Journal of Scientific Computing (2021) 87:87
https://doi.org/10.1007/s10915-021-01506-y

A Stochastic Homotopy Tracking Algorithm for Parametric Systems of Nonlinear Equations

Wenrui Hao[1] · **Chunyue Zheng**[1]

Received: 22 September 2020 / Revised: 24 March 2021 / Accepted: 20 April 2021 / Published online: 4 May 2021

Abstract
The homotopy continuation method has been widely used in solving parametric systems of nonlinear equations. But it can be very expensive and inefficient due to singularities during the tracking even though both start and end points are non-singular. The current tracking algorithms focus on the adaptivity of the stepsize by estimating the distance to the singularities but cannot avoid these singularities during the tracking. We present a stochastic homotopy tracking algorithm that perturbs the original parametric system randomly each step to avoid the singularities. We then prove that the stochastic solution path introduced by this new method is still closed to the original solution path theoretically. Moreover, several homotopy examples have been tested to show the efficiency of the stochastic homotopy tracking method.

Keywords Stochastic homotopy tracking · Nonlinear parametric systems · Convergence analysis

1 Introduction

The homotopy continuation method is the main tool to solve systems of polynomial equations in numerical algebraic geometry (NAG) [6,14,18]. The basic idea is to trace out a one-real-dimensional solution curve described implicitly by a system of equations: given a nonlinear system $\boldsymbol{F}(\boldsymbol{u})$ to solve, one first forms a nonlinear system $\boldsymbol{G}(\boldsymbol{u})$ that is related to $\boldsymbol{F}(\boldsymbol{u})$ in a prescribed way but has known, or easily computable solutions. The systems $\boldsymbol{G}(\boldsymbol{u})$ and $\boldsymbol{F}(\boldsymbol{u})$ are combined to form a homotopy, such as the linear homotopy

$$\boldsymbol{H}(\boldsymbol{u}, t) = \boldsymbol{F}(\boldsymbol{u})(1-t) + t\boldsymbol{G}(\boldsymbol{u}) = 0, \tag{1}$$

where $\boldsymbol{G}(\boldsymbol{u})$ is a start system with known solutions and $\boldsymbol{F}(\boldsymbol{u})$ is the target system we want to solve. Then solutions of $\boldsymbol{F}(\boldsymbol{u}) = 0$ can be solved by tracking t from 1 to 0 via this linear

✉ Wenrui Hao
wxh64@psu.edu

Chunyue Zheng
cmz5199@psu.edu

[1] Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA

homotopy. In NAG, there is a well-developed theory on how to choose the start system $\boldsymbol{G}(\boldsymbol{u})$ to guarantee all the solutions of $\boldsymbol{F}(\boldsymbol{u})$ via this homotopy. Furthermore, by constructing different start systems based on other theories, the homotopy continuation method has also successfully applied to compute solutions of nonlinear systems such as nonlinear PDEs [11,19,20], machine learning [9,10], and nonlinear systems in biology and physics [12]. Moreover, the homotopy continuation method has been also used to explore the general parameter space, so-called paramotopy, as a quite powerful tool for many classes of problems that arise in practice [3].

In the linear homotopy setup, each solution path can be tracked via the prediction/correction algorithm [6,14,18] which is referred as the homotopy tracking algorithm. This algorithm could become very inefficient if the parametric system is singular or near singular. To avoid the singular system, in NAG [18], the gamma trick is proposed to construct a random homotopy setup in (1) by multiplying a random complex number. Then the probability of hitting a singularity during the tracking is zero. Nevertheless, the system could be still near singular so that the homotopy tracking is still time-consuming [6,14,18]. In order to address this numerical challenge, an adaptive multi-precision path tracking algorithm [5] has been developed by adjusting precision in response to step failure according to the error estimates. An adaptive step-size homotopy tracking method [13] has also been developed to control the tracking stepsize each time to compute the bifurcation point. An endgame algorithm [7] has also been widely used to deal with the singularities at $t = 0$. However, all these algorithms could be very slow and inefficient when the size of nonlinear systems becomes large [4].

Stochastic algorithms have been widely used in scientific computing [8,17], e.g., the coordinate gradient descent has been developed for solving large-scale optimization problems [16] and has also been revised for solving the leading eigenvalue problem [15]. Motivated by these stochastic algorithms, in this paper, we present an efficient stochastic homotopy tracking method that gives the original system a random perturbation each step so that it can avoid singularities and improve the efficiency during the tracking. The paper is organized as follows: In Sect. 2, we present a novel stochastic homotopy tracking algorithm; In Sect. 3, we analyze the stochastic homotopy tracking algorithm and show the solution path is close to the original solution path under certain conditions; several numerical examples are presented in Sect. 4 to illustrate the efficiency of the stochastic homotopy tracking method.

2 Stochastic Homotopy Continuation Method

Generally speaking, a nonlinear parametric system is written as $\mathbf{F} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$,

$$\mathbf{F}(\mathbf{u}, p) = \mathbf{0}, \tag{2}$$

where $p \in [a, b]$ is a parameter and $\mathbf{u}$ is the variable vector that depends on the parameter p, i.e., $\mathbf{u} = \mathbf{u}(p)$. Suppose we have a solution at the starting point, namely $\mathbf{u}(a) = \mathbf{u}_0$, the homotopy tracking along the solution path, $\mathbf{u}(p)$, reduces down to solving the Davidenko differential equation [6,18],

$$\begin{cases} \mathbf{F}_{\mathbf{u}}(\mathbf{u}, p)\dfrac{d\mathbf{u}}{dp} + \mathbf{F}_p(\mathbf{u}, p) = \mathbf{0}, \\ \mathbf{u}(a) = \mathbf{u}_0, \end{cases} \tag{3}$$

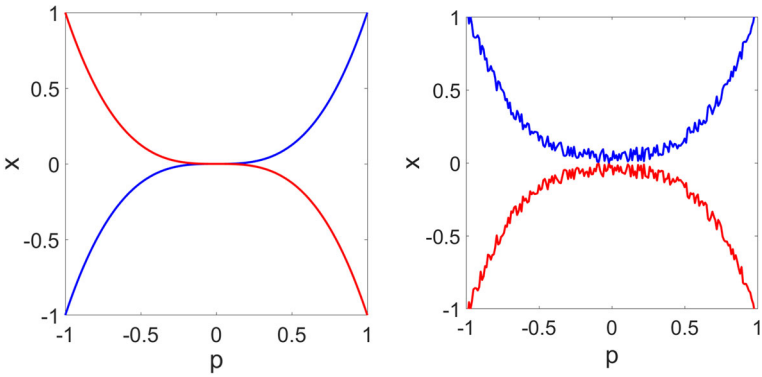

Fig. 1 An illustration example, $x^2 - p^6 = 0$, has two solution paths $x = \pm p^3$ and one bifurcation point at $p = 0$. The traditional homotopy tracking (**Left**) hits the bifurcation point while the stochastic tracking (**Right**) can avoid the bifurcation point by tracking $x = \pm(p^3 + \xi)$, where $\xi \sim \mathcal{N}(0, 0.1)$

where $\mathbf{F_u}(\mathbf{u}, p)$ is the Jacobian matrix and $\mathbf{F}_p(\mathbf{u}, p)$ is the derivative vector with respect to p. If $\mathbf{F_u}(\mathbf{u}, p)$ is nonsingular, the solution path $\mathbf{u}(p)$ is smooth and unique. However, when $\mathbf{F_u}(\mathbf{u}, p)$ becomes singular, the solution path yields different types of bifurcations [6]. Then the numerical homotopy tracking could become very inefficient. In order to solve this numerical issue, a trial-and-error homotopy tracking method [6,18] and an adaptive homotopy tracking method [13] have been developed to control the stepsize of p. However, the computational cost could still be very expensive when the homotopy tracking method is applied to the large-scale nonlinear systems due to the slow tracking near the singularity.

To address this challenge, we propose to solve a stochastic version of the Davidenko differential equation by introducing a noise term, namely

$$\begin{cases} \mathbf{F_u}\big(\mathbf{u}(p,\omega), p\big)d\mathbf{u}(p,\omega) + \mathbf{F}_p(\mathbf{u}(p,\omega), p)dp = \mathbf{g}(\mathbf{u}(p,\omega), p)dW(p,\omega), \\ \mathbf{u}(a,\omega) = \mathbf{u}_0, \end{cases} \tag{4}$$

where ω is a random variable and possesses the initial condition $\mathbf{u}(a, \omega) = \mathbf{u}_0$ with probability one and $dW(p, \omega)$ denotes differential form of the Brownian motion [2]. Then, in this case, the solution path can avoid the singularity with probability one (See Fig. 1 for an illustration).

In order to integrate the idea of the stochastic differential equation into the homotopy tracking, we track the solution $\mathbf{u}(p)$ from $p = a$ to $p = b$ with a stepsize Δp. Then for each $p_k = a + \Delta p \cdot k$, we solve the stochastic system below

$$\tilde{\mathbf{F}}(\mathbf{u}, p_k) = \begin{bmatrix} F_1(\mathbf{u}, p_k) \\ \vdots \\ F_{i-1}(\mathbf{u}, p_k) \\ \mathbf{u}(j) - \tilde{\mathbf{u}}_{k-1}(j) \\ F_{i+1}(\mathbf{u}, p_k) \\ \vdots \\ F_n(\mathbf{u}, p_k) \end{bmatrix} := \mathbf{F}(\mathbf{u}, p_k, \xi = (i, j)) = \mathbf{0}, \tag{5}$$

where $\tilde{\mathbf{u}}_{k-1}$ is the solution from previous step and $\tilde{\mathbf{F}} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ can be viewed as randomly chooses $n - 1$ equations from $\mathbf{F}(\mathbf{u}, p_k)$ and replaces F_i by $\mathbf{u}(j) - \tilde{\mathbf{u}}_{k-1}(j)$. Here the random variable ξ follows the uniform distribution, namely $\mathbb{P}(\xi = (i, j)) = \frac{1}{n^2}$, and

quantifies the perturbations to the original system $\mathbf{F}(\mathbf{u}, p_k)$. More generally, we can randomly replace m ($1 \leq m \leq n$) equations of $\mathbf{F}(\mathbf{u}, p)$ by $\mathbf{u}(\mathcal{J}) = \tilde{\mathbf{u}}_{k-1}(\mathcal{J})$ where $\mathcal{J} = (\mathcal{J}_1, \cdots, \mathcal{J}_m)$ is a m index. Then we define the s-th equation of $\tilde{\mathbf{F}}$ as

$$\tilde{\mathbf{F}}_s(\mathbf{u}, p_k) = \begin{cases} \mathbf{F}_s(\mathbf{u}, p_k), & s \notin \mathcal{I} \\ \mathbf{u}(\mathcal{J}_c) - \tilde{\mathbf{u}}_{k-1}(\mathcal{J}_c), & s \in \mathcal{I} \text{ and } \mathcal{I}_c = s \end{cases}, \tag{6}$$

where $\mathcal{I} = (\mathcal{I}_1, \cdots, \mathcal{I}_m)$ stands for randomly choosing m equations. If $s \in \mathcal{I}$, then we find c such that $\mathcal{I}_c = s$ and replace the s-th equation by the previous value, namely, $\mathbf{u}(\mathcal{J}_c) - \tilde{\mathbf{u}}_{k-1}(\mathcal{J}_c)$. Here $\mathcal{I}$ and $\mathcal{J}$ are randomly drawn from the uniform distribution, namely $\mathbb{P}(\mathcal{I}, \mathcal{J}) = \frac{1}{(C_n^m)^2}$. We denote the set of all possible m indexes as $\mathcal{M}$.

Finally, we summarize the stochastic homotopy tracking algorithm in **Algorithm 1**. In this algorithm, we increase the number of random equations, m, if there is no solution to the stochastic system $\tilde{\mathbf{F}}(\mathbf{u}, p_{k+1}) = 0$. This is equivalent to perform a larger perturbation to the original system by solving fewer equations. Similarly, we could also increase the perturbation by setting an adaptive tolerance for $\|\mathbf{F}(\tilde{\mathbf{u}}_{k+1}, p_{k+1})\| < TOL$ by fixing the number of randomly choosing equation, m.

Algorithm 1: The pseudocode of the stochastic homotopy tracking algorithm.

Input: A step-size Δp, a threshold TOL, and a start point $(\tilde{\mathbf{u}}_0, p_0)$
Output: A nearby solution path $(\tilde{\mathbf{u}}_k, p_k)_{k=1}^N$
for $k = 0, \cdots, N$ **do**
  Set m=1;
  Randomly choose $n - m$ equations and $n - m$ variables to form the stochastic system $\tilde{\mathbf{F}}(\mathbf{u}, p_{k+1})$ (5);
  Solve $\tilde{\mathbf{F}}(\mathbf{u}, p_{k+1}) = \mathbf{0}$ using the predictor-corrector method;
  if $\|\tilde{\mathbf{F}}(\tilde{\mathbf{u}}_{k+1}, p_{k+1})\| < TOL$ **then**
    Update the solution sequence;
  else
    Increase m and solve the stochastic system again.
  end
end

3 Convergence Analysis

We employ the Euler predictor and the Newton corrector [1] for the homotopy tracking algorithm: Given a solution $(\mathbf{u}_0, p_0)$ on the path, that is, $\mathbf{F}(\mathbf{u}_0, p_0) = 0$, an Euler predictor step gives

$$\mathbf{F}_{\mathbf{u}}(\mathbf{u}_0, p_0)\Delta\mathbf{u} = -\mathbf{F}_p(\mathbf{u}_0, p_0)\Delta p, \tag{7}$$

and then letting $\mathbf{u}_1 = \mathbf{u}_0 + \Delta\mathbf{u}$; The Newton corrector reads

$$\mathbf{F}_{\mathbf{u}}(\mathbf{u}_1, p_1)\Delta\mathbf{u} = -\mathbf{F}(\mathbf{u}_1, p_1). \tag{8}$$

Then we repeat this correction until $(\mathbf{u}_1, p_1)$ is on the path. The predictor-corrector method for the stochastic homotopy tracking method needs to replace $\mathbf{F}$ by $\tilde{\mathbf{F}}$ defined in (5) with the corresponding derivatives below:

$$\tilde{\mathbf{F}}_p(\mathbf{u}) = \mathbf{F}_p(\mathbf{u}, \xi = (i, j)) = \mathbf{F}_p(\mathbf{u}) - \frac{\partial F_i}{\partial p}\mathbf{e}_i,$$
$$\tilde{\mathbf{F}}_{\mathbf{u}}(\mathbf{u}) = \mathbf{F}_{\mathbf{u}}(\mathbf{u}, \xi = (i, j)) = \mathbf{F}_{\mathbf{u}}(\mathbf{u}) - \mathbf{e}_i \frac{\partial F_i}{\partial \mathbf{u}}(\mathbf{u}) + E_{ij},$$

where E_{ij} is a matrix with all zero elements except the (i, j)-th element as one. For the general stochastic system (6) with m random equations, we have $\xi = (\mathcal{I}, \mathcal{J})$ and

$$\tilde{\mathbf{F}}_p(\mathbf{u}) = \mathbf{F}_p(\mathbf{u}) - \sum_{i\in\mathcal{I}} \frac{\partial F_i}{\partial p}\mathbf{e}_i \triangleq \mathbf{F}_p(\mathbf{u}) - C(\mathbf{u}, \xi),$$
$$\tilde{\mathbf{F}}_{\mathbf{u}}(\mathbf{u}) = \mathbf{F}_{\mathbf{u}}(\mathbf{u}) - \sum_{i\in\mathcal{I}} \mathbf{e}_i \frac{\partial F_i}{\partial \mathbf{u}}(\mathbf{u}) + \sum_{i\in\mathcal{I}, j\in\mathcal{J}} E_{ij} \triangleq \mathbf{F}_{\mathbf{u}}(\mathbf{u}) - S(\mathbf{u}, \xi).$$

We also define the tensor $\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{u})$ as follows:

$$[\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{u})]_{ijk} = [\nabla^2\mathbf{F}_i(\mathbf{u})]_{jk}, \quad i, j, k \in \{1, 2, \cdots, n\}$$

and define the multiplication of the tensor with a vector, $\mathbf{b} \in \mathbb{R}^n$, as

$$[\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{u})\mathbf{b}]_{ij} = \sum_{k=1}^{n} [\nabla^2\mathbf{F}_i(\mathbf{u})]_{jk}\mathbf{b}_k.$$

Then $\|\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{u})\| = \max_{1\le i\le n} \|\nabla^2\mathbf{F}_i(\mathbf{u})\|$. In this section, we analyze that the solution path guided by the stochastic homotopy tracking is closed to the path guided by the traditional homotopy tracking under certain conditions. This analysis is performed for Euler's prediction in Theorem 3.1 and for Newton's correction in Theorem 3.2.

Theorem 3.1 *(*Euler's Prediction*). Suppose* $\mathbf{u}_0$ *and* $\tilde{\mathbf{u}}_0$ *are the start points for the original system* $\mathbf{F}$ *and the stochastic system* $\tilde{\mathbf{F}}$ *respectively. If we have the following assumptions*

- $\mathbf{F}_{\mathbf{u}}$ *and* $\tilde{\mathbf{F}}_{\mathbf{u}}$ *are invertible and differentiable and*

$$\|\mathbf{F}_{\mathbf{u}}\| \le L_{\mathbf{u}}, \|\mathbf{F}_{\mathbf{u}}^{-1}\| \le M_{\mathbf{u}} \text{ and } \|\tilde{\mathbf{F}}_{\mathbf{u}}^{-1}\| \le M_{\mathbf{u}};$$

- $\nabla\mathbf{F}_{\mathbf{u}}$, $\nabla\tilde{\mathbf{F}}_{\mathbf{u}}$ *are continuous;*
- $\mathbf{F}_p$ *and* $\tilde{\mathbf{F}}_p$ *are differentiable and* $\|\mathbf{F}_p\| \le M_p$*;*
- $\nabla\mathbf{F}_p$ *is continuous and* $\|\nabla\mathbf{F}_p\| \le L_p$*,*

then we have

$$\|\mathbb{E}(\mathbf{u}_N - \tilde{\mathbf{u}}_N)\|^2 \le CS_1\|\mathbb{E}(\mathbf{u}_0 - \tilde{\mathbf{u}}_0)\|^2 + CS_2\frac{m^2}{n^2} + \mathcal{O}(\frac{m^2\Delta p}{n^2}), \tag{9}$$

where CS_1 *and* CS_2 *are constants.*

Proof We compare the predictor step of the traditional and the stochastic homotopy tracking at $p = p_{k-1}$ and obtain

$$\begin{aligned} \mathbf{u}_k &= \mathbf{u}_{k-1} + \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\mathbf{F}_p(\mathbf{u}_{k-1})\Delta p, \\ \tilde{\mathbf{u}}_k &= \tilde{\mathbf{u}}_{k-1} + \tilde{\mathbf{F}}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1})\tilde{\mathbf{F}}_p(\tilde{\mathbf{u}}_{k-1})\Delta p, \end{aligned} \tag{10}$$

which implies

$$\mathbf{u}_k - \tilde{\mathbf{u}}_k = \mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1} + \mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k)\Delta p, \tag{11}$$

where $\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k) = \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\mathbf{F}_p(\mathbf{u}_{k-1}) - \tilde{\mathbf{F}}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1})\tilde{\mathbf{F}}_p(\tilde{\mathbf{u}}_{k-1})$. Then by taking the expectation with respect to ξ, we have

$$\begin{aligned}
&\|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\|^2 = \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}) + \mathbb{E}\big(\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k)\big)\Delta p\|^2 \\
&\quad \le \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2 + \|\mathbb{E}\big(\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k)\big)\|^2\Delta p^2 + 2\|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|\|\mathbb{E}\big(\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k)\big)\|\Delta p \\
&\quad \le (1 + \Delta p)\|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2 + \|\mathbb{E}\big(\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k)\big)\|^2(\Delta p + \Delta p^2).
\end{aligned} \tag{12}$$

Moreover, by Taylor's theorem, there exists $\mathbf{t}_{k-1}$ such that

$$\mathbf{F}_{\mathbf{u}}(\mathbf{u}_{k-1}) = \mathbf{F}_{\mathbf{u}}(\tilde{\mathbf{u}}_{k-1}) + \nabla\mathbf{F}_{\mathbf{u}}(\mathbf{t}_{k-1}) \cdot (\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}). \tag{13}$$

Therefore, we have

$$\begin{aligned}
\mathbf{FP}(\mathbf{u}_k, \tilde{\mathbf{u}}_k) &= \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\big[\mathbf{F}_p(\mathbf{u}_{k-1}) - \mathbf{F}_{\mathbf{u}}(\mathbf{u}_{k-1})\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1}, \xi_k)\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}, \xi_k)\big] \\
&= \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\Big[\mathbf{F}_p(\mathbf{u}_{k-1}) - \Big((\mathbf{F}_{\mathbf{u}} + S)(\tilde{\mathbf{u}}_{k-1}, \xi_k) + \nabla\mathbf{F}_{\mathbf{u}}(\mathbf{t}_{k-1}) \cdot (\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\Big) \\
&\quad \mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1}, \xi_k)\big(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}) - C(\tilde{\mathbf{u}}_{k-1}, \xi_k)\big)\Big] \\
&= \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})[\mathbf{F}_p(\mathbf{u}_{k-1}) - \mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}) + R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k)],
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
&R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k) \\
&\quad = C(\tilde{\mathbf{u}}_{k-1}, \xi_k) - S(\tilde{\mathbf{u}}_{k-1}, \xi_k)\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1}, \xi_k)(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}) - C(\tilde{\mathbf{u}}_{k-1}, \xi_k)) \\
&\qquad - \nabla\mathbf{F}_{\mathbf{u}}(\mathbf{t}_{k-1}) \cdot (\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1}, \xi_k)(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}) - C(\tilde{\mathbf{u}}_{k-1}, \xi_k)).
\end{aligned}$$

Moreover, there exists $\mathbf{s}_{k-1}$ such that

$$\mathbf{F}_p(\mathbf{u}_{k-1}) = \mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1}) + \nabla\mathbf{F}_p(\mathbf{s}_{k-1})(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}),$$

then Eq. (14) becomes

$$\begin{aligned}
&\|\mathbb{E}(\mathbf{FP}(\mathbf{u}_{k-1}, \tilde{\mathbf{u}}_{k-1}))\|^2 \\
&\quad = \|\mathbb{E}(\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\nabla\mathbf{F}_p(\mathbf{s}_{k-1})(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})) + \mathbb{E}(\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k))\|^2 \\
&\quad \le 2\|\mathbb{E}(\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\nabla\mathbf{F}_p(\mathbf{s}_{k-1})(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}))\|^2 + 2\|\mathbb{E}(\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k))\|^2) \\
&\quad \le 2\|\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\|^2\|\nabla\mathbf{F}_p(\cdot)\|^2\|\mathbb{E}((\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}))\|^2 + 2\|\mathbf{F}_{\mathbf{u}}^{-1}(\mathbf{u}_{k-1})\|^2\|\mathbb{E}(R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k))\|^2
\end{aligned} \tag{15}$$

Since $\mathbf{F}_{\mathbf{u}}^{-1}$ and $\nabla\mathbf{F}_p$ are bounded, we have

$$\|\mathbb{E}(\mathbf{FP}(\mathbf{u}_{k-1}, \tilde{\mathbf{u}}_{k-1}))\|^2 \le 2M_{\mathbf{u}}^2L_p^2\|\mathbb{E}((\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1}))\|^2 + 2M_{\mathbf{u}}^2\|\mathbb{E}(R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k))\|^2 \tag{16}$$

Next we estimate $R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k)$:

$$
\begin{aligned}
&\|\mathbb{E}(R(\tilde{\mathbf{u}}_{k-1},\mathbf{u}_{k-1},\xi_k))\|^2 = \|\mathbb{E}_{\xi_0\xi_1\ldots\xi_{k-1}}(\mathbb{E}_{\xi_k}R(\tilde{\mathbf{u}}_{k-1},\mathbf{u}_{k-1},\xi_k))\|^2\\
&= \Big\|\mathbb{E}\big(\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}} R(\tilde{\mathbf{u}}_{k-1},\mathbf{u}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))\big)\Big\|^2\\
&\le 3\Big(\underbrace{\big\|\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(C(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J})))\big\|^2}_{A_1}\\
&\quad+\underbrace{\big\|\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(S(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1})-C(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))))\big\|^2}_{A_2}\\
&\quad+\underbrace{\big\|\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{t}_{k-1})\cdot(\mathbf{u}_{k-1}-\tilde{\mathbf{u}}_{k-1})\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1})}\\
&\quad\underbrace{-C(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J})))\big\|^2}_{A_3}\Big).
\end{aligned}
\tag{17}
$$

Since

$$
\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}} C(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J})) = \sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\sum_{i\in\mathcal{I}}\frac{\partial F_i}{\partial p}\mathbf{e}_i = C_n^m C_{n-1}^{m-1}\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}),
$$

we have

$$
A_1 = \|\frac{C_n^m C_{n-1}^{m-1}}{(C_n^m)^2}\mathbb{E}(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}))\|^2 = \|\frac{C_{n-1}^{m-1}}{C_n^m}\mathbb{E}(\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1}))\|^2 \le \frac{m^2}{n^2}M_p^2.
$$

Moreover, we have

$$
\begin{aligned}
A_2 &\le \|\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(S(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))\|^2\|\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1},\cdot)\|^2\|\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1})\|^2\\
&\le \frac{M_{\mathbf{u}}^2M_p^2}{(C_n^m)^4}\Big\|\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(S(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))\Big\|^2,
\end{aligned}
$$

By the definition of $S(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J}))$, we have

$$
\begin{aligned}
\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\mathbb{E}(S(\tilde{\mathbf{u}}_{k-1},\xi_k=(\mathcal{I},\mathcal{J})) &= \sum_{\mathcal{J}\in\mathcal{M}}\mathbb{E}\Big(\sum_{\mathcal{I}\in\mathcal{M}}\big(\sum_{i\in\mathcal{I}}\mathbf{e}_i\frac{\partial F_i}{\partial\mathbf{u}}(\mathbf{u})-\sum_{i\in\mathcal{I},j\in\mathcal{J}}E_{ij}\big)\Big)\\
&= C_n^m C_{n-1}^{m-1}\mathbb{E}(\mathbf{F}_{\mathbf{u}}(\tilde{\mathbf{u}}_{k-1})) - C_{n-1}^{m-1}C_{n-1}^{m-1}\mathbf{E},
\end{aligned}
$$

where $\mathbf{E}$ is the all-ones matrix. Therefore

$$
A_2 \le \frac{m^2}{n^2}(L_{\mathbf{u}}+1)^2M_{\mathbf{u}}^2M_p^2.
$$

Similarly, we have

$$
\begin{aligned}
A_3 &\le \Big\|\frac{1}{(C_n^m)^2}\sum_{\mathcal{I},\mathcal{J}\in\mathcal{M}}\nabla\mathbf{F}_{\mathbf{u}}(\mathbf{t}_{k-1})\|^2\|\mathbb{E}(\mathbf{u}_{k-1}-\tilde{\mathbf{u}}_{k-1})\|^2\|\mathbf{F}_{\mathbf{u}}^{-1}(\tilde{\mathbf{u}}_{k-1},\cdot)\|^2\|\mathbf{F}_p(\tilde{\mathbf{u}}_{k-1})\Big\|^2\\
&\le L_{\mathbf{u}}^2M_{\mathbf{u}}^2M_p^2\|\mathbb{E}(\mathbf{u}_{k-1}-\tilde{\mathbf{u}}_{k-1})\|^2.
\end{aligned}
$$

Then Eq. (17) becomes

$$\|\mathbb{E}(R(\tilde{\mathbf{u}}_{k-1}, \mathbf{u}_{k-1}, \xi_k))\|^2 \leq \frac{m^2}{n^2} C_1 + C_2 \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2, \tag{18}$$

where $C_1 = 3(M_{\mathbf{u}}^2 M_p^2 (L_{\mathbf{u}} + 1)^2 + M_p^2)$ and $C_2 = 3L_{\mathbf{u}}^2 M_{\mathbf{u}}^2 M_p^2$.

Then we get the estimate below

$$\|\mathbb{E}(\mathbf{FP}(\mathbf{u}_{k-1}, \tilde{\mathbf{u}}_{k-1}))\|^2 \leq \frac{m^2}{n^2} M_1 + M_2 \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2,$$

where

$$M_1 = 2M_{\mathbf{u}}^2 C_1 \text{ and } M_2 = 2M_{\mathbf{u}}^2 L_p^2 + C_2.$$

Plugging the above results into (12), we have

$$\begin{aligned}
&\|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\|^2 \\
\leq &(1 + \Delta p)\|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2 + (\frac{m^2 M_1}{n^2} + M_2 \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2))(\Delta p + \Delta p^2) \\
\leq &\underbrace{(1 + \Delta p + M_2(\Delta p + \Delta p^2))}_{\tilde{M}_1} \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2 + \underbrace{\frac{m^2}{n^2} M_1 (\Delta p + \Delta p^2)}_{\tilde{M}_2}.
\end{aligned} \tag{19}$$

which implies

$$\begin{aligned}
\|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\|^2 \leq &\tilde{M}_1 \|\mathbb{E}(\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_{k-1})\|^2 + \tilde{M}_2 \\
\leq &\tilde{M}_1^2 \|\mathbb{E}(\mathbf{u}_{k-2} - \tilde{\mathbf{u}}_{k-2})\|^2 + \tilde{M}_1 \tilde{M}_2 + \tilde{M}_2 \\
\leq &\tilde{M}_1^k \|\mathbb{E}(\mathbf{u}_0 - \tilde{\mathbf{u}}_0)\|^2 + (1 + \tilde{M}_1 + \cdots + \tilde{M}_1^{k-1}) \tilde{M}_2 \\
= &\tilde{M}_1^k \|\mathbb{E}(\mathbf{u}_0 - \tilde{\mathbf{u}}_0)\|^2 + \frac{1 - \tilde{M}_1^k}{1 - \tilde{M}_1} \tilde{M}_2.
\end{aligned} \tag{20}$$

Then we obtain the estimate of $\frac{1-\tilde{M}_1^k}{1-\tilde{M}_1} \tilde{M}_2$ as follows

$$\begin{aligned}
&\frac{1 - \tilde{M}_1^k}{1 - \tilde{M}_1} \tilde{M}_2 \leq \frac{e^{(1+M_2)(b-a)} - 1}{(1 + M_2)\Delta p + M_2 \Delta p^2} \tilde{M}_2 \\
\leq &\frac{e^{(1+M_2)(b-a)} - 1}{(1 + M_2)\Delta p} (1 - \frac{M_2}{1 + M_2} \Delta p + \mathcal{O}(\Delta p^2)) \frac{m^2 M_1}{n^2} (\Delta p + \Delta p^2) \\
\leq &\frac{m^2 M_1}{n^2} \frac{e^{(1+M_2)(b-a)} - 1}{1 + M_2} + \mathcal{O}(\frac{m^2 \Delta p}{n^2})
\end{aligned}$$

Thus, Eq. (20) becomes

$$\|\mathbb{E}(\mathbf{u}_N - \tilde{\mathbf{u}}_N)\|^2 \leq \underbrace{e^{(1+M_2)(b-a)}}_{CS_1} \|\mathbb{E}(\mathbf{u}_0 - \tilde{\mathbf{u}}_0)\|^2 + \underbrace{M_1 \frac{e^{(1+M_2)(b-a)} - 1}{1 + M_2}}_{CS_2} \frac{m^2}{n^2} + \mathcal{O}(\frac{m^2 \Delta p}{n^2}).$$

□

Remark 1 For large-scale nonlinear parametric problems, when n is large, the error caused by the stochastic homotopy tracking becomes very small due to the $O(\frac{1}{n^2})$ estimate for any

given m. Therefore, the Euler's prediction of the stochastic homotopy tracking stays closed to the prediction by the traditional homotopy tracking.

Theorem 3.2 *(*Newton's correction*). Suppose* $\mathbf{u}_k^i$ *and* $\tilde{\mathbf{u}}_k^i$ *are* i*-th Newton's iterations for solving* $\mathbf{F}(\mathbf{u}, p_k) = 0$ *and* $\tilde{\mathbf{F}}(\mathbf{u}, p_k) = 0$ *respectively. If we have the following assumptions*

- $\mathbf{F}_\mathbf{u}$ *and* $\tilde{\mathbf{F}}_\mathbf{u}$ *are invertible and differentiable and*
$$\|\mathbf{F}_\mathbf{u}^{-1}\| \le M_\mathbf{u} \text{ and } \|\tilde{\mathbf{F}}_\mathbf{u}^{-1}\| \le M_\mathbf{u};$$
- $\nabla\mathbf{F}_\mathbf{u}$, $\nabla\tilde{\mathbf{F}}_\mathbf{u}$ *are continuous and*
$$\|\nabla\mathbf{F}_\mathbf{u}\| \le K_\mathbf{u} \text{ and } \|\nabla\tilde{\mathbf{F}}_\mathbf{u}\| \le K_\mathbf{u};$$
- *The initial guesses* $\mathbf{u}_k^0$ *and* $\tilde{\mathbf{u}}_k^0$ *are in a small neighborhood of the real solutions* $\mathbf{u}_k$ *and* $\tilde{\mathbf{u}}_k$,

then we have

$$\lim_{i\to\infty} \|\mathbb{E}(\mathbf{u}_k^i - \tilde{\mathbf{u}}_k^i)\| \le \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\|. \tag{21}$$

Proof We consider the i-th iteration of Newton's correction for $\mathbf{F}(\mathbf{u}, p_k) = 0$ and $\tilde{\mathbf{F}}(\mathbf{u}, p_k) = 0$. There exists $\mathbf{t}_k$ and $\tilde{\mathbf{t}}_k$ such that the following Taylor expansions hold

$$\begin{aligned}
0 = \mathbf{F}(\mathbf{u}_k, p_k) &= \mathbf{F}(\mathbf{u}_k^i) + \mathbf{F}_\mathbf{u}(\mathbf{u}_k^i)(\mathbf{u}_k - \mathbf{u}_k^i) + \frac{1}{2}(\mathbf{u}_k - \mathbf{u}_k^i)^T\nabla\mathbf{F}_\mathbf{u}(\mathbf{t}_k)(\mathbf{u}_k - \mathbf{u}_k^i),\\
0 = \tilde{\mathbf{F}}(\tilde{\mathbf{u}}_k, p_k) &= \tilde{\mathbf{F}}(\tilde{\mathbf{u}}_k^i) + \tilde{\mathbf{F}}_\mathbf{u}(\tilde{\mathbf{u}}_k^i)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i) + \frac{1}{2}(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i)^T\nabla\tilde{\mathbf{F}}_\mathbf{u}(\tilde{\mathbf{t}}_k)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i).
\end{aligned}$$

Thus the Newton's schemes are re-written as

$$\begin{aligned}
\mathbf{u}_k^{i+1} &= \mathbf{u}_k^i - \mathbf{F}_\mathbf{u}^{-1}(\mathbf{u}_k^i)\mathbf{F}(\mathbf{u}_k^i) = \mathbf{u}_k + \frac{1}{2}\mathbf{F}_\mathbf{u}^{-1}(\mathbf{u}_k^i)(\mathbf{u}_k - \mathbf{u}_k^i)^T\nabla\mathbf{F}_\mathbf{u}(\mathbf{t}_k)(\mathbf{u}_k - \mathbf{u}_k^i),\\
\tilde{\mathbf{u}}_k^{i+1} &= \tilde{\mathbf{u}}_k^i - \tilde{\mathbf{F}}_\mathbf{u}^{-1}(\tilde{\mathbf{u}}_k^i)\tilde{\mathbf{F}}(\tilde{\mathbf{u}}_k^i) = \tilde{\mathbf{u}}_k + \frac{1}{2}\tilde{\mathbf{F}}_\mathbf{u}^{-1}(\tilde{\mathbf{u}}_k^i)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i)^T\nabla\tilde{\mathbf{F}}_\mathbf{u}(\tilde{\mathbf{t}}_k)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\mathbb{E}(\mathbf{u}_k^{i+1} - \tilde{\mathbf{u}}_k^{i+1})\| =& \Big\|\mathbb{E}\Big(\big((\mathbf{u}_k^i - \mathbf{F}_\mathbf{u}^{-1}(\mathbf{u}_k^i)\mathbf{F}(\mathbf{u}_k^i)\big) - \big(\tilde{\mathbf{u}}_k^i - \tilde{\mathbf{F}}_\mathbf{u}^{-1}(\tilde{\mathbf{u}}_k^n)\tilde{\mathbf{F}}(\tilde{\mathbf{u}}_k^n)\big)\big)\Big)\Big\|\\
=& \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k) + \mathbb{E}(\frac{1}{2}\mathbf{F}_\mathbf{u}^{-1}(\mathbf{u}_k^i)(\mathbf{u}_k - \mathbf{u}_k^i)^T\nabla\mathbf{F}_\mathbf{u}(\mathbf{t}_k)(\mathbf{u}_k - \mathbf{u}_k^i))\\
&- \mathbb{E}(\frac{1}{2}\tilde{\mathbf{F}}_\mathbf{u}^{-1}(\tilde{\mathbf{u}}_k^i)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i)^T\nabla\tilde{\mathbf{F}}_\mathbf{u}(\tilde{\mathbf{t}}_k)(\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i))\|\\
\le& \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\| + \frac{1}{2}\|\mathbf{F}_\mathbf{u}^{-1}(\mathbf{u}_k^i)\|\|\nabla\mathbf{F}_\mathbf{u}(\mathbf{t}_k)\|\mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^i\|^2)\\
&+ \frac{1}{2}\|\tilde{\mathbf{F}}_\mathbf{u}^{-1}(\tilde{\mathbf{u}}_k^i)\|\|\nabla\tilde{\mathbf{F}}_\mathbf{u}(\tilde{\mathbf{t}}_k)\|\mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i\|^2)\\
\le& \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\| + M_\mathbf{u}K_\mathbf{u}\big(\mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^i\|^2) + \mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i\|^2)\big).
\end{aligned} \tag{22}$$

Due to the local assumption of the initial guesses, then we have the quadratic convergence of Newton's method, namely,

$$\begin{aligned}
\mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^i\|) &\le \alpha\mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^{i-1}\|^2),\\
\mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i\|) &\le \tilde{\alpha}\mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^{i-1}\|^2).
\end{aligned} \tag{23}$$

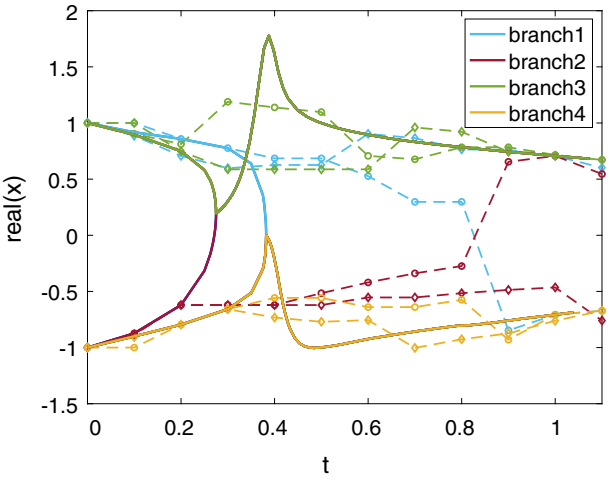

Fig. 2 An illustration of the stochastic homotopy tracking method for tracking the solution path $x(t)$ of (24) on four solution branches. The solid lines are for the traditional homotopy tracking while the dashed lines are for stochastic homotopy tracking

Therefore

$$\begin{aligned}\|\mathbb{E}(\mathbf{u}_k^{i+1} - \tilde{\mathbf{u}}_k^{i+1})\| \leq & \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\| + M_{\mathbf{u}} K_{\mathbf{u}} \big(\alpha \mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^{i-1}\|^4) + \tilde{\alpha} \mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i\|^4)\big) \\ \leq & \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\| + M_{\mathbf{u}} K_{\mathbf{u}} \big(\alpha^i \mathbb{E}(\|\mathbf{u}_k - \mathbf{u}_k^0\|^{2^{i+1}}) + \tilde{\alpha}^n \mathbb{E}(\|\tilde{\mathbf{u}}_k - \tilde{\mathbf{u}}_k^i\|^{2^{i+1}})\big).\end{aligned}$$

By taking the limit on both sides, we have

$$\lim_{i \to \infty} \|\mathbb{E}(\mathbf{u}_k^i - \tilde{\mathbf{u}}_k^i)\| \leq \|\mathbb{E}(\mathbf{u}_k - \tilde{\mathbf{u}}_k)\|.$$

□

Remark 2 The difference of Newton's corrections between the traditional and the stochastic homotopy tracking is bounded by the difference of the solutions between the original and the stochastic systems which is pretty small for large scale systems. Thus Newton's corrections by two different homotopy tracking algorithms are near each other.

4 Numerical Examples

In this section, we compare the stochastic homotopy tracking with the traditional homotopy tracking on the Matlab platform. We use the stopping criteria of $\Delta p < 10^{-7}$ for the traditional homotopy tracking method to detect the bifurcation points.

4.1 Example 1

We first consider a homotopy setup for solving a system of polynomial equations with the total degree start system, namely,

$$H(x, y, z; t) = t \begin{bmatrix} x^2 + y^2 + z^2 - 1 \\ x^2 - y^2 - z^2 \\ x + y + z \end{bmatrix} + (1 - t) \begin{bmatrix} x^2 - 1 \\ y^2 - 1 \\ z - 1 \end{bmatrix} = 0. \tag{24}$$

Table 1 Timing comparison between traditional and stochastic homotopy tracking methods on different branches shown in Fig. 2

	Traditional homotopy tracking	Stochastic homotopy tracking
Branch 1	1.05s (259 steps)	0.24s (11 steps)
Branch 2	0.59s (221 steps)	0.24s (11 steps)
Branch 3	0.91s (246 steps)	0.17s (11 steps)
Branch 4	0.84s (237 steps)	0.18s (11 steps)

When $t = 0$, the solutions of $H(x, y, z; 0) = 0$ are known explicitly. The solutions of the target system, $H(x, y, z; 1) = 0$, are revealed by tracking t from 0 to 1 on the complex field. There are four solution paths needed to track from 0 to 1 for $\mathbf{u} = [x, y, z]^T$ shown in Fig. 2. The solid lines indicate the solution path of $x(t)$ for the traditional homotopy tracking, while the dashed lines represent the solution paths guided by stochastic homotopy tracking.

The timing data is compared between two tracking methods is shown in Table 1 with $\Delta t = 0.1$ which clearly demonstrates that the stochastic homotopy tracking method is more efficient with fewer steps from $t = 0$ to $t = 1$.

4.2 Example 2

We consider the following 1D nonlinear boundary value problem.

$$\begin{cases} u_{xx} = u^2(u^2 - p), \\ u_x(0) = 0, u(1) = 0, \end{cases} \tag{25}$$

where p is the parameter. The multiple solutions become more as p gets larger. Therefore, turning points happen when p is tracked. We discretize (25) by using the finite difference method and have the following discretized polynomial system

$$\mathbf{F}(\mathbf{u}, p) := \begin{pmatrix} \frac{1}{h^2}(\mathbf{u}_1 - 2\mathbf{u}_1 + \mathbf{u}_2) - \mathbf{u}_1^2(\mathbf{u}_1^2 - p) \\ \frac{1}{h^2}(\mathbf{u}_{i-1} - 2\mathbf{u}_i + \mathbf{u}_{i+1}) - \mathbf{u}_i^2(\mathbf{u}_i^2 - p) \\ \frac{1}{h^2}(\mathbf{u}_{n-2} - 2\mathbf{u}_{n-1}) - \mathbf{u}_{n-1}^2(\mathbf{u}_{n-1}^2 - p) \end{pmatrix} = 0. \tag{26}$$

where $h = \frac{1}{n}$, $\mathbf{u} \in \mathbb{R}^{n-1}$ and $\mathbf{u}_i = u(\frac{i}{n})$ for $i = 1, 2, \cdots, n-1$. We track the parameter p from 14 down to 2 with $\Delta p = -1$ for one solution path with a turning point shown in Fig 3. Since the lower solution branch is close to the constant solution branch (the red line in Fig. 3, the stochastic homotopy tracking just switches to the constant solution branch when it is close to the turning point. Moreover, the stochastic homotopy tracking is much efficient than the traditional method by comparing the average tracking time shown in Table 2 for different grid points n. For the upper solution branch, since no nearby solution branch exists, the stochastic homotopy tracking has to deal with a stochastic system with a large perturbation, namely increasing m in **Algorithm 1**.

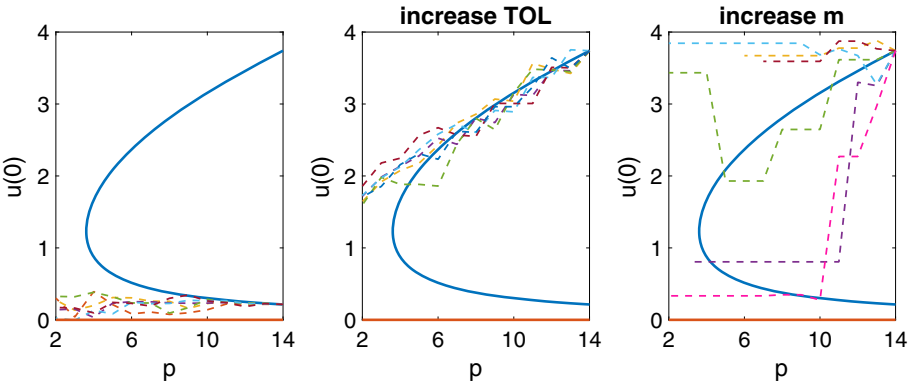

Fig. 3 An illustration of stochastic homotopy tracking for tracking (25) with respect to p from 14 to 2. The lower solution branch is switched to the constant solution branch (**Left**); The upper solution branch needs a large TOL (**Middle**) or a large m (**Right**) in **Algorithm 1**

Table 2 Comparison between the traditional and the stochastic homotopy tracking with different number of grid points n

n	Traditional	Stochastic
10	0.027s (24 steps)	0.013s (12 steps)
20	0.051s (22 steps)	0.022s (12 steps)
40	0.141s (30 steps)	0.076s (12 steps)
80	0.530s (29 steps)	0.272s (12 steps)

4.3 Example 3

Last we consider the Schnakenberg model which is a system of partial differential equations shown below [12]:

$$\begin{cases} \dfrac{\partial u}{\partial t} = \Delta u + \eta(a - u + u^2 v), \\ \dfrac{\partial v}{\partial t} = d\Delta v + \eta(b - u^2 v), \end{cases} \tag{27}$$

where u is an activator and v is a substrate. The steady-state system of (27) with non-flux boundary condition has been well-studied in [12] and shown multiple steady-state solutions and the bifurcation structure to the diffusion parameter d. In this example, we consider the discretized steady-state system on a 1D domain $x \in [0, 1]$ with no-flux boundary conditions:

$$\mathbf{F}(\mathbf{u}, \mathbf{v}, d) := \begin{pmatrix} \frac{1}{h^2}(2\mathbf{u}_2 - 2\mathbf{u}_1) + \eta(a - \mathbf{u}_1 + \mathbf{u}_1^2\mathbf{v}_1) \\ \frac{1}{h^2}(\mathbf{u}_{i-1} - 2\mathbf{u}_i + \mathbf{u}_{i+1}) + \eta(a - \mathbf{u}_i + \mathbf{u}_i^2\mathbf{v}_i) \\ \frac{1}{h^2}(2\mathbf{u}_n - 2\mathbf{u}_{n+1}) + \eta(a - \mathbf{u}_{n+1} + \mathbf{u}_{n+1}^2\mathbf{v}_{n+1}) \\ \frac{d}{h^2}(2\mathbf{v}_2 - 2\mathbf{v}_1) + \eta(b - \mathbf{u}_1^2\mathbf{v}_1) \\ \frac{d}{h^2}(\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}) + \eta(b - \mathbf{u}_i^2\mathbf{v}_i) \\ \frac{d}{h^2}(2\mathbf{v}_n - 2\mathbf{v}_{n+1}) + \eta(b - \mathbf{u}_{n+1}^2\mathbf{v}_{n+1}) \end{pmatrix} = 0. \tag{28}$$

where $h = \frac{1}{n}$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{n+1}$ with $\mathbf{u}_i = u(\frac{i-1}{n})$ and $\mathbf{v}_i = v(\frac{i-1}{n})$ for $i = 1, 2, \cdots, n+1$. We introduce ghost points $\mathbf{u}_0$, $\mathbf{v}_0$, $\mathbf{u}_{n+2}$, and $\mathbf{v}_{n+2}$ at $x = 0$ and $x = 1$. The nonflux boundary conditions imply that $\mathbf{u}_0 = \mathbf{u}_2$, $\mathbf{v}_0 = \mathbf{v}_2$,$\mathbf{u}_{n+2} = \mathbf{u}_n$, and $\mathbf{v}_{n+2} = \mathbf{v}_n$. We choose $a =$

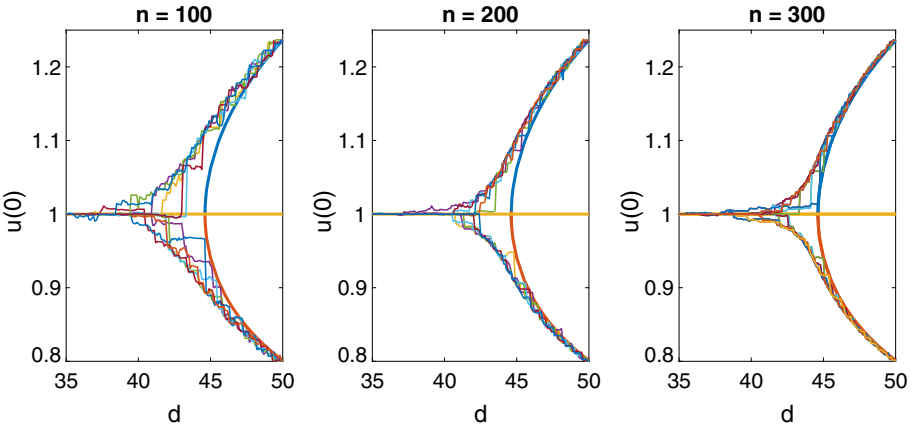

Fig. 4 Traditional and stochastic homotopy tracking methods with different number of grid points

Table 3 Comparison between traditional and stochastic homotopy tracking with different number of grid points n and different step-sizes Δd

		Lower branch		Upper branch	
n	Δd	Traditional	Stochastic	Traditional	Stochastic
100	-0.5	2.76s(32steps)	2.30s(31steps)	2.49s(28steps)	1.87s(31steps)
	-1	3.23s(59steps)	1.35s(16steps)	2.38s(34steps)	0.93s(16steps)
200	-0.5	12.88s(53steps)	8.83s(31steps)	10.62s(35steps)	8.93s(31steps)
	-1	9.36s(53steps)	3.08s(16steps)	7.61s(21steps)	3.88s(16steps)
300	-0.5	77.9s(90steps)	34.1s(31steps)	40.2s(34steps)	36.9s(31steps)
	-1	40.3s(90steps)	16.5s(16steps)	30.1s(34steps)	15.6s(16steps)

$1/3,\ b = 2/3,\ \eta = 50$ and track d from 50 to 35 with different number of grid points n. As shown in Fig. 4, the traditional homotopy tracking method stops near the bifurcation around $d \approx 45$ with a very small tracking stepsize. However, the stochastic homotopy tracking method can avoid the bifurcation point and track down to 35. Moreover, as n goes larger, the solution path guided by the stochastic homotopy tracking gets closer to the original path. Detailed iteration comparison between two tracking methods is shown in Table 3 for the different number of grid points n and different tracking stepsizes Δd. It clearly shows that the stochastic homotopy tracking method becomes more efficient compared to the traditional one as the size of the system gets larger.

5 Conclusion

By taking the path tracking from a stochastic differential equation point of view, we have developed a stochastic homotopy path tracking algorithm that perturbs the nonlinear parametric system by randomly removing m equation each step. In this paper, we also proved that the solution path guided by the stochastic homotopy algorithm is nearby the original solution path but can avoid the singularities during the tracking. Several numerical examples are used to demonstrate the efficiency of this new method through comparison with the traditional homotopy tracking method. However, the efficiency of the stochastic homotopy

tracking depends on the solution landscaping of the original system: if there exists a nearby solution path for bifurcation points, then the stochastic homotopy tracking can switch to the nearby solution paths and keep tracking. Otherwise, the computational cost might be still expensive since it keeps solving stochastic systems by increasing perturbations. In the future, we will improve the efficiency of stochastic homotopy tracking further by exploring the optimal perturbation.

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Funding This research is supported by NSF via DMS-1818769. The authors declare that there is no conflict of interest.

References

1. Allgower, E.L., Georg, K.: Introduction to numerical continuation methods SIAM (2003)
2. Arnold, L.: Stochastic differential equations, New York (1974)
3. Bates, D., Brake, D., Niemerg, M.: Paramotopy: Parameter homotopies in parallel. In: International Congress on Mathematical Software, pp. 28–35. Springer (2018)
4. Bates, D., Hauenstein, J., Sommese, A., Wampler, C. Bertini: Software for numerical algebraic geometry. J Softw Algeb Geom **3**(1), 5–10 (2006)
5. Bates, D., Hauenstein, J., Sommese, A., Wampler, C.: Adaptive multiprecision path tracking. SIAM J Num Anal **46**(2), 722–746 (2008)
6. Bates, D.J., Hauenstein, J.D., Sommese, A., Wampler, C.: Numerically solving polynomial systems with Bertini,. 25. SIAM (2013)
7. Bates, D.J., Hauenstein, J.D., Sommese, A.: A parallel endgame. Contemp Math **556**, 25–35 (2011)
8. Cauwenberghs, G.: A fast stochastic error-descent algorithm for supervised learning and optimization. In: Advances in neural information processing systems, pp. 244–251 (1993)
9. Chen, Q., Hao, W.: A homotopy training algorithm for fully connected neural networks. *Submitted*
10. Hao, W., Harlim, J.: An equation-by-equation method for solving the multidimensional moment constrained maximum entropy problem. Commun Appl Math Comput Sci **13**(2), 189–214 (2018)
11. Hao, W., Hauenstein, J., Shu, C.-W., Sommese, A., Xu, Z., Zhang, Y.-T.: A homotopy method based on weno schemes for solving steady state problems of hyperbolic conservation laws. J Comput Phys **250**, 332–346 (2013)
12. Hao, W., Xue, C.: Spatial pattern formation in reaction-diffusion models: a computational approach. J Math Biol **80**(1), 521–543 (2020)
13. Hao, W., Zheng, C.: An adaptive homotopy method for computing bifurcations of nonlinear equations. *Submitted*
14. Leykin, A.: Numerical algebraic geometry. J Softw Algeb Geom **3**(1), 5–10 (2011)
15. Li, Y., Lu, J., Wang, Z.: Coordinatewise descent methods for leading eigenvalue problem. SIAM J Sci Comput **41**(4), A2681–A2716 (2019)
16. Nesterov, Yu.: Efficiency of coordinate descent methods on huge-scale optimization problems. SIAM J Optim **22**(2), 341–362 (2012)
17. Nguyen, L., Schmidt, H., Von Haeseler, A., Minh, B.: Iq-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol **32**(1), 268–274 (2015)
18. Wampler, C., Sommese, A.: The Numerical solution of systems of polynomials arising in engineering and science. World Scientific, Singapore (2005)
19. Wang, Y., Hao, W., Lin, G.: Two-level spectral methods for nonlinear elliptic equations with multiple solutions. SIAM J Sci Comput **40**(4), B1180–B1205 (2018)
20. Yang, Y., Hao, W.: Convergence of a homotopy finite element method for computing steady states of burgers' equation. Mathematical Modelling and Numerical Analysis, ESAIM (2018)