

# Chetaev Instability Framework for Kinetostatic Compliance-Based Protein Unfolding

Alireza Mohammadi and Mark W. Spong

**Abstract**—Understanding the process of protein unfolding plays a crucial role in various applications such as design of folding-based protein engines. Using the well-established kinetostatic compliance (KCM)-based method for modeling of protein conformation dynamics and a recent nonlinear control theoretic approach to KCM-based protein folding, this paper formulates protein unfolding as a destabilizing control analysis/synthesis problem. In light of this formulation, it is shown that the Chetaev instability framework can be used to investigate the KCM-based unfolding dynamics. In particular, a Chetaev function for analysis of unfolding dynamics under the effect of optical tweezers and a class of control Chetaev functions for synthesizing control inputs that elongate protein strands from their folded conformations are presented. Based on the presented control Chetaev function, an unfolding input is derived from the Artstein-Sontag universal formula and the results are compared against optical tweezer-based unfolding.

## I. INTRODUCTION

Protein molecules make conformational transitions in-between two or more native conformations to perform a plethora of crucial biological functions [1]. These transitions usually involve two distinct processes of folding and unfolding. During the folding/unfolding process, a polypeptide chain folds/unfolds into/from a biologically active molecule (see Supplementary Material for an energy landscape schematic of folding/unfolding). In addition to its role for understanding the protein structures, unfolding is used for engineering materials such as hydrogels and anti-viral drugs as well as designing protein-based nanomachines [2]–[4].

Algorithmic prediction of pathways through which a folded protein molecule unfolds to its denatured state has been among the main tools for investigating the biochemical process of unfolding. The AI-based approaches to predicting the folded structure of proteins such as Google AlphaFold [5], which are rooted in pattern recognition, cannot address issues such as folding/unfolding pathway prediction. Indeed, AlphaFold’s main task has been to predict the most likely structures of folded proteins while ignoring the kinetics and stability of folding/unfolding processes (see, e.g., [6] for further details). Alternatively, dynamics-based approaches rooted in physical first-principles provide a more preferred way for predicting the folding/unfolding pathways [7].

To address the high computational burden of physics-based approaches, the framework of kinetostatic compli-

This work is supported by the National Science Foundation (NSF) under Grants 2153744 and 2153901. A. Mohammadi is with the Department of Electrical & Computer Engineering, University of Michigan-Dearborn, Dearborn, MI 48128 USA. Mark W. Spong is with the University of Texas at Dallas, Richardson, TX 75080 USA. Emails: {amohmmad@umich.edu\*, mspong@utdallas.edu}.

\*Corresponding Author: A. Mohammadi

ance method (KCM), which is based on modeling protein molecules as mechanisms with a large number of rigid nano-linkages that fold under the nonlinear effect of electrostatic and van der Waals interatomic forces, has been developed in the literature [8]–[12]. The KCM framework has been successful in investigating the effect of hydrogen bond formation on protein kinematic mobility [13] and synthesizing molecular nano-linkages [14], [15]. Furthermore, as demonstrated by [16], this framework lends itself to a nonlinear control theoretic interpretation, where entropy-loss constraints can be encoded in KCM-based folding via proper quadratic optimization-based nonlinear control techniques. Despite its success for investigating the structure of proteins, the KCM framework except for a simulation study in [17] has not been systematically utilized for studying unfolding.

In this paper, we formulate the KCM-based protein unfolding as a destabilizing control analysis/synthesis problem by utilizing the Chetaev instability framework. A systemic development of Chetaev functions (CFs) and control Chetaev functions (CCFs), which are rooted in the seminal work of N. G. Chetaev [18], has been recently initiated by Efimov and collaborators [19]–[21] and Kellett and collaborators [22]. CFs and CCFs are not only useful for studying instability of nonlinear control systems but also for synthesis of destabilizing control inputs to induce oscillations via static feedback and safety critical control applications. CFs and CCFs in this paper are employed to investigate the instability properties of the folded protein conformations under study.

**Contributions of the paper.** This paper has the following contributions. First, this paper bridges the two fields of protein unfolding in biochemistry and the Chetaev instability framework in nonlinear control theory. Second, this paper provides a nonlinear control theoretic foundation for single-molecule force-based protein denaturation. The single-molecule protein unfolding using devices such as optical tweezers [23] and atomic force microscopes (AFMs) [24] has proven to be useful in nano-/molecular-robotics [25], [26]. Third, this paper adds to the body of knowledge on KCM-based modeling for further investigation of structural properties of protein molecules.

The rest of this paper is organized as follows. First, we review the KCM framework for modeling protein molecules, the control theoretic interpretation of it, and some basics from the Chetaev instability theory in Section II. Thereafter, we model the dynamics of protein unfolding under optical tweezers in Section III. Next, a CF for analysis of optical tweezer-based unfolding and a class of CCFs for synthesizing unfolding control inputs, which elongate protein molecules,

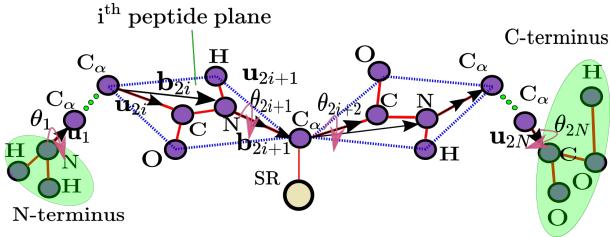


Fig. 1: The protein molecule kinematic structure.

are presented in Section IV. After presenting the simulation results that compare optical tweezer-based unfolding against a synthesized Artstein-Sontag destabilizing feedback in Section V, we conclude the paper in Section VI.

**Notation.** We let  $\mathbb{R}_+$  denote the set of all non-negative real numbers. We let  $(\cdot)^\top$  denote the vector/matrix transpose operator. Given  $\mathbf{x} \in \mathbb{R}^M$ , we let  $|\mathbf{x}| := \sqrt{\mathbf{x}^\top \mathbf{x}}$  denote the Euclidean norm of  $\mathbf{x}$ . Given  $\varepsilon \in \mathbb{R}_+$ , we let  $\mathcal{B}_x(\varepsilon)$  denote the open ball centered at  $\mathbf{x}$  with radius  $\varepsilon$ . Given two vectors  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$ , we let  $\mathbf{v}_1 \times \mathbf{v}_2$  denote their cross product.

## II. PRELIMINARIES

### A. Kinematic linkage representation of protein molecules

In this section, we review the KCM-based modeling of the main chain *in vacuo* (see [11], [12] for further details). Proteins essentially consist of peptide planes that are rigidly joined together in a long chain (see Figure 1). Central carbon atoms denoted by  $C_\alpha$  play the role of hinges in-between the amino acids. The red line segments in Figure 1 represent covalent atomic bonds. The first  $C_\alpha$  is connected to the N-terminus, and one other peptide plane. Similarly, the last  $C_\alpha$  hinges to the C-terminus, and one other peptide plane.

The backbone conformation of protein molecules can be described by using a collection of bond lengths and two sets of rotation angles around  $N - C_\alpha$  and  $C_\alpha - C$ , which are known as the dihedral angles. Hence, the vector of dihedral angles  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{2N}]^\top \in \mathbb{R}^{2N}$  represents the protein configuration with  $N - 1$  peptide planes. Along the rotation axis corresponding to each DOF of the protein molecule, a unit vector denoted by  $\mathbf{u}_j$ ,  $1 \leq j \leq 2N$ , can be considered.

In addition to the unit vectors  $\mathbf{u}_j$ , Kazerounian and collaborators [9]–[11] utilize the so-called **body vectors** to represent the spatial orientation of the peptide planes. These body vectors are denoted by  $\mathbf{b}_j$ ,  $1 \leq j \leq 2N$ , and describe the relative position of any two atoms in the chain. Using the vectors  $\mathbf{u}_j$  and  $\mathbf{b}_j$ , the conformation of a protein molecule can be completely described. In particular, after designating the zero position conformation with  $\boldsymbol{\theta} = \mathbf{0}$  corresponding to the biological reference position of the chain (see [9]), the transformations

$$\mathbf{u}_j(\boldsymbol{\theta}) = \Xi(\boldsymbol{\theta}, \mathbf{u}_j^0) \mathbf{u}_j^0, \mathbf{b}_j(\boldsymbol{\theta}) = \Xi(\boldsymbol{\theta}, \mathbf{u}_j^0) \mathbf{b}_j^0, \quad (1)$$

where

$$\Xi(\boldsymbol{\theta}, \mathbf{u}_j^0) = \prod_{r=1}^j R(\theta_r, \mathbf{u}_r^0), \quad (2)$$

describe the molecule kinematic configuration. In (2), the matrix  $R(\theta_j, \mathbf{u}_j^0) \in \text{SO}(3)$  represents the rotation about the unit vector  $\mathbf{u}_j^0$  with angle  $\theta_j$ . Having obtained the body vectors  $\mathbf{b}_j(\boldsymbol{\theta})$  from (1) and under the assumption that the N-terminus nitrogen atom is fixed at the origin, the coordinates of the backbone chain atoms in the  $k$ -th peptide plane are found from  $\mathbf{r}_i(\boldsymbol{\theta}) = \sum_{j=1}^i \mathbf{b}_j(\boldsymbol{\theta})$ .

### B. Control theoretic interpretation of KCM-based dynamics

The KCM framework in [10]–[12] builds upon the experimental fact that the inertial forces during the folding of a protein chain can be neglected in the presence of the interatomic forces (see, e.g., [10], [27]). Instead, the dihedral angles rotate under the kinetostatic influence of these forces.

To present the KCM-based protein dynamics, let us consider a polypeptide chain with  $N - 1$  peptide planes and conformation vector  $\boldsymbol{\theta} \in \mathbb{R}^{2N}$ . The molecule aggregated free energy is given by (see [12] for detailed derivations)

$$\mathcal{G}(\boldsymbol{\theta}) := \mathcal{G}^{\text{elec}}(\boldsymbol{\theta}) + \mathcal{G}^{\text{vdw}}(\boldsymbol{\theta}), \quad (3)$$

where  $\mathcal{G}^{\text{elec}}(\boldsymbol{\theta})$  and  $\mathcal{G}^{\text{vdw}}(\boldsymbol{\theta})$  represent the the electrostatic and van der Waals potential energies, respectively.

The resultant forces and torques on each of the  $N - 1$  rigid peptide links can be computed from the molecule free energy in (3) and the protein chain kinematics. These forces and torques can then be appended in the generalized force vector  $\mathcal{F}(\boldsymbol{\theta}) \in \mathbb{R}^{6N}$ . Following the steps of derivation in [11], one can utilize a proper mapping to relate  $\mathcal{F}(\boldsymbol{\theta})$  to the equivalent torque vector  $\boldsymbol{\tau}(\boldsymbol{\theta}) \in \mathbb{R}^{2N}$  acting on the dihedral angles at conformation  $\boldsymbol{\theta}$ , which is given by

$$\boldsymbol{\tau}(\boldsymbol{\theta}) = \mathcal{J}^\top(\boldsymbol{\theta}) \mathcal{F}(\boldsymbol{\theta}). \quad (4)$$

In (4), the matrix  $\mathcal{J}(\boldsymbol{\theta}) \in \mathbb{R}^{6(2N) \times 2N}$  represents the Jacobian of the protein chain at conformation  $\boldsymbol{\theta}$ . The vector  $\mathcal{F}(\boldsymbol{\theta})$  is due to the torques and forces, which arise from the interatomic electrostatic and van der Waals effects, at conformation  $\boldsymbol{\theta}$ . As it is demonstrated in [10], [11], for a protein molecule with  $N - 1$  peptide planes, the matrix  $\mathcal{J}^\top(\boldsymbol{\theta}) \in \mathbb{R}^{2N \times 6(2N)}$  has the structure

$$\mathcal{J}^\top(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{J}_1^\top(\boldsymbol{\theta}) & \mathbf{J}_1^\top(\boldsymbol{\theta}) & \cdots & \mathbf{J}_1^\top(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{J}_2^\top(\boldsymbol{\theta}) & \cdots & \mathbf{J}_2^\top(\boldsymbol{\theta}) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{2N}^\top(\boldsymbol{\theta}) \end{bmatrix}, \quad (5)$$

where

$$\mathbf{J}_k(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{u}_k(\boldsymbol{\theta}) \\ -\mathbf{u}_k(\boldsymbol{\theta}) \times \mathbf{r}_k(\boldsymbol{\theta}) \end{bmatrix} \in \mathbb{R}^6, 1 \leq k \leq 2N, \quad (6)$$

in which  $\mathbf{u}_k(\boldsymbol{\theta})$  denotes the unit vector introduced in Section II-A and  $\mathbf{r}_k(\boldsymbol{\theta})$  connects the N-terminus to either  $C_\alpha$  atom (even  $k$ ) or  $N$  atom (odd  $k$ ) on the  $k$ -th peptide plane.

Using the torque vector in (4) and starting from an initial conformation  $\boldsymbol{\theta}_0$ , the KCM-based iteration is given by  $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + h \mathbf{f}_{\text{KCM}}(\boldsymbol{\tau}(\boldsymbol{\theta}_i))$ , where  $i$  is a non-negative integer,  $h$  is a positive real constant that tunes the maximum dihedral angle rotation magnitude in each step, and  $\mathbf{f}_{\text{KCM}}(\cdot)$  is a proper mapping, which can be chosen to be the identity map in its

simplest form (see [11] for further details). The successive iteration of the KCM-based folding is performed until all of the kinetostatic torques converge to a minimum corresponding to a local minimum of the free energy. Furthermore,  $\mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$  is a locally Lipschitz continuous function in a neighborhood of a given folded conformation.

In [16], it has been demonstrated that the KCM-based iteration can be interpreted as a control synthesis problem. In particular, the work in [16] considered the following nonlinear control-affine system

$$\dot{\boldsymbol{\theta}} = \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta}) + \mathbf{u}_c, \quad (7)$$

where the Jacobian matrix  $\mathcal{J}(\boldsymbol{\theta})$  and the generalized force vector  $\mathcal{F}(\boldsymbol{\theta})$  are the same as in (4). It was demonstrated that forward Euler discretization of the closed-loop dynamics in (7) under a proper closed-loop feedback control input  $\mathbf{u}_c(\boldsymbol{\theta})$  restores the KCM-based iteration for protein folding.

### C. Chetaev instability theory preliminaries

Consider the nonlinear dynamical system

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \quad (8)$$

where  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a locally Lipschitz continuous function. We say that the equilibrium  $\mathbf{x}^* \in \mathbb{R}^n$  is unstable for (8) if for any  $\delta > 0$  there exists  $\mathbf{x}_0 \in \mathcal{B}_{\mathbf{x}^*}(\delta)$  such that  $\mathbf{x}(T_{\mathbf{x}_0}, \mathbf{x}_0) \notin \mathcal{B}_{\mathbf{x}^*}(\delta)$  for some  $T_{\mathbf{x}_0} \in \mathbb{R}_+$ , where the solution of (8) with  $\mathbf{x}(0) = \mathbf{x}_0$  is given by  $\mathbf{x}(\cdot, \mathbf{x}_0)$  on the solution maximal interval of existence.

*Theorem 2.1* ([18]): Consider the system in (8) and the equilibrium  $\mathbf{x}^*$ . Let  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  be a locally Lipschitz continuous function such that  $V(\mathbf{x}^*) = 0$ . Suppose that there exists  $\epsilon_0 > 0$  such that  $\mathcal{V}^+ \cap \mathcal{B}_{\mathbf{x}^*}(\epsilon) \neq \emptyset$  for any  $\epsilon \in (0, \epsilon_0]$ , where  $\mathcal{V}^+ := \{\mathbf{x} \in \mathcal{B}_{\mathbf{x}^*}(\epsilon) : V(\mathbf{x}) > 0\}$ . If the lower Dini derivative of  $V$  along  $\mathbf{F}$  satisfies  $D^-V(\mathbf{x})\mathbf{F}(\mathbf{x}) > 0$  for all  $\mathbf{x} \in \mathcal{V}^+$ , then the equilibrium  $\mathbf{x}^*$  is unstable for (8).

A function  $V(\cdot)$  that satisfies the conditions in Proposition 2.1 is called a CF for system (8). While the classical theorem by Chetaev [18] only provided sufficient conditions for instability, Efimov and collaborators provided the necessary part of the Chetaev's theorem in [21]. Furthermore, they developed the CCF concept as a counterpart of the Control Lyapunov function (CLF) theory [21]. Kellett and collaborators [22] extended the work in [21] to differential inclusions. See Supplementary Material for the required background on CCFs.

## III. KCM-BASED MODELING OF PROTEIN UNFOLDING UNDER OPTICAL TWEEZERS

In this section we present the dynamics of protein unfolding under the effect of an optical tweezer. We remark that the derivation of KCM-based dynamical model of protein unfolding under optical tweezers has not been presented elsewhere in the literature.

Optical tweezers can be utilized for unfolding protein molecules by chemically attaching a latex bead (microsphere) to the protein under study, where the bead can be trapped and used as a handle for stretching the molecule.

The reader is referred to Supplementary Material for further details. The optical tweezer tension force on the molecule strand can be modeled by the Hookean force [23]

$$\mathbf{F}_{\text{twz}}(\boldsymbol{\theta}) = \kappa(\boldsymbol{\theta})\mathbf{x}_{\text{twz}}, \quad (9)$$

where  $\mathbf{x}_{\text{twz}}$  represents the displacement of the microsphere from the center of the beam and  $\kappa(\boldsymbol{\theta})$  is the optical trap stiffness modulated by the protein conformation using a proper technique such as linear polarization [28], [29]. In particular, we assume that the trap stiffness can be modulated as a smooth function of  $\boldsymbol{\theta}$  in a way that at a given folded conformation  $\boldsymbol{\theta}^*$ ,  $\kappa(\boldsymbol{\theta}^*) = 0$ .

*Remark 3.1:* For a given positive integer  $m$  and a positive constant  $\kappa_0$ , a candidate optical trap modulated stiffness is

$$\kappa(\boldsymbol{\theta}) = \kappa_0|\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*) - \mathbf{r}_{\text{NC}}(\boldsymbol{\theta})|^m, \quad (10)$$

where at any conformation  $\boldsymbol{\theta}$ , the vector  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}) \in \mathbb{R}^3$  connects the nitrogen atom in the N-terminus to the carbon atom in the C-terminus (see Figure 1).

Let us consider the optical tweezer force in (9) and a protein molecule chain with  $N - 1$  peptide planes. Without loss of generality, we assume that this force is being exerted on the C-terminus of the protein molecule. Accordingly, the resultant torque applied to the last link  $\text{AA}_{\text{last}}$  of the molecule at conformation  $\boldsymbol{\theta}$ , is given by  $\mathbf{T}_{\text{twz}}(\boldsymbol{\theta}) = \sum_{a_i \in \text{AA}_{\text{last}}} \mathbf{r}_i(\boldsymbol{\theta}) \times \kappa(\boldsymbol{\theta})\mathbf{x}_{\text{twz}}$ , where  $\mathbf{r}_i(\boldsymbol{\theta})$  denotes the position of atom  $a_i$ , which is the  $i$ -th atom located in peptide plane  $\text{AA}_{\text{last}}$ . Therefore, the force-torque couple due to the effect of the optical tweezer is  $\mathbf{TF}_{\text{twz}}(\boldsymbol{\theta}) := [\mathbf{T}_{\text{twz}}(\boldsymbol{\theta})^\top, \mathbf{F}_{\text{twz}}(\boldsymbol{\theta})^\top]^\top \in \mathbb{R}^6$ .

Consequently, the optical tweezer-based force-torque couple, which is exerted on the last peptide plane of the protein molecule, gives rise to the equivalent joint torque vector

$$\mathbf{u}_{\text{unfold}}(\boldsymbol{\theta}) = \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}_{\text{twz}}(\boldsymbol{\theta}), \quad (11)$$

where  $\mathcal{F}_{\text{twz}}(\boldsymbol{\theta})$  is the generalized force vector

$$\mathcal{F}_{\text{twz}}(\boldsymbol{\theta}) = [\mathbf{0}^\top, \dots, \mathbf{0}^\top, \mathbf{TF}_{\text{twz}}(\boldsymbol{\theta})^\top]^\top \in \mathbb{R}^{6(2N)}, \quad (12)$$

which gets mapped through (11) to the vector  $\mathbf{u}_{\text{unfold}}(\boldsymbol{\theta}) \in \mathbb{R}^{2N}$  affecting the dihedral angles in the molecule.

## IV. DESTABILIZING CONTROL ANALYSIS/SYNTHESIS FOR STUDYING THE KCM-BASED PROTEIN UNFOLDING

In this section we first formulate the KCM-based protein unfolding as a destabilizing control analysis/synthesis problem for (7). Next, we present a CF for analysis of a folded conformation instability under optical tweezers. Thereafter, we introduce a class of CCFs for synthesis of destabilizing inputs that elongate protein strands from their folded conformations.

### A. Control problem formulation and a trivial solution

It is known that the KCM-based iteration when initiated in a vicinity of a stable conformation  $\boldsymbol{\theta}^*$  converges to it [10], [11]. The conformation  $\boldsymbol{\theta}^*$  corresponds to a local minimum of the aggregated free energy function  $\mathcal{G}(\cdot)$  in (3), where the equivalent torque in (4) becomes  $\boldsymbol{\tau}(\boldsymbol{\theta}^*) = \mathcal{J}^\top(\boldsymbol{\theta}^*)\mathcal{F}(\boldsymbol{\theta}^*) =$

0. Indeed, the conformation  $\boldsymbol{\theta}^*$  is an isolated and locally asymptotically stable equilibrium point for  $\dot{\boldsymbol{\theta}} = \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$  (see, e.g., [10], [11], [16]). In light of the properties of a folded conformation  $\boldsymbol{\theta}^*$ , we can state our control objective for unfolding as follows.

**Protein Unfolding Control Problem (PUCP).** Consider the KCM-based protein conformation dynamics given by (7) and a stable folded conformation given by  $\boldsymbol{\theta}^*$ . Find a closed-loop feedback control input  $\mathbf{u}_c = \mathbf{u}_{\text{unfold}}(\boldsymbol{\theta})$  for (7) such that the folded conformation  $\boldsymbol{\theta}^*$  becomes an unstable equilibrium for

$$\dot{\boldsymbol{\theta}} = \mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}) := \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta}) + \mathbf{u}_{\text{unfold}}(\boldsymbol{\theta}). \quad (13)$$

A trivial solution to PUCP is given as follows.

*Proposition 4.1:* Consider  $\mathbf{u}_{\text{unfold}}(\boldsymbol{\theta}) = -2\mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$  for (13). Under this control input the folded conformation  $\boldsymbol{\theta}^*$  will become a purely repulsing equilibrium for the resulting closed-loop dynamics given by  $\dot{\boldsymbol{\theta}} = -\mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$ .

*Proof:* The folded conformation  $\boldsymbol{\theta}^*$  is an isolated and locally asymptotically stable equilibrium point for  $\dot{\boldsymbol{\theta}} = \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$ . Therefore, we can invoke a converse argument [30] to conclude that there exists a Lyapunov function

$$V^{\text{fold}} : \mathbb{R}^{2N} \rightarrow \mathbb{R}_+, \boldsymbol{\theta} \mapsto V^{\text{fold}}(\boldsymbol{\theta}) \quad (14)$$

for these dynamics from which the local asymptotic stability of  $\boldsymbol{\theta}^*$  follows. It can be easily seen that  $V^{\text{fold}}(\cdot)$  is indeed an anti-control Lyapunov function (ALF) [20] for  $\dot{\boldsymbol{\theta}} = -\mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta})$ , where  $V^{\text{fold}}(\boldsymbol{\theta}^*) = 0$ ,  $V^{\text{fold}}(\boldsymbol{\theta}) > 0$  on  $\mathcal{N}_{\boldsymbol{\theta}^*} \setminus \{\boldsymbol{\theta}^*\}$ , where  $\mathcal{N}_{\boldsymbol{\theta}^*}$  is an open neighborhood of  $\boldsymbol{\theta}^*$ , and  $-\frac{\partial V^{\text{fold}}}{\partial \boldsymbol{\theta}} \mathcal{J}^\top(\boldsymbol{\theta})\mathcal{F}(\boldsymbol{\theta}) > 0$  for all  $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$  on an open neighborhood of  $\boldsymbol{\theta}^*$  that is a subset of  $\mathcal{N}_{\boldsymbol{\theta}^*}$ . ■

*Remark 4.2:* There are some drawbacks associated with considering  $V^{\text{fold}}(\cdot)$  as a candidate Chetaev function for protein unfolding. First, no closed-form expression for  $V^{\text{fold}}(\cdot)$  is readily available. Second,  $V^{\text{fold}}(\cdot)$  does not inherently possess the notion of direction of the unfolding forces.

### B. Instability of folded conformations under optical tweezers

In this section, we present an analysis to guarantee that the forces due to optical tweezers can be considered as solutions to PUCP formulated in Section IV-A. Such an analysis provides numerical conditions for real-time control of optical tweezer-based protein unfolding. Using the torque in (11), the unfolding conformation dynamics in (13) become

$$\dot{\boldsymbol{\theta}} = \mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}), \quad (15)$$

where  $\mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}) := \mathcal{J}^\top(\boldsymbol{\theta})\{\mathcal{F}(\boldsymbol{\theta}) + \mathcal{F}_{\text{twz}}(\boldsymbol{\theta})\}$ . Accordingly, if the protein molecule is under the effect of the forces generated by an optical tweezer as in (9), then the unfolding dynamics are given by (13), where the unfolding control torque vector  $\mathbf{u}_{\text{unfold}}(\boldsymbol{\theta})$  is given by (11) and (12). Since the trap stiffness is modulated such that  $\kappa(\boldsymbol{\theta}^*) = 0$  (see Remark 3.1), the conformation  $\boldsymbol{\theta}^*$  is an equilibrium of (15).

To analyze a folded conformation instability under the effect of optical tweezers, we are interested in deriving numerical conditions guaranteeing that PUCP objectives are met for the unfolding dynamics in (15). It is natural to expect that these conditions should depend on the direction

of the applied force in (9). Considering the fact that a protein molecule extends under an optical tweezer-based force from its folded conformation  $\boldsymbol{\theta}^*$ , we consider the candidate CF

$$C_{\text{twz}}(\boldsymbol{\theta}) = |\Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)|^2 |\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)|^2 \cos^2(\alpha_C) - (\Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*))^2, \quad (16)$$

for the unfolding dynamics in (15). At any conformation  $\boldsymbol{\theta}$ , the vector  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}) \in \mathbb{R}^3$  in (16) connects the nitrogen atom in the N-terminus to the carbon atom in the C-terminus. Furthermore, the difference vector is given by  $\Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*) := \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}) - \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)$ . Finally, the constant  $\alpha_C$  is any angle in-between 0 and  $\frac{\pi}{2}$ . Details on the geometric interpretation of this candidate CF is provided in Supplementary Material. As it will be seen in the next section (Propositions 4.7 and 4.8), this candidate CF is also a CCF. The proof of the next lemma is given in Supplementary Material.

*Lemma 4.3:* Consider the candidate CF  $C_{\text{twz}} : \mathbb{R}^{2N} \rightarrow \mathbb{R}$  in (16). We have  $C_{\text{twz}}(\boldsymbol{\theta}^*) = 0$ . Moreover, there exists  $\epsilon_0 > 0$  such that  $\mathcal{C}_{\text{twz}}^+ \cap \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) \neq \emptyset$ , where  $\mathcal{C}_{\text{twz}}^+ := \{\boldsymbol{\theta} \mid C_{\text{twz}}(\boldsymbol{\theta}) > 0\}$ , for all  $\epsilon \in (0, \epsilon_0]$ .

Lemma 4.3 is used in the following proposition whose proof is given in Supplementary Material.

*Proposition 4.4:* Consider the optical tweezer-based unfolding dynamics in (15). The folded conformation  $\boldsymbol{\theta}^*$  is an unstable equilibrium for (15) if

$$P(\boldsymbol{\theta}) := \frac{\partial C_{\text{twz}}}{\partial \boldsymbol{\theta}} \mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}) > 0 \text{ for all } \boldsymbol{\theta} \in \mathcal{C}_{\text{twz}}^+, \quad (17)$$

where  $\mathcal{C}_{\text{twz}}^+$  is defined in Lemma 4.3, and

$$\frac{\partial C_{\text{twz}}}{\partial \boldsymbol{\theta}} = 2\Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \mathbf{M}(\boldsymbol{\theta}^*) \frac{\partial \mathbf{r}_{\text{NC}}}{\partial \boldsymbol{\theta}}, \quad (18)$$

where  $\mathbf{M}(\boldsymbol{\theta}^*) := |\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)|^2 \cos^2(\alpha_C) \mathbf{I}_3 - \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*) \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)^\top$ .

The condition given by (17) can be numerically checked by investigating the Hessian of the function  $P(\cdot)$  at the folded conformation under study as follows.

*Proposition 4.5:* Consider the folded protein molecule conformation  $\boldsymbol{\theta}^*$  for the closed-loop dynamics in (15) and the set  $\mathcal{C}_{\text{twz}}^+$  in Lemma 4.3. Considering the Hessian matrix  $\text{Hess}(P)|_{\boldsymbol{\theta}^*}$ , the condition given by (17) holds if and only if  $\delta \boldsymbol{\theta}^\top \text{Hess}(P)|_{\boldsymbol{\theta}^*} \delta \boldsymbol{\theta} > 0$  for all  $\delta \boldsymbol{\theta}$  such that  $\boldsymbol{\theta}^* + \delta \boldsymbol{\theta} \in \mathcal{C}_{\text{twz}}^+ \cap \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon)$  for some  $\epsilon > 0$ .

*Remark 4.6:* The condition in Proposition 4.5, which relies on numerical computation of  $\text{Hess}(P)$  at the folded conformation  $\boldsymbol{\theta}^*$ , can be used for real-time control of optical tweezers and AFMs in protein unfolding applications. In particular, to approximate each entry  $\frac{\partial^2 P}{\partial \theta_i \partial \theta_j}$ , one can use

$$\frac{P(\boldsymbol{\theta}^* + \epsilon_i \mathbf{1}_i + \epsilon_j \mathbf{1}_j) - P(\boldsymbol{\theta}^* + \epsilon_i \mathbf{1}_i) - P(\boldsymbol{\theta}^* + \epsilon_j \mathbf{1}_j) + P(\boldsymbol{\theta}^*)}{\epsilon_i \epsilon_j}, \quad (19)$$

where  $P(\boldsymbol{\theta}^*) = 0$  because  $\mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}^*) = \mathbf{0}$  (see Supplementary Material for analytical expression of these entries). Moreover,  $\epsilon_i, \epsilon_j$  are some small positive constants and the vectors  $\mathbf{1}_i, \mathbf{1}_j$  are the  $i$ -th,  $j$ -th columns of the identity matrix  $\mathbf{I}_{2N}$ , respectively (see, e.g., [31]). For real-time implementation, which is beyond the scope of the current article, there is a need for high-speed scanning and high-bandwidth control

at nanoscale, where specialized DSP/FPGA-based embedded hardware have been developed (see, e.g., [24], [32]).

### C. Synthesis of destabilizing control inputs for unfolding

In this section we present a class of CCFs for the KCM-based molecular dynamics. We then utilize these CCFs in the Artstein-Sontag universal formula, which was first presented in [21] in the context of CCFs, to synthesize destabilizing state feedback control torques for folded protein conformations. The following propositions provide two types of CCFs for the KCM-based dynamics of protein molecules, where the CCF in the second proposition is the candidate CF given by (16).

*Proposition 4.7:* Consider the KCM-based dynamics in (7) and a folded conformation  $\boldsymbol{\theta}^*$ . Consider the function  $C : \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon_0) \rightarrow \mathbb{R}$  given by

$$C(\boldsymbol{\theta}) = g(|\Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)|^2), \quad (20)$$

where  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is any smooth strictly monotonic function such that  $g(0) = 0$ . The function  $C(\cdot)$  is a CCF for the nonlinear control-affine system in (7).

*Proof:* Consider the set  $\mathcal{C}^+ := \{\boldsymbol{\theta} \in \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon_0) : C(\boldsymbol{\theta}) > 0\}$ . From (20), we have  $C(\boldsymbol{\theta}^*) = 0$ . Also,  $\mathcal{C}^+ \cap \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) \neq \emptyset$  for all  $\epsilon \in (0, \epsilon_0]$ . Furthermore,  $\frac{\partial C}{\partial \boldsymbol{\theta}} = g'(|\Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)|^2) \Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \frac{\partial \mathbf{r}_{NC}}{\partial \boldsymbol{\theta}}$ , where  $g'(\cdot)$  is the derivative of  $g(\cdot)$ . We claim that

$$\partial \mathcal{C}^+ \cap \mathcal{C}^+ \cap \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) = \emptyset, \quad (21)$$

where  $\partial \mathcal{C}^+ := \{\boldsymbol{\theta} : |\frac{\partial C}{\partial \boldsymbol{\theta}}| = 0\}$ . To see this, note that

$$\boldsymbol{\theta} \in \partial \mathcal{C}^+ \text{ if and only if } \Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \frac{\partial \mathbf{r}_{NC}}{\partial \boldsymbol{\theta}} = \mathbf{0}. \quad (22)$$

On the set  $\mathcal{C}^+$ ,  $\Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \neq \mathbf{0}$ . Therefore, the relation in (22) holds if and only if  $\Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \frac{\partial \mathbf{r}_{NC}}{\partial \theta_j} = \mathbf{0}$  for all  $1 \leq j \leq 2N$ . Due to the kinematic structure of the protein molecule, which can be classified as a manipulator with hyper degrees of freedom (see [33]), we have

$$\frac{\partial \mathbf{r}_{NC}}{\partial \theta_j} = \Xi(\boldsymbol{\theta}, \mathbf{u}_j^0) \mathbf{u}_j^0 \times (\mathbf{r}_{NC}(\boldsymbol{\theta}) - \mathbf{r}_j(\boldsymbol{\theta})), \quad 1 \leq j \leq 2N, \quad (23)$$

where  $\Xi$  is given by (2) and  $\mathbf{r}_j(\boldsymbol{\theta})$  is the position vector of the backbone chain atom ( $C_\alpha$  atom for even  $j$  and  $N$  atom for odd  $j$ ) in the  $j$ -th peptide plane. Therefore, the relation in (22) holds if and only if the nonzero vector  $\Delta \mathbf{r}_{NC}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$  is perpendicular to the collection of  $2N$  three-dimensional vectors  $\frac{\partial \mathbf{r}_{NC}}{\partial \theta_j}$ . This geometric relation holds only if the protein molecule conformation  $\boldsymbol{\theta}$  is such that all the amino-acid linkages are co-located on the same two-dimensional plane corresponding to an unfolded structure, which is impossible in a neighborhood of a folded conformation  $\boldsymbol{\theta}^*$ . Due to (21), we conclude that  $\sup_{\mathbf{u}_c \in \mathbb{R}^{2N}} \{ \frac{\partial C}{\partial \boldsymbol{\theta}}(\mathcal{J}^\top(\boldsymbol{\theta}) \mathcal{F}(\boldsymbol{\theta}) + \mathbf{u}_c) \} > 0$  for all  $\boldsymbol{\theta} \in \mathcal{C}^+$ . Therefore,  $C(\cdot)$  is a CCF for (7). ■

*Proposition 4.8:* Consider the KCM-based dynamics in (7) and a folded conformation  $\boldsymbol{\theta}^*$ . The function  $C_{twz}(\cdot)$  in (16) is a CCF for (7).

Having a family of CCFs afforded by Propositions 4.7 and 4.8, we can use the following CCF-based Artstein-Sontag universal formula from [21]

$$\mathbf{u}_c(\boldsymbol{\theta}) = -\phi[a(\boldsymbol{\theta}), |B(\boldsymbol{\theta})|] B(\boldsymbol{\theta}), \quad (24)$$

to synthesize unfolding control inputs that solve the PUCP formulated in Section IV-A, where  $a(\boldsymbol{\theta}) := \frac{\partial C}{\partial \boldsymbol{\theta}} \mathcal{J}^\top(\boldsymbol{\theta}) \mathcal{F}(\boldsymbol{\theta})$ ,  $B^\top(\boldsymbol{\theta}) := \frac{\partial C}{\partial \boldsymbol{\theta}}$ , and  $\phi(a, b) := \frac{a - \sqrt{|a|^p + b^q}}{b^2}$  if  $b \neq 0$ , and  $\phi(a, b) := 0$  if  $b = 0$ , with any  $2q \geq p > 1$ , and  $q > 1$ .

## V. SIMULATION RESULTS

In this section we present simulation results to validate our proposed approach to protein unfolding using Chetaev instability framework. We consider a protein molecule chain consisting of  $N - 1 = 10$  peptide planes with a 22-dimensional dihedral angle space in (7). Our implementation follows the PROTOFOLD I guidelines [10], [12].

We have performed the unfolding simulations on the folded protein conformation  $\boldsymbol{\theta}^{*1}$  that has been obtained using the KCM-based iteration in [10]. The candidate Chetaev function  $C_{twz} : \mathbb{R}^{22} \rightarrow \mathbb{R}_+$  is chosen according to (16) with  $\alpha_C = \frac{1}{4}$ . To visualize this function, we have plotted it using the dihedral angle vectors of the form  $\boldsymbol{\theta} = \boldsymbol{\theta}^* + [\delta\theta_1, \delta\theta_2, \delta\theta_3, 0, \dots, 0]^\top$ , where the triplet  $(\delta\theta_1, \delta\theta_2, \delta\theta_3)$  belong to the three-dimensional sphere of radius 0.1 radians and centered at the origin. Figure 2 depicts both the folded conformation at  $\boldsymbol{\theta}^*$  and the visualization of the candidate Chetaev function when translated to the origin.

Using the setup above, we conduct two numerical simulations. The first one entails the optical tweezer-based unfolding while the second one entails an unfolding simulation using the Artstein-Sontag universal formula in (24) with  $p = q = 2$  and  $C_{twz}$  in Proposition 4.8 as the CCF. In the first simulation, where the numerical condition of Proposition 4.5 is verified, we let the optical tweezer displacement be  $\mathbf{x}_{twz} = x_0 \frac{\mathbf{r}_{NC}(\boldsymbol{\theta}^*)}{\|\mathbf{r}_{NC}(\boldsymbol{\theta}^*)\|}$ , where  $x_0$  is chosen to be 51 nm. In other words, we elongated the protein molecule along the vector that connects the  $N$ -terminus to the  $C$ -terminus. Furthermore, we utilized a modulated optical trap stiffness of  $\kappa(\boldsymbol{\theta}) = \kappa_0 |\mathbf{r}_{NC}(\boldsymbol{\theta}^*) - \mathbf{r}_{NC}(\boldsymbol{\theta})|^2$ , where  $\kappa_0 = 0.16$  pN/nm. Figure 3 depicts the free energy of the molecule under the optical tweezer-based and the CCF-based unfolding control inputs as well as several protein conformations along their respective unfolding pathways (see Supplementary Material for further discussion).

## VI. CONCLUSION

This paper formulated the important biochemical process of KCM-based protein unfolding as a destabilizing control analysis/synthesis problem. In light of this formulation and using the Chetaev instability framework, numerical conditions on optical tweezer forces that lead to protein unfolding

<sup>1</sup> $\boldsymbol{\theta}^* = [1.34 \ 1.37 \ 1.26 \ 1.29 \ 1.42 \ 1.73 \ 1.65 \ 1.65 \ 1.46 \ 1.62 \ 1.49 \ 2.01 \ 1.31 \ 0.99 \ 1.98 \ 1.9 \ 1.59 \ 1.56 \ 1.57 \ 0.93 \ 1.22 \ 1.29]^\top$  in radians.

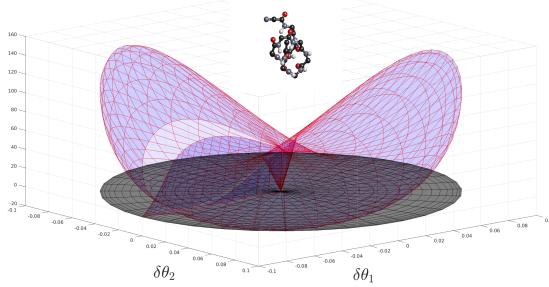


Fig. 2: Visualization of the candidate Chetaev function  $C_{twz} : \mathbb{R}^{22} \rightarrow \mathbb{R}$  based on the depicted folded conformation at  $\theta^*$ .

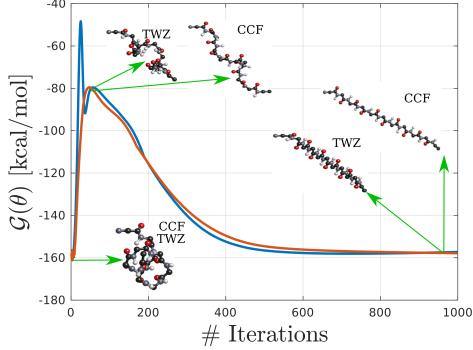


Fig. 3: Energy profiles from unfolding numerical simulations under optical tweezer-based control inputs (red curve; molecule labeled by TWZ) and a theoretically-motivated input generated by the Artstein-Sontag universal formula (blue curve; molecule labeled by CCF).

have been derived. Furthermore, after presenting two types of CCFs, a theoretically-motivated CCF-based Artstein-Sontag unfolding input has been derived. These findings lead us to further research avenues such as encoding entropy-loss constraints for protein unfolding using the obtained CCFs, investigating folding pathways for two-state proteins, and studying protein folding/unfolding in solutions.

## REFERENCES

- [1] A. V. Finkelstein and O. Ptitsyn, *Protein physics: a course of lectures*. Elsevier, 2016.
- [2] F. Bergasa-Caceres and H. A. Rabitz, “Interdiction of protein folding for therapeutic drug development in SARS CoV-2,” *J. Phys. Chem. B*, vol. 124, no. 38, pp. 8201–8208, 2020.
- [3] J. Fang, A. Mehlrich, N. Koga, J. Huang, R. Koga, X. Gao, C. Hu, C. Jin, M. Rief, J. Kast, D. Baker, and H. Li, “Forced protein unfolding leads to highly elastic and tough protein hydrogels,” *Nature Commun.*, vol. 4, no. 1, pp. 1–10, 2013.
- [4] L. Fu, H. Wang, and H. Li, “Harvesting mechanical work from folding-based protein engines: from single-molecule mechanochemical cycles to macroscopic devices,” *CCS Chem.*, vol. 1, no. 1, pp. 138–147, 2019.
- [5] M. AlQuraishi, “A watershed moment for protein structure prediction,” *Nature*, vol. 577, pp. 627–628, 2020.
- [6] A. R. Fersht, “AlphaFold—A personal perspective on the impact of machine learning,” *J. Mol. Biol.*, vol. 433, no. 20, p. 167088, 2021.
- [7] L. Heo, C. F. Arbour, and M. Feig, “Driven to near-experimental accuracy by refinement via molecular dynamics simulations,” *Proteins: Struct. Funct. Bioinform.*, vol. 87, no. 12, pp. 1263–1275, 2019.
- [8] C. Alvarado and K. Kazerounian, “On the rotational operators in protein structure simulations,” *Prot. Eng.*, vol. 16, no. 10, pp. 717–720, 2003.
- [9] K. Kazerounian, K. Latif, K. Rodriguez, and C. Alvarado, “Nanokinematics for analysis of protein molecules,” *ASME J. Mech. Des.*, vol. 127, no. 4, pp. 699–711, 2005.
- [10] K. Kazerounian, K. Latif, and C. Alvarado, “PROTOFOLD: A successive kinetostatic compliance method for protein conformation prediction,” *ASME J. Mech. Des.*, vol. 127, no. 4, pp. 712–717, 2005.
- [11] P. Tavousi, M. Behandish, H. T. Ilieş, and K. Kazerounian, “PROTOFOLD II: Enhanced model and implementation for kinetostatic protein folding,” *ASME J. Nanotechnol. Eng. Med.*, vol. 6, no. 3, 2015.
- [12] P. Tavousi, “On the systematic design and analysis of artificial molecular machines,” Ph.D. dissertation, University of Connecticut, 2016.
- [13] Z. Shahbazi, H. T. Ilieş, and K. Kazerounian, “Hydrogen bonds and kinematic mobility of protein molecules,” *ASME J. Mech. Robot.*, vol. 2, no. 2, 2010.
- [14] P. Tavousi, K. Kazerounian, and H. Ilieş, “Synthesizing functional mechanisms from a link soup,” *ASME J. Mech. Des.*, vol. 138, no. 6, 2016.
- [15] M. T. Chorsi, P. Tavousi, C. Mundrane, V. Gorbatyuk, K. Kazerounian, and H. Ilieş, “Kinematic design of functional nanoscale mechanisms from molecular primitives,” *ASME J. Micro-Nano-Manuf.*, vol. 9, no. 2, p. 021005, 2021.
- [16] A. Mohammadi and M. W. Spong, “Quadratic optimization-based nonlinear control for protein conformation prediction,” *IEEE Control Syst. Lett.*, vol. 6, pp. 373–378, 2022, doi: 10.1109/LCSYS.2021.3076869.
- [17] H.-J. Su, J. Parker, K. Kazerounian, and H. Ilieş, “A comparison of kinetostatic and multibody dynamics models for simulating protein structures,” in *ASME Int. Des. Eng. Tech. Comput. Inf. Eng. Conf. (IDETC-CIE)*, vol. 48043, 2007, pp. 897–904.
- [18] N. G. Chetaev, *The stability of motion*. New York: Pergamon Press, 1961.
- [19] D. V. Efimov and A. L. Fradkov, “Oscillatory of nonlinear systems with static feedback,” *SIAM J. Contr. Optim.*, vol. 48, no. 2, pp. 618–640, 2009.
- [20] D. Efimov and W. Perruquetti, “Oscillating system design applying universal formula for control,” in *2011 50<sup>th</sup> IEEE Conf. Dec. Contr. Eur. Contr. Conf. (CDC-ECC)*, 2011, pp. 1747–1752.
- [21] D. Efimov, W. Perruquetti, and M. Petreczky, “On necessary conditions of instability and design of destabilizing controls,” in *53<sup>rd</sup> IEEE Conf. Dec. Contr. (CDC)*. IEEE, 2014, pp. 3915–3917.
- [22] P. Braun, L. Grüne, and C. M. Kellett, “Complete instability of differential inclusions using Lyapunov methods,” in *57<sup>th</sup> IEEE Conf. Dec. Contr. (CDC)*. IEEE, 2018, pp. 718–724.
- [23] C. Bustamante, L. Alexander, K. Maciuba, and C. M. Kaiser, “Single-molecule studies of protein folding with optical tweezers,” *Annu. Rev. Biochem.*, vol. 89, pp. 443–470, 2020.
- [24] M. R. Ragazzon, J. T. Gravdahl, and K. Y. Pettersen, “Model-based identification of nanomechanical properties in atomic force microscopy: Theory and experiments,” *IEEE Trans. Contr. Syst. Technol.*, vol. 27, no. 5, pp. 2045–2057, 2018.
- [25] N. Thammawongsa, F. D. Zainol, S. Mitatha, J. Ali, and P. P. Yupapin, “Nanorobot controlled by optical tweezer spin for microsurgical use,” *IEEE Trans. Nanotechnol.*, vol. 12, no. 1, pp. 29–34, 2012.
- [26] Y. Wen, H. Lu, Y. Shen, and H. Xie, “Nanorobotic manipulation system for 360° characterization atomic force microscopy,” *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 2916–2924, 2019.
- [27] D. B. Adolf and M. D. Ediger, “Brownian dynamics simulations of local motions in polyisoprene,” *Macromolecules*, vol. 24, no. 21, pp. 5834–5842, 1991.
- [28] E. Schonbrun and K. Crozier, “Spring constant modulation in a zone plate tweezer using linear polarization,” *Opt. Lett.*, vol. 33, no. 17, pp. 2017–2019, 2008.
- [29] X. Li and C. C. Cheah, “Tracking control for optical manipulation with adaptation of trapping stiffness,” *IEEE Trans. Contr. Syst. Technol.*, vol. 24, no. 4, pp. 1432–1440, 2015.
- [30] Y. Lin, E. D. Sontag, and Y. Wang, “A smooth converse Lyapunov theorem for robust stability,” *SIAM J. Contr. Optim.*, vol. 34, no. 1, pp. 124–160, 1996.
- [31] J. E. Dennis Jr and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM, 1996.
- [32] Y. Sun, Y. Fang, Y. Zhang, and X. Dong, “Field programmable gate array (FPGA) based embedded system design for AFM real-time control,” in *2010 IEEE Int. Conf. Contr. Appl. (CCA)*. IEEE, 2010, pp. 245–250.
- [33] H. Mochiyama and H. Kobayashi, “The shape Jacobian of a manipulator with hyper degrees of freedom,” in *Proc. 1999 IEEE Int. Conf. Robot. Automat. (ICRA)*, vol. 4. IEEE, 1999, pp. 2837–2842.

SUPPLEMENTARY MATERIAL TO “CHETAEV INSTABILITY FRAMEWORK FOR KINETOSTATIC COMPLIANCE-BASED PROTEIN UNFOLDING”

**Notation.** In these supplementary notes, we will number the equations using (s#), the references using [S#], and the figures using SF#. Therefore, (s1) refers to Equation 1 in the supplementary notes while (1) refers to Equation 1 in the original article. The same holds true for the references and the figures.

*A. Schematic of protein folding/unfolding against the free energy landscape*

Figure SF1 depicts a schematic of folding/unfolding against the free energy landscape of the protein molecule.

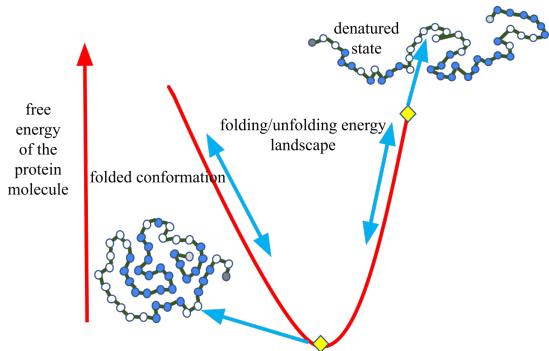


Fig. SF1: Protein folding/unfolding against the free energy landscape of the molecule.

*B. Principles of operation of optical tweezers*

Optical tweezers, which were discovered and developed by Ashkin [S2], operate based on the principle of force exertion by light on matter due to the change in momentum that light carries. Figure SF2 depicts a schematic of optical tweezer-based protein unfolding. Optical tweezers can be utilized for unfolding protein molecules by chemically attaching a latex bead (microsphere) to the protein under study. The bead could subsequently be trapped and utilized both as a handle for manipulating the protein molecule and as a tool for measuring displacements at a very small scale. Under the assumptions of a nearly Gaussian beam and a locally harmonic potential energy for a given optical tweezer, the optical trap can be modeled as a Hookean spring [23].

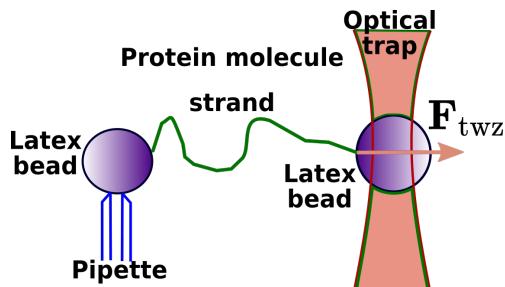


Fig. SF2: Schematic of optical tweezer-based protein unfolding (recreated from [S1]).

*C. On the choice of the candidate CF*

Figure SF3 represents the geometric interpretation associated with the three dimensional vectors utilized for defining  $C_{\text{twz}}(\cdot)$  in (16). The vectors  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta})$  and  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)$  represent the “length” of the molecule at a general conformation  $\boldsymbol{\theta}$  and the folded conformation  $\boldsymbol{\theta}^*$ , respectively. The positive level sets of the candidate CF  $C_{\text{twz}}(\cdot)$  in (16), which depends on  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta})$  and  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)$ , correspond to conformations at which the protein molecule has been elongated from the folded conformation along desired directions. Indeed, when  $\dot{C}_{\text{twz}} > 0$  along the trajectories of the unfolding dynamics on the positive level sets of  $C_{\text{twz}}(\cdot)$ , the molecule is being elongated away from its folded conformation.

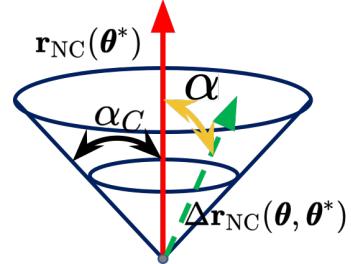


Fig. SF3: When  $C_{\text{twz}}(\boldsymbol{\theta}) = 0$ ,  $\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$  belongs to a right circular cone with vertex at the origin, axis parallel to  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)$ , and aperture  $2\alpha_C$ .

*D. Proof of Technical Results*

**Proof of Lemma 4.3.** Since  $\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*) = \mathbf{0}$ , we conclude that  $C_{\text{twz}}(\boldsymbol{\theta}^*) = 0$ . Let us define the mapping

$$f : \mathbb{R}^{2N} \setminus \{\boldsymbol{\theta}^*\} \rightarrow \mathbb{R}, \boldsymbol{\theta} \mapsto \cos^{-1} \left( \frac{\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)}{\|\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)\| \|\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)\|} \right), \quad (\text{s1})$$

where  $f(\boldsymbol{\theta})$  is the angle between  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)$  and  $\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ . We claim

$$C_{\text{twz}}^+ = f^{-1} \left( (\alpha_C, \frac{\pi}{2}) \cap (-\frac{\pi}{2}, -\alpha_C) \right). \quad (\text{s2})$$

Consider an arbitrary  $\boldsymbol{\theta}_0 \in \mathbb{R}^{2N}$  such that  $C_{\text{twz}}(\boldsymbol{\theta}_0) > 0$ . From (16) and (s1), it can be seen that

$$C_{\text{twz}}(\boldsymbol{\theta}_0) = |\Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}_0, \boldsymbol{\theta}^*)|^2 |\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)|^2 (\cos^2(\alpha_C) - \cos^2(f(\boldsymbol{\theta}_0))). \quad (\text{s3})$$

Therefore,  $C_{\text{twz}}(\boldsymbol{\theta}_0) > 0$  if and only if  $\boldsymbol{\theta}_0 \in C_{\text{twz}}^+$ . Since  $f(\cdot)$  is a continuous function, we conclude that  $C_{\text{twz}}^+$  in (s2) is an open set and  $\boldsymbol{\theta}^*$  belongs to the boundary of  $C_{\text{twz}}^+$ . Therefore,  $C_{\text{twz}}^+ \cap \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) \neq \emptyset$  for all  $\epsilon \in (0, \epsilon_0]$  for some  $\epsilon_0 > 0$ . ■

**Proof of Proposition 4.4.** Considering the smooth candidate Chetaev function in (16), the directional derivative  $DC_{\text{twz}}(\boldsymbol{\theta}) \mathbf{f}^{\text{unfold}}(\boldsymbol{\theta})$  in the direction of  $\mathbf{f}^{\text{unfold}}(\cdot)$  is given by  $\frac{\partial C_{\text{twz}}}{\partial \boldsymbol{\theta}} \mathbf{f}^{\text{unfold}}(\boldsymbol{\theta})$ , where

$$\frac{\partial C_{\text{twz}}}{\partial \boldsymbol{\theta}} = 2 \left\{ |\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)|^2 \cos^2(\alpha_C) \Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top - \Delta\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*) \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)^\top \right\} \frac{\partial \mathbf{r}_{\text{NC}}}{\partial \boldsymbol{\theta}}. \quad (\text{s4})$$

Therefore, due to Lemma 4.3, the statement of the proposition follows from Theorem 2.1.  $\blacksquare$

**Proof of Proposition 4.5.** Given the  $C^\infty$ -smoothness of  $P(\cdot)$  in a neighborhood of  $\boldsymbol{\theta}^*$ , we can utilize the Taylor's theorem for multivariate functions to arrive at

$$P(\boldsymbol{\theta}) = P(\boldsymbol{\theta}^*) + \frac{\partial P}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}^*} \delta \boldsymbol{\theta} + \delta \boldsymbol{\theta}^\top \text{Hess}(P)|_{\boldsymbol{\theta}^*} \delta \boldsymbol{\theta} + \text{H.O.T}, \quad (\text{s5})$$

where  $\boldsymbol{\theta} \in \mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) \cap \mathcal{V}^+$  for some  $\epsilon > 0$  and  $\delta \boldsymbol{\theta} := \boldsymbol{\theta} - \boldsymbol{\theta}^*$ . It can be seen that  $P(\boldsymbol{\theta}^*) = 0$  because  $\boldsymbol{\theta}^*$  is an equilibrium for the dynamics in (15). Furthermore,  $\frac{\partial P}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}^*} = \mathbf{0}$  because  $\mathbf{f}^{\text{unfold}}(\boldsymbol{\theta}^*) = \mathbf{0}$  and  $\frac{\partial C_{\text{twz}}}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}^*} = \mathbf{0}$  according to (18). Therefore,  $P(\boldsymbol{\theta}) > 0$  on  $\mathcal{B}_{\boldsymbol{\theta}^*}(\epsilon) \cap \mathcal{V}^+$  if and only if the stated condition in the proposition holds.  $\blacksquare$

**Proof of Proposition 4.8.** The proof is almost verbatim the proof of Proposition 4.7 with the difference that the relation in (s6) should be replaced with

$$\boldsymbol{\theta} \in \partial \mathcal{C}^+ \text{ if and only if } \Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)^\top \mathbf{M}(\boldsymbol{\theta}^*) \frac{\partial \mathbf{r}_{\text{NC}}}{\partial \boldsymbol{\theta}} = \mathbf{0}. \quad (\text{s6})$$

Since  $\mathbf{M}(\boldsymbol{\theta}^*)$  is a symmetric and non-singular matrix with  $\det(\mathbf{M}) = -\sin^2(\alpha_c) |\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)|^6$ , the relation in (s6) holds if and only if the nonzero vector  $\mathbf{M}(\boldsymbol{\theta}^*) \Delta \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$  is perpendicular to the collection of  $2N$  three-dimensional vectors  $\frac{\partial \mathbf{r}_{\text{NC}}}{\partial \theta_j}$ . The rest of the proof follows verbatim the proof of Proposition 4.7.  $\blacksquare$

**The Analytical Expression for the Entries of Hessian.** The entries of the Hessian of the function  $P(\cdot)$  given by (17) at a folded conformation  $\boldsymbol{\theta}^*$  take the form

$$\begin{aligned} \frac{\partial^2 P}{\partial \theta_j \partial \theta_k} \Big|_{\boldsymbol{\theta}^*} = & 2 \sum_{i=1}^{2N} \frac{\partial \mathbf{r}_{\text{NC}}}{\partial \theta_i}(\boldsymbol{\theta}^*) \mathbf{M}(\boldsymbol{\theta}^*) \left\{ \frac{\partial \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)}{\partial \theta_j} \frac{\partial f_{i, \text{unfold}}(\boldsymbol{\theta}^*)}{\partial \theta_k} + \right. \\ & \left. \frac{\partial \mathbf{r}_{\text{NC}}(\boldsymbol{\theta}^*)}{\partial \theta_k} \frac{\partial f_{i, \text{unfold}}(\boldsymbol{\theta}^*)}{\partial \theta_j} \right\}, \quad 1 \leq j, k \leq 2N, \end{aligned} \quad (\text{s7})$$

where  $\mathbf{M}(\boldsymbol{\theta}^*)$  is defined in Proposition 4.4, the vector  $\mathbf{r}_{\text{NC}}(\boldsymbol{\theta}) \in \mathbb{R}^3$  connects the nitrogen atom in the N-terminus to the carbon atom in the C-terminus, and  $f_{i, \text{unfold}}(\boldsymbol{\theta})$  is the  $i$ -th element of the vector  $\mathbf{f}^{\text{unfold}}(\boldsymbol{\theta})$  in (15).

#### E. Control Chetaev Functions

In this section we provide some preliminaries on control Chetaev functions due to Efimov and collaborators [S3]. The reader is referred to the line of work by Efimov, Kellett, and collaborators [S3] and [S4] for further details.

Consider the nonlinear control-affine system

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}, \quad (\text{s8})$$

where  $\mathbf{x} \in \mathbb{R}^n$  is the state vector,  $\mathbf{u} \in \mathbb{R}^m$  is the control input,  $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$ , and  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\mathbf{G} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  are locally Lipschitz continuous functions. A smooth function  $V : \mathcal{B}_{\mathbf{x}^*}(\epsilon) \rightarrow \mathbb{R}$ , such that  $V(\mathbf{x}^*) = 0$  and  $\mathcal{V}^+ \cap \mathcal{B}_{\mathbf{x}^*}(\epsilon) \neq \emptyset$ , where  $\mathcal{V}^+ := \{\mathbf{x} : V(\mathbf{x}) > 0\}$  and for any  $\epsilon \in (0, \epsilon_0]$ , is called a CCF for the control system in (s8) if

$$\sup_{\mathbf{u} \in \mathbb{R}^m} \{D^- V(\mathbf{x}) \mathbf{F}(\mathbf{x}) + (D^- V(\mathbf{x}) \mathbf{G}(\mathbf{x}))^\top \mathbf{u}\} > 0, \quad (\text{s9})$$

for all  $\mathbf{x} \in \mathcal{V}^+$ . As it is argued in [S3], if for all  $\mathbf{x} \in \mathcal{V}^+$  at which  $|D^- V(\mathbf{x}) \mathbf{G}(\mathbf{x})| = 0$ , it holds that  $|D^- V(\mathbf{x}) \mathbf{F}(\mathbf{x})| > 0$ , then  $V(\cdot)$  is a CCF for (s8). Furthermore, in the special case when  $\mathbf{G}(\mathbf{x}) = \mathbf{I}_{n \times n}$  and  $\mathbf{u} \in \mathbb{R}^n$ , if  $|D^- V(\mathbf{x})| \neq 0$  for all  $\mathbf{x} \in \mathcal{V}^+$ , then  $V(\cdot)$  is a CCF for (s8).

#### SUPPLEMENTARY REFERENCES

- [S1] J. L. Killian, F. Ye, and M. D. Wang, "Optical tweezers: a force to be reckoned with," *Cell*, vol. 175, no. 6, pp. 1445–1448, 2018.
- [S2] A. Ashkin, "Acceleration and trapping of particles by radiation pressure," *Phys. Rev. Lett.*, vol. 24, no. 4, p. 156, 1970.
- [S3] D. Efimov, W. Perruquetti, and M. Petreczky, "On necessary conditions of instability and design of destabilizing controls," in *53rd IEEE Conf. Dec. Contr. (CDC)*. IEEE, 2014, pp. 3915–3917.
- [S4] P. Braun, L. Grüne, and C. M. Kellett, "Complete instability of differential inclusions using Lyapunov methods," in *57th IEEE Conf. Dec. Contr. (CDC)*. IEEE, 2018, pp. 718–724.