Using the Helmholtz decomposition of the vector field of folding fluxes in a reduced space of collective variables, a potential of the driving force for protein folding is determined. The potential has two components and can be written as a complex function. One component is responsible for the source and sink of the folding flows (representing, respectively, the unfolded states and the native state of the protein), and the other accounts for the vorticity of the flow that is produced at the boundaries of the main flow by the contact of the moving folding "fluid" with the quiescent surroundings. The theoretical consideration is illustrated by calculations for a model β-hairpin protein.

Given the flows, it is possible to construct vector fields and streamlines of the folding flows, similar to how it is done in hydrodynamics [28]. Then, the process of protein folding can be viewed as a steady flow of a folding "fluid" from the unfolded states of the protein to its native state. The hydrodynamic approach has been successfully applied to the study of folding dynamics of several model proteins - an α-helix (off-lattice model) [21], a SH3 domain (Cα model) [23, 25], and beta3s-miniprotein (all-atom simulations) [26]. It has been found that although the overall direction of the folding flow is consistent with the FES landscape, the distribution of the local flows is very different from what the FES suggests. In other words, the FES does not present a true potential for folding flows.

The knowledge of the local flows offers a possibility to determine such a potential on the basis of the Helmholtz decomposition [27] of the vector field of the flows. In the present Letter, using a model β-hairpin protein as an example, we show that this potential has two components and can be written as a complex function. One component is responsible for the source and sink of the folding flows, which correspond, respectively, to the initiation of the folding trajectories in the unfolded states and their termination in the native state, and the other accounts for the vorticity of the flow that is produced at the boundaries of the main flow by the contact of the moving folding "fluid" with the quiescent surroundings.

Determining the potential with the Helmholtz decomposition requires a smooth vector field of folding flows and, accordingly, a large number of folding trajectories to be simulated. To speed up the simulations, we considered a short fast-folding protein in the framework of a minimalist model. The goal was to have a typical picture of the folding dynamics rather than to describe folding of a particular protein. Specifically, a twelve-residue β-hairpin protein 2evq [29] (Supplementary material, Fig. S1) was chosen as a model system, and the interactions between the residues were characterized using a Cα-based Gö-like model, i.e. the residues were represented by monomers (beads) centered at the Cα atoms, with the interaction between the monomers determined by the structure of the native state of the protein [29].
cally, the interaction potential of Ref. [30] was used. It consists of three terms; one term accounts for rigidity of the protein backbone, and the other two describe the contribution of the native and non-native contacts, assuming that the monomers interact through the Lennard-Jones 12-6 potential (with different parametrization). The contact between two monomers was considered to be a native contact if the monomers were not immediate neighbors along the chain, and the distance between these monomers in the native state was less than 7.5 Å. In the native state, the number of such contacts is equal to 27. The simulations were performed using a constant-temperature molecular dynamics based on the Langevin equation; the equation was solved with the numerical algorithm of Ref. [31]. (For details about the interaction potential and Langevin dynamics see Supplementary Material). In what follows, we use dimensionless units with the length, energy and time scales, respectively, to be \( l = 7.5 \text{ Å}, \varepsilon \) and \( t = (m l^2/\varepsilon)^{1/2} \), where \( \varepsilon \) is the Lennard-Jones energy constant, and \( m \) is the mass of a monomer.

There were simulated \( 1 \times 10^5 \) folding trajectories at a temperature equal to \( T = 0.17 \) (in units \( \varepsilon/\k_B \), where \( \k_B \) is the Boltzmann constant). Each trajectory started at an unfolded state of the protein and terminated upon reaching the native state. The unfolded states were prepared by thermalization of the native conformation at \( T = 0.5 \) for \( 1 \times 10^4 \) time steps; if the number of native contacts did not exceed 4, the conformation was accepted, if not, the thermalization was continued. The native state was considered to be reached when the root-mean square deviation (RMSD) of the protein conformation from the native conformation was less than 1 Å. The simulations have shown that the present protein model gives a quite typical picture of protein folding: the FES depending on the fraction of native contacts and radius of gyration is "L-shaped" [10], the dependence of the mean first-passage time (MFPT) on the temperature is "U-shaped" [20,21,22,23], and the folding kinetics at temperatures close to the temperature at which the MFPT has the minimum value are essentially single-exponential [20], similar to those at \( T = 0.17 \) (see Figs. S2, S3 and S4, respectively). The value of the MFPT at \( T = 0.17 \) is \( t_f \approx 70 \).

To characterize protein conformations, we employed a set of bond distances between the monomers which are not immediate neighbors along the protein chain; they formed a 55-dimensional conformation space. Using the principal component analysis (PCA) [24], this space was reduced to a two-dimensional space of collective variables \( g = (g_1, g_2) \). The variable \( g_1 \) was directed along the eigenvector for the largest eigenvalue, and the variable \( g_2 \) along the vector calculated as the linear combination of the rest of the eigenvectors, which contributed with the weights corresponding to their eigenvalues [35].

Figure 1 shows the FES as a function of \( g_1 \) and \( g_2 \). The free energy was calculated as \( F(g_1, g_2) = -k_B T \ln P(g_1, g_2) \), where \( P(g_1, g_2) \) is the probability for the system to be found at the point \((g_1, g_2)\). Similar to the other FESs depending on the PCA components [36,37], the present FES is characteristically different from the FES depending on the fraction of native contacts and radius of gyration (Fig. S2) in that it is not "L-shaped".

The hydrodynamic description of protein folding [21] is based on the calculation of the transitions in a space of collective variables (the space \( g \) in the present case). These transitions are organized into transition probability local flows (fluxes) \( j(g) \). In the case of two variables, \( g = (g_1, g_2) \), the \( g_1 \)-component of the flux at a point \( g \) is determined as

\[
\begin{align*}
  j_{g_1}(g) = & \sum_{g'' \subset g'} n(g'', g') \\
  & - \sum_{g'' \supset g'} n(g'', g')/(M t_f)
\end{align*}
\]

where \( M \) is the total number of simulated trajectories, \( t_f \) is the MFPT, \( n(g'', g') \) is the number of transitions from state \( g' \) to \( g'' \), and \( g \subset g' \) is a symbolic designation of the condition that the transitions included in the sum have the straight line connecting points \( g' \) to \( g'' \), which crosses the line \( g_1 = \text{const} \) within the segment of the length of \( \Delta g_1 \) centered at the point \( g \). The first term on the right-hand side of the equation corresponds to

![FIG. 1:](image1.png) (Color online) Free energy surface as a function of collective variables \( g_1 \) and \( g_2 \).

![FIG. 2:](image2.png) (Color online) The distributions of the (a) divergence and (b) vorticity of the fluxing folds.
the transitions in the positive direction of \( g_1 \), and the second term to those in the negative direction. The \( g_2 \)-component of \( \mathbf{j}(\mathbf{g}) \) is determined in a similar way, except that one selects the transitions crossing the line \( g_2 = \text{const} \). The calculations were performed on a grid with discretization \( \Delta g_1 = \Delta g_2 = 0.12 \).

Figure 2 shows how the divergence of the folding flow \( (q = \partial j_{g_1}/\partial g_1 + \partial j_{g_2}/\partial g_2, \text{ panel a}) \) and its vorticity \( (\omega = \partial j_{g_2}/\partial g_1 - \partial j_{g_1}/\partial g_2, \text{ panel b}) \) are distributed in the \((g_1, g_2)\) space. Except for two localized regions at the unfolded and native states, which, represent, respectively, the source and sink of the folding flow, the flow is divergence-free. The vorticity, in contrast, spans the entire flow region between the unfolded and native states (cf. Fig. 1). It is produced by the contact of the moving folding fluid with the quiescent surroundings and, correspondingly, has different signs in the upper and lower parts of the region.

Figure 3 depicts the distribution of \( \mathbf{j}(\mathbf{g}) \) in the form of vector field (for illustrative purpose, the lengths of the vectors are equally increased by factor \( 3.5 \times 10^2 \)). The boundaries of the flow region, i.e. the region where the folding fluxes are essentially non-zero, approximately correspond to the FES boundaries (Fig. 1). Integration of the \( g_1 \)-component of \( \mathbf{j}(\mathbf{g}) \) over \( g_2 \) in each cross-section \( g_1 = \text{const} \) gives the total flow from the unfolded states to the native state as a function of \( g_1 \), i.e. \( G(g_1) = \int j_{g_1}(g_1, g_2)dg_2 \): it indicates that the flow is essentially constant in the region between the source and sink of the flow (Fig. S6).

Comparison of Fig. 3 to Fig. 1 shows that the folding flux distribution is consistent with the FES landscape in that the folding flow follows the valley of the FES that connects the unfolded states to the native state. At the same time, \( F(\mathbf{g}) \) does not present a true potential of folding flows, i.e. determining the fluxes as \( j_{g_1} = -\partial F(\mathbf{g})/\partial g_1 \) and \( j_{g_2} = -\partial F(\mathbf{g})/\partial g_2 \) leads to a vector flow field drastically different from that of Fig. 3 (Fig. S7). In fact, this can be seen immediately from the FES that the flow vectors will be generally directed perpendicular to the valley because of the very steep sides of the valley. Calculation of the corresponding total flow shows that the flow is not constant but fluctuates around zero value (Fig. S6).

To determine an actual potential of the driving force, we use the Helmholtz decomposition theorem [27], according to which any smooth vector field can be uniquely represented as a sum of two terms: a curl-free field and a divergence-free field. Then

\[
\mathbf{j} = \mathbf{j}_{\text{cf}} + \mathbf{j}_{\text{df}} \quad (2)
\]

where \( \mathbf{j}_{\text{cf}} \) is the curl-free component, and \( \mathbf{j}_{\text{df}} \) is the divergence-free component, i.e. \( \nabla \times \mathbf{j}_{\text{df}} = 0 \) and \( \nabla \cdot \mathbf{j}_{\text{df}} = 0 \), respectively. These conditions allow introducing the potentials of the vector fields. In the case of two dimensions, vectors \( \mathbf{j}_{\text{cf}} \) and \( \mathbf{j}_{\text{df}} \) can be written as

\[
\mathbf{j}_{\text{cf}} = -\frac{\partial \Phi}{\partial g_1} \mathbf{k}_1 - \frac{\partial \Phi}{\partial g_2} \mathbf{k}_2
\]

\[
\mathbf{j}_{\text{df}} = \frac{\partial \Psi}{\partial g_2} \mathbf{k}_1 - \frac{\partial \Psi}{\partial g_1} \mathbf{k}_2
\]

where \( \Phi = \Phi(\mathbf{g}) \) is the potential for the curl-free component, \( \Psi = \Psi(\mathbf{g}) \) is the potential for the divergence-free component, and \( \mathbf{k}_1 \) and \( \mathbf{k}_2 \) are the unit vectors for the variables \( g_1 \) and \( g_2 \), respectively. The curl-free component is expected to contain sources and sinks of the flow, and the divergence-free component to contain vortices (see below). Substituting these expressions for \( \mathbf{j}_{\text{cf}} \) and \( \mathbf{j}_{\text{df}} \) into Eq. (2) and regrouping the terms, one obtains

\[
\mathbf{j} = j_{g_1} \mathbf{k}_1 + j_{g_2} \mathbf{k}_2
\]

where

\[
j_{g_1} = -\frac{\partial \Phi}{\partial g_1} + \frac{\partial \Psi}{\partial g_1}, \quad j_{g_2} = -\frac{\partial \Phi}{\partial g_2} - \frac{\partial \Psi}{\partial g_2} \quad (3)
\]

We note that the functions \( \Phi(\mathbf{g}) \) and \( \Psi(\mathbf{g}) \) present potentials for the drift flows, because the velocity of motion is assumed to be proportional to the forces produced by these potentials. In other words, although the inertia of the monomers is present in the Langevin equation we used for the simulations, the resulting protein dynamics is expected to be an overdamped diffusion, which is characteristic of protein folding [28].

To find the functions \( \Phi(\mathbf{g}) \) and \( \Psi(\mathbf{g}) \), we minimized the functional

\[
Q = \int \left[ (j_{g_1} + \frac{\partial \Phi}{\partial g_1} - \frac{\partial \Psi}{\partial g_2})^2 + (j_{g_2} + \frac{\partial \Phi}{\partial g_2} + \frac{\partial \Psi}{\partial g_1})^2 \right] dg_1 dg_2 \quad (4)
\]

with respect to \( \Phi(\mathbf{g}) \) and \( \Psi(\mathbf{g}) \); here \( j_{g_1} \) and \( j_{g_2} \) are the folding fluxes obtained with the MD simulations. Integration in Eq. (4) was performed numerically on a grid with the same discretization as for Eq. (1). To avoid
boundary effects, the region of integration was extended to \(-21.8 \leq g_1 \leq 7.0\) and \(-13.7 \leq g_2 \leq 12.7\) (cf. Figs. 1 and 2). At the boundaries of the region, no-flux conditions were imposed, i.e. at the left and right boundaries \((g_1 = \text{const})\) it was assumed \(j_{g_1} = 0\), and at the lower and upper boundaries \((g_2 = \text{const})\) \(j_{g_2} = 0\). The minimization was performed using the steepest-descent method with a variable step-size. Starting with \(\Phi(g) = \Psi(g) = 0\) as an initial guess for these functions, in which case \(Q \approx 1 \times 10^{-3}\), the process was continued until \(Q\) was as small as \(\approx 1 \times 10^{-11}\), i.e. even in the worst case, when the deviation of the fluxes determined by Eqs. 3 from the fluxes obtained in the simulations was concentrated at a single point, it would not exceed \(\sim 0.01\%\).

![FIG. 4: (Color online) The functions \(\Phi(g_1, g_2)\) (panel a) and \(\Psi(g_1, g_2)\) (panel b).](image)

The results of the calculations are shown in Fig. 4. In agreement with the Helmholtz decomposition \[27\], the functions \(\Phi(g)\) and \(\Psi(g)\) are characteristically different in that the former accounts for the source and sink of the flows, and the latter for the vorticity effects. More specifically, as can be seen from comparison of Eqs. 3 with Fig. 4, the function \(\Phi(g)\) determines the intensity of the flow in the vicinity of the source and sink, and the function \(\Psi(g)\) determines it in the region between the source and sink. The function \(\Psi(g)\) also provides the canalization of the flow within a relatively narrow region of the conformation space, which can be associated with a folding funnel.

Considering the \((g_1, g_2)\) space as a complex number space, the functions \(\Phi(g)\) and \(\Psi(g)\) can be written in the form of a complex potential

\[
\Theta(g) = \Phi(g) + i\Psi(g)
\]

where \(i\) is the imaginary unit. Then, introducing the gradient operator in this space as \(\nabla = \partial / \partial g_1 + i \partial / \partial g_2\), we have

\[
j(g) = j_{g_1} + ij_{g_2} = -\nabla \Theta(g)
\]

It should be noted that since the flow is not divergence- and curl-free in the whole region of the conformation space (Fig. 2), the Cauchy-Riemann conditions, which present the necessary and sufficient condition for a function to be analytic \[27\], are not satisfied, i.e. the functions \(j(g)\) and \(\Theta(g)\) are not functions of complex variable \(g_1 + ig_2\). Therefore, the term "complex potential", which we use to refer to the function \(\Theta(g)\), should be considered not in the narrow sense, as in hydrodynamics, where the complex potential is an analytic function \[23\], but in a general sense, i.e. as a function that has real and imaginary components.

In summary, we have shown that a true potential of the driving force for protein folding is a function that has two components. One component accounts for the source and sink of the folding flows, representing, respectively, the unfolded and native states of the protein, and the other accounts for the vorticity effects at the boundaries of the main flow and provides the canalization of the flow within a relatively narrow region of conformation space that connects the unfolded and native states, i.e. the folding funnel.

---

[1] J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. USA 84, 7524 (1987).
[2] E. Shakhnovich, G. Farztdinov, A. M. Gutin, and M. Karplus, Phys. Rev. Lett. 67, 1665 (1991).
[3] P. E. Leopold, M. Montal, J. N. Onuchic, Proc. Natl. Acad. Sci. USA 89, 8721 (1992).
[4] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins: Structure, Function, and Bioinformatics 21, 167 (1995).
[5] K. A. Dill and H. S. Chan, Nat. Struct. Biol. 4, 10 (1997).
[6] E. I. Shakhnovich, Curr. Opin. Struct. Biol. 7, 20 (1997).
[7] J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, Annu.

[8] C. M. Dobson, A. Sali, and M. Karplus, Angew. Chem. Int. Ed. 37, 869 (1998).
[9] A. R. Dinner, A. Šali, L. J. Smith, C. M. Dobson, and M. Karplus, Trends Biochem. Sci. 25, 331 (2000).
[10] J.-E. Shea and C. L. Brooks III, Annu. Rev. Phys. Chem. 52, 499 (2001).
[11] B. Schuler, E. A. Lipman, and W. A. Eaton, Nature 419, 743 (2002).
[12] M. Gruebele, Curr. Opinion Struct. Biol. 12, 161 (2002).
[13] C. C. Mello and D. Barrick, Proc. Natl. Acad. Sci. USA 101, 14102 (2004).
[14] M. Karplus, Nat. Chem. Biol. 7, 401 (2011).
[15] K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl, Annu. Rev. Biophys. 37, 289316 (2008).
[16] K. A. Dill and J. L. MacCallum, Science 338, 1042 (2012).
[17] A. V. Finkelstein and O. Ptitsyn, Protein Physics: A Course of Lectures (Academic Press: London, 2002).
[18] E. M. Bocchio and C. L. Brooks III, Science 269, 393 (1995).
[19] N. D. Socci, J. N. Onuchic, and P. G. Wolynes, Proteins: Structure, Function, and Genetics 32, 136 (1998).
[20] S. F. Chekmarev, S. V. Krivov, and M. Karplus, J. Phys. Chem. B 109, 5312 (2005).
[21] S. F. Chekmarev, A. Yu. Palyanov, and M. Karplus, Phys. Rev. Lett. 100, 018107 (2008).
[22] F. Rao and A. Caflisch, J. Mol. Biol. 342, 299 (2004).
[23] L. D. Landau and E. M. Lifshitz, Fluid Mechanics (Pergamon: New York, 1987).
[24] I. V. Kalgin, M. Karplus, and S. F. Chekmarev, J. Phys. Chem. B 113, 12759 (2009).
[25] I. V. Kalgin and S. F. Chekmarev, Phys. Rev. E 83, 011920 (2011).
[26] I. V. Kalgin, A. Caflisch, S. F. Chekmarev, and M. Karplus, J. Phys. Chem. B (to be published).
[27] G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists, 4th ed. (Academic Press: San Diego, 1995).
[28] N. H. Andersen, K. A. Olsen, R. M. Fesinmeyer, X. Tan, F. M. Hudson, L. A. Eidenschink, and S. R. Farazi, J. Am. Chem. Soc. 128, 6101 (2006).
[29] N. Gö, Annu. Rev. Biophys. Bioeng. 12, 183 (1983).
[30] T. X. Hoang and M. Cieplak, J. Chem. Phys. 112, 6851 (2000).
[31] R. Biswas and D. R. Hamann, Phys. Rev. B 34, 895 (1986).
[32] M. Oliveberg, Y.-J. Yan, and A. R. Fersht, Proc. Natl. Acad. Sci. U.S.A. 92, 8926 (1995).
[33] M. Karplus, Fold. Des. 2, 569 (1997).
[34] I. T. Jolliffe, Principal Component Analysis; 2nd ed. (Springer: New York, 2002).
[35] Although the second principal component could also be chosen as $g_2$, because the first two eigenvalues are well separated from the others (Fig. S5), the present choice makes it possible to take into account all effects missed by the first component.
[36] Y. Levy, J. Jortner, and O. M. Becker, Proc. Natl. Acad. Sci. U.S.A. 98, 2188 (2001).
[37] S. V. Krivov and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. 101, 14766 (2004).
[38] M. A. Rohrdanz, W. Zheng, and C. Clementi, Annu. Rev. Phys. Chem. 64, 295 (2013).
Supplementary Material for “Protein folding: A complex potential for the driving force” by S. F. Chekmarev

I. NATIVE STATE OF THE PROTEIN

FIG. S1: The “bead” model of the native conformation of the 2evq de novo protein [28].

II. PROTEIN MODEL AND SIMULATION METHOD

The interaction potential of Ref. [30] is written as follows

$$U = \sum_{i=1}^{N-1} [k_1(r_{i,i+1} - d_0)^2 + k_2(r_{i,i+1} - d_0)^4]$$

$$+ \sum_{i+1<j}^{\text{NAT}} 4\varepsilon \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

$$\times \left[ \left( \frac{\sigma_0}{r_{ij}} \right)^{12} - \left( \frac{\sigma_0}{r_{ij}} \right)^6 + 1 \right] \Delta(r_{ij} - d_{\text{nat}})$$

where $N$ is the number of monomers (residues), $r_{ij}$ is the current distance between two monomers $i$ and $j$, and $\varepsilon$ is the characteristic attractive energy. The first term represents rigidity of the backbone; here $d_0 = 3.8\text{Å}$, $k_1 = \varepsilon/\text{Å}^2$, and $k_2 = 100\varepsilon/\text{Å}^4$. The second and third terms describe contributions of the native and non-native contacts, respectively. The contact between two monomers $i$ and $j$ was considered as a native contact if $|j-i| > 1$ and the distance between these monomers in the native state was less than $7.5\text{Å}$. In the second term $\sigma_{ij} = 2^{1/6}d_{ij}$, where $d_{ij}$ is the distance between monomers $i$ and $j$ in the native state, and in the third term $d_{\text{nat}} = \langle d_{ij} \rangle$, $\sigma_0 = 2^{1/6}d_{\text{nat}}$, and $\Delta(r_{ij} - d_{\text{nat}})$ is the cutoff function which is equal to 1 for $r_{ij} < d_{\text{nat}}$ and 0 otherwise.

The simulations were performed using a constant-temperature molecular dynamics based on the Langevin equation

$$m \frac{d^2r_i}{dt^2} + \gamma \frac{dr_i}{dt} = -\frac{\partial U}{\partial r_i} + \Phi_i(t)$$

where $r_i$ is the radius-vector of $i$ monomer representing a residue, $m$ is the monomer mass, $U$ is the potential energy of the system, $\Gamma_i$ is a random force from the surroundings (a solvent that is not considered explicitly), and $\gamma$ is the friction coefficient that introduces viscosity of the surroundings to balance the random force and dissipation. The random forces have the Gaussian distribution with zero mean and variance $\langle \Phi_i^2(t)\Phi_i^2(t + \tau) \rangle = 2\gamma k_B T \delta_{ij}\delta_{ij}\delta(\tau)$, where the angular brackets denote an ensemble average, $T$ is the temperature, $k_B$ is the Boltzmann constant, the upper index at $\Phi$ stands for the vector component, and $\delta_{ik}$ and $\delta(\tau)$ are the Kronecker and Dirac delta functions, respectively. The equation was numerically integrated using the algorithm of Ref. [31] with the time step $\Delta t = 0.005\tau$ and $\gamma = 3m/\tau$. With the length scale $l = 7.5\text{Å}$ and the attractive energy $\varepsilon \sim 1\text{kcal/mol}$, the characteristic time scale $\tau = (m\varepsilon/\gamma)^{1/2}$ is $\sim 1\text{ps}$, where $m = 110\text{Da}$.

III. FREE ENERGY SURFACE IN “PHYSICAL” COLLECTIVE VARIABLES

FIG. S2: The free energy surface as a function of the fraction of native contacts and radius of gyration, $T = 0.17$.

IV. FIRST-PASSAGE TIMES

FIG. S3: The mean first-passage time as a function of temperature. The arrow indicates the temperature ($T = 0.17$) at which the study was performed.
FIG. S4: The distribution of the first-passage times, $T = 0.17$.

V. SPECTRUM OF EIGENVALUES

FIG. S5: Spectrum of eigenvalues.

VI. TOTAL FOLDING FLOWS

FIG. S6: The total flow from the denatured states to the native state as a function of the longitudinal variable $g_1$. The blue curve depicts the total flow as it was obtained in the simulations. The product of the maximum value of the flow ($G \approx 1.4338 \times 10^{-2}$) by the MFPT ($t_f \approx 69.7368$) is close to 1 with very good accuracy. The red curve shows the total flow calculated from the fluxes of Fig. S7; the value of the flow is reduced by $1 \times 10^4$ times. The average value of the flow is very close to zero.

VII. VECTOR FLOW FIELD FROM THE FES

FIG. S7: Vector flow field calculated from the FES of Fig. S2 with the fluxes determined as $j_{g_1} = -\partial F(g)/\partial g_1$ and $j_{g_2} = -\partial F(g)/\partial g_2$, where $F(g)$ is the free energy function of Fig. S2.