = \frac{1}{n}\mathrm{trace}(A)$*, where $\mathscr{P}_0$ is the projection onto H_0.*

Note that only symmetry is needed in Theorem 4, whereas the notion of symmetry depends on the inner product being used. The theorem is applicable to any prescribed inner product in $\mathbb{R}^n$. For our application, we shall employ the standard Euclidean inner product only. Under such a situation, Theorem 4 can be interpreted as follows, whereas our goal is to construct the orthogonal matrix U stated in the corollary.

Corollary 1 *Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$, there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that the quantity* $\mathrm{trace}(A)$ *is redistributed through the diagonal of the matrix $T := U^\top AU$ in the following way:*

1. *From the bottom up, each 2×2 diagonal blocks of T shares equally $\frac{2}{n}$ portion of* $\mathrm{trace}(A)$*.*
2. *If n is odd, then the $(1,1)$ element of T is precisely* $\frac{1}{n}\mathrm{trace}(A)$*.*

Our contribution is at converting the otherwise a theoretical proof by induction of the above theorem into a recursive algorithm that computes the needed perturbation effectively. The basic idea is to split the basis of orthonormal eigenvectors (with respect to the inner product reduced $[\cdot,\cdot]$) as evenly as possible into two groups each of which spans lower dimensional subspaces. If the matrix representation of A with respect to these subspaces

can be found, then the lower dimensional decompositions of the restricted transformations are guaranteed to be solvable according to the induction hypothesis. After finding the decompositions for the lower dimensional problems, a mechanism is proposed to revet the two subproblems together to form the higher dimensional problem.

By repeating the above argument, each of the two subproblems can further be downsized. In this way, the original problem is first *divided* into subproblems whose matrix representations are 2×2 or 3×3 blocks of which solutions are explicitly known. We then solve the original problem by *conquering* these small blocks to build up the original size. This divide and conquer process is similar to that occurred in the radix-2 fast Fourier transform. With the aid of modern programming languages that allow a subprogram to invoke itself recursively, induction proofs can often be implementable for numerical computation. Many classical results, such the Sing-Thompson theorem and the Weyl-Horn theorem [22], whose original proof involves induction can be adopted into a similar recursive approach for numerical construction [8, 9].

A working copy of our recursive algorithm is displayed in Table 1. A key ingredient in our computation is the following observation concerning the projection and the representation of the restricted map.

Lemma 2 *Suppose that $\mathcal{V} \subset \mathbb{R}^n$ is a k-dimensional invariant subspace under A. Let columns of the matrix $V \in \mathbb{R}^{n\times k}$ denote an orthonormal basis of the subspace $\mathcal{V}$. Then*

1. *With respect to the original basis of $\mathbb{R}^{n\times n}$, the orthogonal projection $\mathscr{P}_{\mathcal{V}}(\mathbf{x})$ of any $\mathbf{x} \in \mathbb{R}^n$ onto the subspace $\mathcal{V}$ is given by $VV^\top \mathbf{x} \in \mathbb{R}^n$.*
2. *With respect to the specific basis V, the matrix representation of the restricted map $A|_{\mathcal{V}}$ of A is given by $V^\top AV \in \mathbb{R}^{k\times k}$.*

For clarity, we briefly explain the structure of the algorithm below. Though these statements constitute essentially the induction proof in [15, Lemma 4.1], note how we interpret them in simple yet effective computational terms in the algorithm.

(a) If the size of the underlying matrix is 2×2, the goal is automatically accomplished. We take U to be the orthogonal matrix of orthonormal eigenvectors.
(b) If the size of the underlying matrix is 3×3,
  (i) Take the (normalized) average of its orthonormal eigenvectors as the first vector of U. This will ensure the second bullet in Corollary 1.
  (ii) Compute an orthonormal basis for the 2-dimensional subspace which is orthogonal to the normalized average. These two basis vectors constitute the remaining two columns of the 3×3 orthogonal matrix U.
(c) If the size of the underlying matrix is $(2m) \times (2m)$,
  (i) Split the orthonormal eigenvectors evenly into two groups of size m.
  (ii) Compute the $m \times m$ matrix representation of the restriction of the underlying matrix to the subspace spanned by each of the subgroup of orthonormal eigenvectors.

(iii) Apply the recursion to obtain two $m \times m$ orthogonal matrices associated with these restricted maps.
(iv) Express these orthogonal matrices in terms of the original basis of orthonormal eigenvectors. (The resulting matrices are of size $(2m) \times m$ with orthogonal columns. Matrices of this type are called Stiefel matrices.)
(v) Depending on whether m is even or odd, assemble the two $2m \times m$ Stiefel matrices into an $(2m) \times (2m)$ orthogonal matrix in accordance with the specific recipe described in the algorithm. (This is the process of conquering.)

(d) If the size of the underlying is $(2m+1) \times (2m+1)$,
(i) Take the (normalized) average of all its orthonormal eigenvectors as the first vector of U.
(ii) Compute the $2m \times 2m$ matrix representation of the underlying matrix when restricted to the subspace orthogonal to the normalized average.
(iii) Apply the recursion to compute the $(2m) \times (2m)$ orthogonal matrix associated with this restricted map.
(iv) Express this orthogonal matrix in terms of the original basis of orthonormal eigenvectors.

The recursive nature of the above algorithm is intriguing because it self-organizes an otherwise fairly complicated computation. Consider a problem of size $n = 2^t$, for example, the recursion breaks down the matrix into 2^{t-1} blocks of size 2×2 through $2^{t-1} - 1$ self-callings. It appears that each calling should involve an eigenvalue computation. However, we hasten to point out that such a computation is redundant. For instance, as soon as `[U,D] = eig(A)` is calculated, the matrix representations `A1` and `A2` of the restricted maps are necessarily diagonal, respectively, and the spectral decompositions in the next levels of splitting can be totally spared. In other words, throughout the process of dividing only one eigenvalue computation is needed, whereas throughout the process of conquering most calculations involve only matrix arithmetic such as addition or augmentation. With more carefully programming, the overhead can be significantly reduced. We choose not to exploit these features in details for the sake of making the code more self-explanatory. Even so, our experiments in Section 3.3 clearly evidence the computational advantages of the H approach over the Q approach.

We now explain how the optimal skew-symmetric perturbation $\mathcal{K}$ can be constructed. Suppose the matrix $T = U^\top AU$ is now in hand. Write

$$T = T_- + T_0 + T_-^\top,$$

where T_0 denotes the block diagonal of T referred to in Corollary 1 and T_- denotes the lower triangular portion of T below the block diagonal T_0. Recall that T_0 consists of all 2×2 diagonal blocks starting from the lower right corner of T and, if n is odd, the $(1,1)$ entry of T. By construction, each of these 2×2 blocks has trace equal to $\frac{2}{n}\text{trace}(A)$ and, if n is odd, the $(1,1)$ entry of T is precisely $\frac{1}{n}\text{trace}(A)$. Let $\mathfrak{K}_0$ denote a block diagonal skew-symmetric matrix whose block structure is conformal to that of T_0. The off-diagonal entries of

$\mathfrak{K}_0$ will be specified later. Define

$$\mathfrak{K} := T_- + \mathfrak{K}_0 - T_-^\top. \tag{46}$$

Then $\mathfrak{K}$ is skew-symmetric. Consider the matrix

$$T - \mathfrak{K} = (T_0 - \mathfrak{K}_0) + 2T_-^\top \tag{47}$$

which is pseudo-upper triangular and, hence, has eigenvalues precisely equal to those of its diagonal blocks. A typical 2×2 block is of the form

$$\begin{bmatrix} a & b+k \\ b-k & c \end{bmatrix},$$

where $a+c = \frac{2}{n}\text{trace}(A)$ and k is an unspecific real number. Trivially, if $|k|$ is large enough, then then both eigenvalues in this block will be complex with real parts equal to $\frac{1}{n}\text{trace}(A)$. In this way, there exists a skew-symmetric matrix $\mathfrak{K}$ such that the real part of every eigenvalue of $T - \mathfrak{K}$ is equal to precisely $\frac{1}{n}\text{trace}(A)$. Taking

$$\mathcal{K} = U\mathfrak{K}U^\top, \tag{48}$$

we have thus answered an interesting inverse eigenvalue problem — perturb a given symmetric matrix A, not necessarily negative definite, by an additive skew-symmetric matrix $\mathcal{K}$ so that the real part of every eigenvalue of $A - \mathcal{K}$ is driven to precisely the same value $\frac{1}{n}\text{trace}(A)$.

We remark that such a result might be of interest in other fields. For instance, it can be considered as special type of the pole assignment problem in control theory as well as a dual problem mentioned in [29] for the stabilization of a positive-definite matrix by matrix multiplication.

For the application to our Gaussian diffusion process (15), we have now a constructive way to select the perturbation so as to achieve the optimal rate of convergence.

Theorem 5 *In the Ornstein-Uhlenbeck process (15), there exists a skew-symmetric perturbation $\mathcal{K}$ such that when the drift term is $B = D - \mathcal{K}$ with D a negative-definite matrix, the optimal rate of convergence can be reach precisely, that is, $\Re\lambda_k(D - \mathcal{K}) = \frac{\text{trace}(D)}{n}$ for every $1 \le k \le n$.*

Example 2. Corresponding to the same diagonal matrix D in Example 1, we find that the recursive algorithm returns the orthogonal matrix

$$U = \begin{bmatrix} 0.4472 & -0.0204 & 0.6460 & 0.1881 & 0.5890 \\ 0.4472 & 0.5476 & 0.3138 & -0.1697 & -0.6106 \\ 0.4472 & -0.4213 & -0.2061 & 0.6790 & -0.3449 \\ 0.4472 & -0.5612 & -0.0960 & -0.6890 & -0.0336 \\ 0.4472 & 0.4553 & -0.6577 & -0.0085 & 0.4002 \end{bmatrix}$$

which leads to

$$T = \begin{bmatrix} -2.4936 & -0.1894 & -0.4054 & -0.6259 & -0.2413 \\ -0.1894 & -2.0966 & 0.0000 & 0.6186 & 0.0000 \\ -0.4054 & 0.0000 & -2.8905 & -0.0000 & -0.5270 \\ -0.6259 & 0.6186 & -0.0000 & -2.0966 & -0.0000 \\ -0.2413 & 0.0000 & -0.5270 & -0.0000 & -2.8905 \end{bmatrix}.$$

The matrix $\mathfrak{K}$ in (46) is of the form

$$\mathfrak{K} = \begin{bmatrix} 0 & 0.1894 & 0.4054 & 0.6259 & 0.2413 \\ -0.1894 & 0 & -k_1 & -0.6186 & -0.0000 \\ -0.4054 & k_1 & 0 & 0.0000 & 0.5270 \\ -0.6259 & 0.6186 & -0.0000 & 0 & -k_2 \\ -0.2413 & 0.0000 & -0.5270 & k_2 & 0 \end{bmatrix},$$

where k_1, k_2 are arbitrary numbers with modulus greater than 0.7939. Once k_1, k_2 are chosen, all eigenvalues of $D - U\mathfrak{K}U^\top$ have real part equal to -2.4936.

3.3 A comparison

Our H approach differs from the Q approach in several aspects. It might worth mentioning a few the most significant distinctions in this section.

First, we have already pointed out that the repeated re-orthogonalization needed in the Q approach is potentially unstable. The numerical instability of the Q approach can be seen in Figure 2 where we plot the loss of orthogonality $\|\Psi^\top\Psi - I\|$ of the matrix $\Psi = [\psi_1, \ldots, \psi_n]$ constructed from Algorithm 1 for 20 runs of the case $n = 500$. Despite that the Gram-Schmidt process originally suggested in [20] for the segment of vectors $\{\widehat{\psi}, \psi_{k+1}, \ldots \psi_n\}$ has been replaced by the more stable QR decomposition, the resulting basis gradually deviates from being orthogonal to the segment of vectors $\{\psi_1, \ldots, \psi_{k-1}\}$. The consequence of this loss of orthogonality is that the resulting skew-symmetric perturbation will also drift from being optimal. This is a major structural disadvantage inherent in the Q approach. Such a phenomenon does not occur in the H approach.

Second, calculating the sequence of QR decompositions required by the Q approach alone is expensive. Excluding other overhead incurred in the algorithm, merely this part of orthogonalization will cost approximately $O(\frac{n^4}{2})$ floating-point operations (flops) by using the formula given in [30]. In contrast, the divide-and-conquer nature in the H approach is similar to that of the fast Fourier transform. Although lots of matrix multiplications and additions are involved in the H approach, the matrix sizes are continually halved. The overhead of the H approach consists of approximately $O(2n^3)$ flops from matrix to matrix arithmetic plus those involved in one eigenvalue computation of A. The CPU time ratio of the Q approach versus the H approach measured by the tic-toc function in Matlab for the portion of orthonormal basis computation on the platform, iMac, OSX 10.8.4, 3.4GHz Quad-core i7, 16 GB 1333MHz DDR3, is plotted in Figure 3. For this comparison, we randomly generate the test matrix D with sizes varying from 100 to 1000 and repeat the experiments 20 times for each size. Counting the flops alone, the ratio should be a fraction of $O(n)$ in theory. However, the CPU time includes carriages of other loads, such as the extensive memory swapping and I/O tasks needed by the H approach, which are too complicated to enumerate. Counting all, we observe two interesting phenomena — that the H approach consistently uses less than one hundredth CPU time than the Q approach

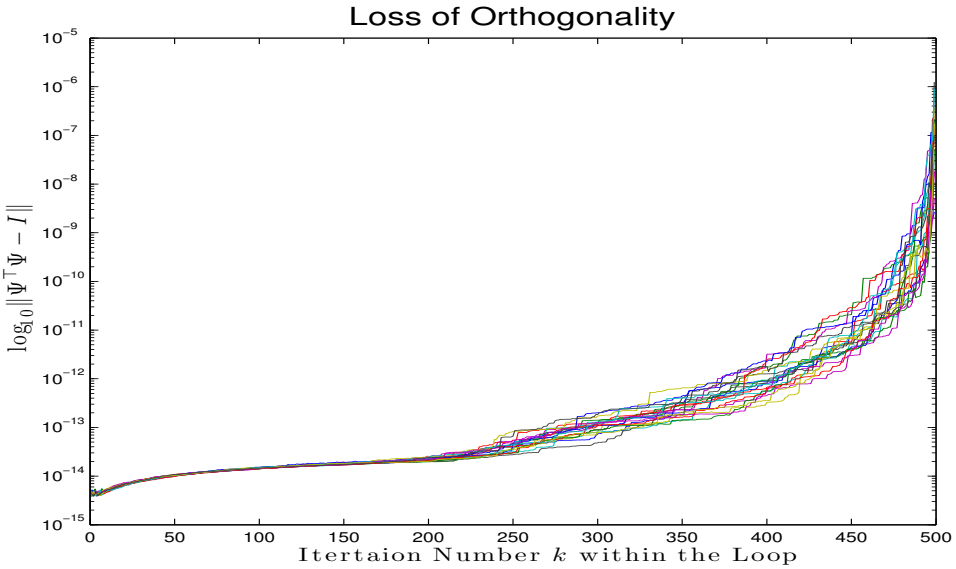


Fig. 2 Loss of orthogonality from 20 repeated experiments of the case $n = 500$.

does, which is significant, and that the ratio is minimum at $n = 500$. Indeed, the average CPU time ratios for $n = 100, 200, \ldots, 1000$ are

$$178.8274, 159.2413, 133.4445, 123.8665, 111.7581, 126.1098, 123.4083, 149.5137, 146.5807, 171.5224,$$

respectively. We think that the convex shape of CPU time ratios when n is increased in this particular application is probably related to the internal system architecture.

Finally, in Example 2, we see that the H approach allows arbitrary but large enough parameters k_1 and k_2, implying infinitely many 2-parameter ways to accelerate to achieve the optimal rate of convergence, provided the orthogonal matrix U is found. In general, the H approach allows $\lfloor \frac{n}{2} \rfloor$ free parameters to define the skew-symmetric perturbation matrix. The Q approach, on the other hand, has the freedom of selecting arbitrary distinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$, provided the orthogonal eigenvectors $\{\psi_1, \ldots, \psi_n\}$ are determined.

4 Conclusion

The problem of skew-symmetric perturbation of a reversible system originates from the desire to accelerate approximation of diffusions to its equilibrium. In the case of Gaussian diffusions (15), the dynamics of spectral gaps lead to an interesting inverse eigenvalue problem — perturbing a symmetric and negative definite matrix with skew-symmetric matrices to achieve optimal convergence. Two types of perturbations are investigated in this paper.

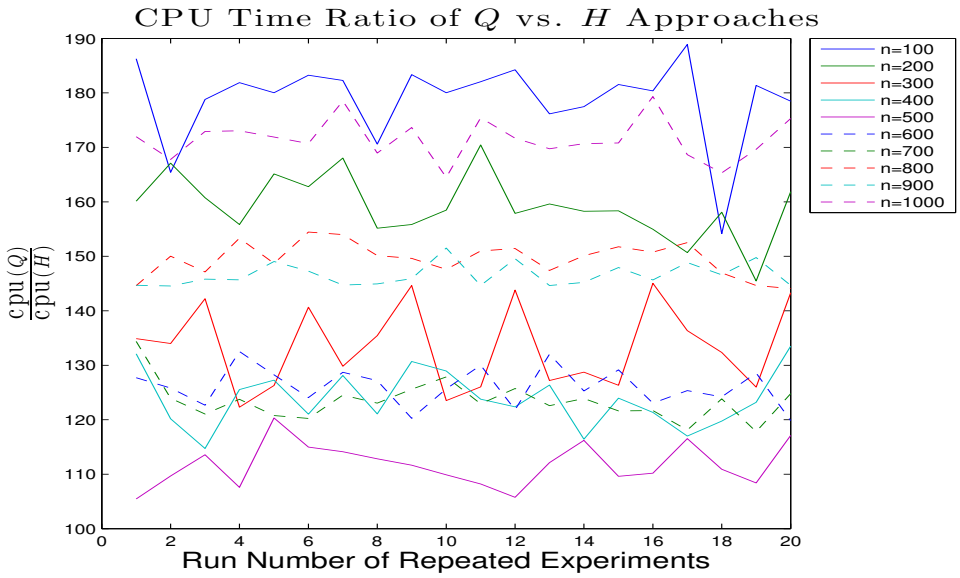


Fig. 3 CPU time ratio of Q approach to the H approach from 20 repeated experiments of sizes 100 to 1000.

One is to study the limiting behavior of growing perturbation of the form $D - \alpha K$ as α goes to infinity, where K comes from a specially selected class of skew-symmetric, (-1)-circulant matrices. This class of matrices can be regarded as representing the best direction of perturbation in the sense that each of its members works universally for all D.

The other is to study the attainability of optimal rate by one direct additive perturbation in the form $D - \mathcal{K}$, where $\mathcal{K}$ is fixed, problem dependent, and no limiting behavior is involved. A novel divide-and-conquer algorithm is proposed for the construction of an optimal $\mathcal{K}$ and is shown empirically superior in both stability and overhead to the recently proposed algorithm in [20].

This research on both the asymptotic approach and the recursive approach leads to results that are innovative and interesting in the field.

References

1. Acker, A.F.: Absolute continuity of eigenvectors of time-varying operators. Proc. Amer. Math. Soc. **42**, 198–201 (1974)
2. Amit, Y.: On rates of convergence of stochastic relaxation for Gaussian and non-Gaussian distributions. J. Multivariate Anal. **38**(1), 82–99 (1991). DOI 10.1016/0047-259X(91)90033-X. URL http://dx.doi.org/10.1016/0047-259X(91)90033-X
3. Amit, Y.: Convergence properties of the Gibbs sampler for perturbations of Gaussians. Ann. Statist. **24**(1), 122–140 (1996). DOI 10.1214/aos/1033066202. URL http://dx.doi.org/10.1214/aos/1033066202

4. Amit, Y., Grenander, U.: Comparing sweep strategies for stochastic relaxation. J. Multivariate Anal. **37**(2), 197–222 (1991). DOI 10.1016/0047-259X(91)90080-L. URL http://dx.doi.org/10.1016/0047-259X(91)90080-L
5. Bhattacharya, R., Denker, M., Goswami, A.: Speed of convergence to equilibrium and to normality for diffusions with multiple periodic scales. Stochastic Process. Appl. **80**(1), 55–86 (1999). DOI 10.1016/S0304-4149(98)00069-6. URL http://dx.doi.org/10.1016/S0304-4149(98)00069-6
6. Chen, T.L., Chen, W.K., Hwang, C.R., Pai, H.M.: On the optimal transition matrix for Markov chain Monte Carlo sampling. SIAM J. Control Optim. **50**(5), 2743–2762 (2012)
7. Chen, T.L., Hwang, C.R.: Accelerating reversible markov chains. Stat. and Probab. Letters **83**, 1956–1962 (2013)
8. Chu, M.T.: On constructing matrices with prescribed singular values and diagonal elements. Linear Algebra Appl. **288**(1-3), 11–22 (1999). DOI 10.1016/S0024-3795(98)10124-6. URL http://dx.doi.org/10.1016/S0024-3795(98)10124-6
9. Chu, M.T.: A fast recursive algorithm for constructing matrices with prescribed eigenvalues and singular values. SIAM J. Numer. Anal. **37**(3), 1004–1020 (electronic) (2000). DOI 10.1137/S0036142998339301. URL http://dx.doi.org/10.1137/S0036142998339301
10. Cline, R.E., Plemmons, R.J., Worm, G.: Generalized inverses of certain Toeplitz matrices. Linear Algebra and Appl. **8**, 25–33 (1974)
11. Davis, P.J.: Circulant matrices. John Wiley & Sons, New York-Chichester-Brisbane (1979). A Wiley-Interscience Publication, Pure and Applied Mathematics
12. Franke, B., Hwang, C.R., Pai, H.M., Sheu, S.J.: The behavior of the spectral gap under growing drift. Trans. Amer. Math. Soc. **362**(3), 1325–1350 (2010). DOI 10.1090/S0002-9947-09-04939-3. URL http://dx.doi.org/10.1090/S0002-9947-09-04939-3
13. Frigessi, A., Hwang, C.R., Younes, L.: Optimal spectral structure of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields. Ann. Appl. Probab. **2**(3), 610–628 (1992)
14. Hermosilla, A.Y.: Skew-symmetric Gaussian diffusions. In: Proceedings of the Third Sino-Philippine Symposium in Analysis (Quezon City, 2000), Special Issue, pp. 47–59 ((2000))
15. Hwang, C.R., Hwang-Ma, S.Y., Sheu, S.J.: Accelerating Gaussian diffusions. Ann. Appl. Probab. **3**(3), 897–913 (1993). URL http://links.jstor.org/sici?sici=1050-5164(199308)3:3¡897:AGD¿2.0.CO;2-Borigin=MSN
16. Hwang, C.R., Hwang-Ma, S.Y., Sheu, S.J.: Accelerating diffusions. Ann. Appl. Probab. **15**(2), 1433–1444 (2005). DOI 10.1214/105051605000000025. URL http://dx.doi.org/10.1214/105051605000000025
17. Hwang, C.R., Sheu, S.J.: On some quadratic perturbation of ornstein-uhlenbeck processes. Soochow Journal of Mathematics **26**(3), 205–244 (2000)
18. Kato, T.: Perturbation theory for linear operators. Classics in Mathematics. Springer-Verlag, Berlin (1995). Reprint of the 1980 edition
19. Kontoyiannis, I., Meyn, S.P.: Geometric ergodicity and the spectral gap of non-reversible Markov chains. Probab. Theory Related Fields **154**(1-2), 327–339 (2012). DOI 10.1007/s00440-011-0373-4. URL http://dx.doi.org/10.1007/s00440-011-0373-4
20. Lelièvre, T., Nier, F., Pavliotis, G.A.: Optimal Non-reversible Linear Drift for the Convergence to Equilibrium of a Diffusion. Journal of Statistical Physics **152**, 237–274 (2013). DOI 10.1007/s10955-013-0769-x
21. Markowich, P.A., Villani, C.: On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis. Mat. Contemp. **19**, 1–29 (2000)
22. Marshall, A.W., Olkin, I.: Inequalities: theory of majorization and its applications, *Mathematics in Science and Engineering*, vol. 143. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York (1979)

23. Mengersen, K.L., Tweedie, R.L.: Rates of convergence of the Hastings and Metropolis algorithms. Ann. Statist. **24**(1), 101–121 (1996). DOI 10.1214/aos/1033066201. URL http://dx.doi.org/10.1214/aos/1033066201
24. Metafune, G., Pallara, D., Priola, E.: Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures. J. Funct. Anal. **196**(1), 40–60 (2002). DOI 10.1006/jfan.2002.3978. URL http://dx.doi.org.prox.lib.ncsu.edu/10.1006/jfan.2002.3978
25. Mira, A.: Ordering and improving the performance of Monte Carlo Markov chains. Statist. Sci. **16**(4), 340–350 (2001). DOI 10.1214/ss/1015346319. URL http://dx.doi.org/10.1214/ss/1015346319
26. Pai, H.M., Hwang, C.R.: Accelerating brownian motion on n-torus. Stat. and Probab. Letters **83**, 1443–1447 (2013)
27. Roberts, G.O., Rosenthal, J.S.: Variance bounding Markov chains. Ann. Appl. Probab. **18**(3), 1201–1214 (2008). DOI 10.1214/07-AAP486. URL http://dx.doi.org/10.1214/07-AAP486
28. Sramer, O., Tweedie, R.L.: Geometric and subgeometric convergence of diffusions with given distributions, and their discretizations ((1997)). Unpublished manuscript
29. Taussky, O.: Positive-definite matrices and their role in the study of the characteristic roots of general matrices. Advances in Math. **2**, 175–186 (1968)
30. Trefethen, L.N., Bau III, D.: Numerical linear algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1997). DOI 10.1137/1.9780898719574. URL http://dx.doi.org.prox.lib.ncsu.edu/10.1137/1.9780898719574
31. Villani, C.: Hypocoercivity, vol. 202. Memoirs Amer. Math. Soc. (2009)

Table 1 A MATLAB code implementing Theorem 4.

```
function [U,first] = split04(A,first);
%
% Given a symmetric matrix A, it computes the orthogonal
% matrix U such that the 2x2 diagonal blocks of the matrix
%
%   T = U'*A*U
%
% shares equally the portion trace(A)/length(A).
%

k = size(A,2);
m = floor(k/2);

if first == 'y'
    [U,D] = eig(A);
    first = 'n';
else
    U = eye(size(A));
    D = A;
end

if k == 2
    U = U;
elseif k == 3
    U = [sum(U,2)/sqrt(3), U*null(ones(1,3)/sqrt(3))];
else

    if k == 2*m
        U1 = U(:,1:m);
        A1 = U1'*A*U1;
        [H1,first] = split04(A1,first);
        H1 = U1*H1;
        U2 = U(:,m+1:k);
        A2 = U2'*A*U2;
        [H2,first] = split04(A2,first);
        H2 = U2*H2;

        if floor(m/2)*2 < m
            U = [H1(:,1),H2(:,1),(H1(:,2:m)-H2(:,2:m))/sqrt(2),...
                (H1(:,2:m)+H2(:,2:m))/sqrt(2)];
        else
            U = [(H1+H2)/sqrt(2), (H1-H2)/sqrt(2)];
        end
    else
        V = U*null(ones(1,k)/sqrt(k));
        A1 =  V'*A*V;
        [H1,first] = split04(A1,first);
        H1 = V*H1;
        U = [sum(U,2)/sqrt(k), H1];
    end
end
```