A Constrained Saddle Dynamics for Computing Electronic Excited States

Yukuan Hu

Centre d'Enseignement et de Recherche en MathématIques et Calcul Scientifique
École nationale des ponts et chaussées, Institut Polytechnique de Paris

CERMICS Young Researchers Seminar
13 January, 2026

Time-Independent Many-Body Schrödinger Equation (SE)

Pauli '25; Schrödinger '26; Born, Oppenheimer '27

$$\underbrace{\left(-\frac{1}{2}\sum_{i=1}^{N}\Delta_{\boldsymbol{r}_i} - \sum_{i=1}^{N}\sum_{I=1}^{M}\frac{z_I}{|\boldsymbol{r}_i - \mathbf{R}_I|} + \sum_{1\leq i<j\leq N}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\right)}_{\widehat{H}_{N,\{\mathbf{R}_I\}} \equiv: \widehat{H}} \Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = E\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$$

- $\{\mathbf{R}_I\}_{I=1}^{M} \subseteq \mathbb{R}^3$: prefixed nuclear configuration, with charges $\{z_I\}_{I=1}^{M} \subseteq \mathbb{N}_+$
- $\Psi \in \wedge^N L^2(\mathbb{R}^3; \mathbb{C})$: normalized antisymmetric wavefunction
- $\widehat{H}$: Hermitian electronic Hamiltonian
- $E \in \sigma(\widehat{H})$: energy associated with Ψ

Time-Independent Many-Body Schrödinger Equation (SE)

Pauli '25; Schrödinger '26; Born, Oppenheimer '27

$$\underbrace{\left(-\frac{1}{2}\sum_{i=1}^{N}\Delta_{\boldsymbol{r}_i} - \sum_{i=1}^{N}\sum_{I=1}^{M}\frac{z_I}{|\boldsymbol{r}_i - \mathbf{R}_I|} + \sum_{1\le i<j\le N}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\right)}_{\widehat{H}_{N,\{\mathbf{R}_I\}}\equiv:\widehat{H}}\Psi(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N) = E\Psi(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N)$$

Fig. 1. Water molecule (cr. CORDIS)

Example. Water molecule H_2O

$M = 3,\ N = 10,\ z_1 = 8,\ z_2 = z_3 = 1$

Electronic Ground and Excited States

Zhislin's theorem Zhislin '60

If $N \leq \sum_{I=1}^{M} z_I$ (neutral or positively charged), then $\sigma(\widehat{H})$ looks like

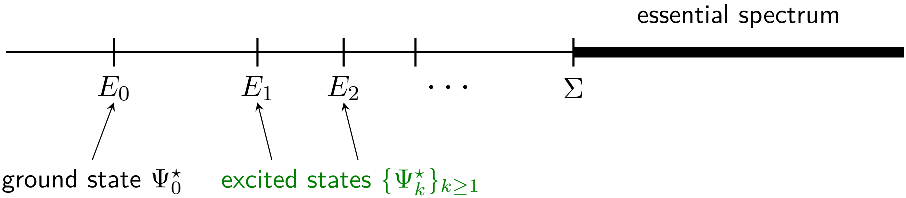

Fig. 2. Spectrum of $\widehat{H}$

Excitation energies: $E_1 - E_0,\ E_2 - E_0,\ \ldots$

Electronic Ground and Excited States

Ground/excited states are critical points

$$\mathcal{E}(\Gamma) := \mathrm{Tr}(\widehat{H}\Gamma) \quad \text{with} \quad \Gamma \in \mathcal{M}^{\mathrm{exact}} := \mathbf{P}\Big(\bigwedge^{N} L^2(\mathbb{R}^3;\mathbb{C})\Big)$$

- $\Gamma := \Psi\Psi^*$: density matrix

Electronic Ground and Excited States

Ground/excited states are critical points

$$\mathcal{E}(\Gamma) := \mathrm{Tr}(\widehat{H}\Gamma) \quad \text{with} \quad \Gamma \in \mathcal{M}^{\text{exact}} := \mathbf{P}\Big(\bigwedge^{N} L^2(\mathbb{R}^3;\mathbb{C})\Big)$$

- Riemannian gradient and stationary condition

$$\mathrm{grad}_{\mathcal{M}^{\text{exact}}}\,\mathcal{E}(\Gamma) = [\Gamma,[\Gamma,\widehat{H}]] = 0 \quad \Leftrightarrow \quad \widehat{H}\Psi = \mathcal{E}(\Gamma)\Psi \ \text{(many-body SE)}$$

- Spectrum of the Riemannian Hessian at the critical point $\Gamma_k^\star := \Psi_k^\star(\Psi_k^\star)^*$

$$\sigma\big(\mathrm{Hess}_{\mathcal{M}^{\text{exact}}}\,\mathcal{E}(\Gamma_k^\star)\big) = \{E_0 - E_k, E_1 - E_k, \ldots\} \quad \text{with} \quad k \text{ of them} < 0$$

💡 The k-th excited state $=$ The index-k saddle point of $\mathcal{E}$

Electronic Ground and Excited States

Photo-induced applications Turro, Ramamurthy, Scaiano '09; Balzani, Ceroni, Juris '14; Mai, González '20; ...

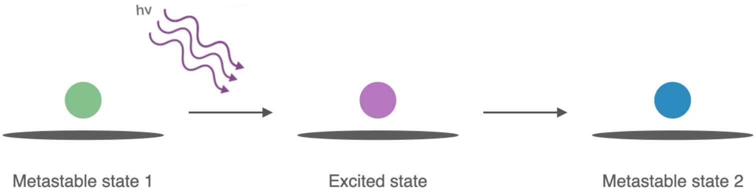

Fig. 3. Photochemical reactions

Electronic Ground and Excited States

Photo-induced applications Turro, Ramamurthy, Scaiano '09; Balzani, Ceroni, Juris '14; Mai, González '20; ...

Curse of dimensionality

$$\sum_{1 \le i < j \le N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \quad \rightsquigarrow \quad \text{Many-body SE is not decomposable}$$

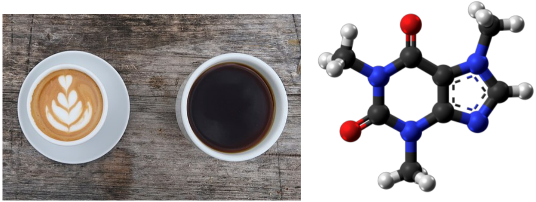

Fig. 4. Coffee and the caffeine molecule (cr. Wikipedia)

Caffeine molecule $C_8H_{10}N_4O_2$

$N = 102,\ M = 24$

↳ 306-dim problem

Approximate Levels of Theory in Quantum Chemistry (QC)

Geometrical idea behind variational approximations

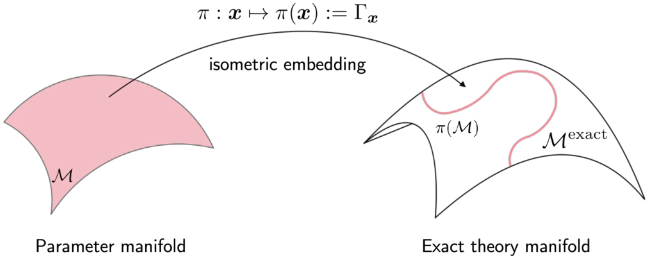

Fig. 5. Isometric embedding from the parameter manifold to the exact theory manifold

💡 Finding suitable embedded Riemannian submanifolds of $\mathcal{M}^{\text{exact}}$

Approximate Levels of Theory in Quantum Chemistry (QC)

- Wavefunction methods
  - Hartree-Fock (HF) methods Hartree '28; Fock '30

    ↪ Grassmann manifold

$$\mathcal{M}^{\mathrm{HF}} := \mathrm{Gr}_N(\mathbb{C}^{n_\mathrm{b}}) := \left\{ P \in \mathbb{C}^{n_\mathrm{b} \times n_\mathrm{b}}_{\mathrm{herm}} \mid P^2 = P^* = P,\ \mathrm{rank}(P) = N \right\}$$

where $n_\mathrm{b} \in \mathbb{N}$ denotes the number of basis functions

Variants: restricted/unrestricted/restricted open-shell/generalized HF, ...

Approximate Levels of Theory in Quantum Chemistry (QC)

- Wavefunction methods
  - HF methods $\rightsquigarrow \mathcal{M}^{\mathrm{HF}}$
  - Post-HF methods: full configuration-interaction (FCI)

    $\hookrightarrow$ Projective space $\mathcal{M}^{\mathrm{FCI}} := \mathbf{P}\left(\mathbb{S}_{\mathbb{C}}^{n_{\mathrm{det}}-1}\right)$, where $n_{\mathrm{det}} := \binom{n_{\mathrm{b}}}{N}$

  Others: truncated CI, coupled cluster methods, Møller-Plesset perturbation theory, ...

Approximate Levels of Theory in Quantum Chemistry (QC)

- Wavefunction methods
  - HF methods $\rightsquigarrow \mathcal{M}^{\mathrm{HF}}$
  - Post-HF methods: FCI $\rightsquigarrow \mathcal{M}^{\mathrm{FCI}}$
  - Multi-reference methods: complete active space self-consistent field (CASSCF) Roos et al. '80

    $\hookrightarrow \mathcal{M}^{\mathrm{CAS}} := \left(\mathbb{S}_{\mathbb{C}}^{n_{\mathrm{det}}-1} \times \mathcal{U}(n_{\mathrm{b}})\right) / \left(\mathcal{U}(1) \times \mathcal{U}(n_{\mathrm{int}}) \times \mathcal{U}(n_{\mathrm{act}}) \times \mathcal{U}(n_{\mathrm{b}} - n_{\mathrm{int}} - n_{\mathrm{act}})\right)$

    where $n_{\mathrm{int}}, n_{\mathrm{act}} \in \mathbb{N}$ $(n_{\mathrm{int}} < N < n_{\mathrm{int}} + n_{\mathrm{act}} < n_{\mathrm{b}})$, $n_{\mathrm{det}} := \binom{n_{\mathrm{act}}}{N - n_{\mathrm{int}}}$

    Others: multi-reference CI/coupled cluster/perturbation theory, ...

Approximate Levels of Theory in Quantum Chemistry (QC)

- Wavefunction methods
  - HF methods $\rightsquigarrow \mathcal{M}^{\mathrm{HF}}$
  - Post-HF methods: FCI $\rightsquigarrow \mathcal{M}^{\mathrm{FCI}}$
  - Multi-reference methods: CASSCF $\rightsquigarrow \mathcal{M}^{\mathrm{CAS}}$
- Density functional theory (DFT) Hohenberg, Kohn '64; Kohn, Sham '65
  - Mean-field model + Semilocal exchange-correlation functionals

    $\hookrightarrow$ Grassmann manifold $\mathcal{M}^{\mathrm{DFT}} := \mathcal{M}^{\mathrm{HF}}$

Approximate Levels of Theory in Quantum Chemistry (QC)

- Wavefunction methods
  - HF methods $\rightsquigarrow \mathcal{M}^{\mathrm{HF}}$
  - Post-HF methods: FCI $\rightsquigarrow \mathcal{M}^{\mathrm{FCI}}$
  - Multi-reference methods: CASSCF $\rightsquigarrow \mathcal{M}^{\mathrm{CAS}}$
- Density functional theory (DFT) $\rightsquigarrow \mathcal{M}^{\mathrm{DFT}}$
- Other theories not covered by this talk
  - Quantum Monte Carlo McMillan '65
  - Many-body Green's function metods Luttinger, Ward '60
  - Semi-empirical and machine learning force fields Thiel '14; von Lilienfeld et al. '20; ...

🥳 More details in Laura's talk

Approximate Levels of Theory in Quantum Chemistry (QC)

Consequence of nonlinearity – Stationary conditions

- Stationary condition in the exact theory ($\mathcal{M}^{\text{exact}}$)

$$0 = \operatorname{grad}_{\mathcal{M}^{\text{exact}}} \mathcal{E}(\Gamma) = [\Gamma, [\Gamma, \widehat{H}]] \quad \Leftrightarrow \quad \widehat{H}\Psi = \mathcal{E}(\Gamma)\Psi$$

- Stationary condition in approximate theories ($\pi(\mathcal{M}) \subsetneq \mathcal{M}^{\text{exact}}$)

$$0 = \operatorname{grad}_{\pi(\mathcal{M})} \mathcal{E}(\Gamma) = [\Gamma, [\Gamma, \widehat{H}_\Gamma]] \quad \Leftrightarrow \quad \widehat{H}_\Gamma\Psi = \operatorname{Tr}(\widehat{H}_\Gamma\Gamma)\Psi$$

where $\widehat{H}_\Gamma := \operatorname{Proj}_\Gamma \circ \widehat{H} \circ \operatorname{Proj}_\Gamma$ is the Hamiltonian projected onto $\mathrm{T}_\Gamma\pi(\mathcal{M})$

Linear eigenvalue problem $\xrightarrow[\text{theories}]{\text{approximation}}$ Nonlinear eigenvalue problem

Approximate Levels of Theory in Quantum Chemistry (QC)

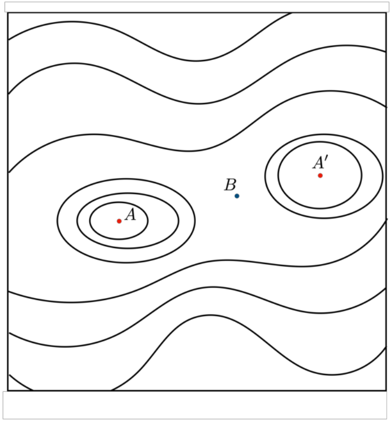

Fig. 6. Saddle points of approximations (cr. Laura)

Consequence of approximations – Critical points

$$\mathcal{E}(\Gamma) \quad \text{with} \quad \Gamma \in \mathcal{M}^{\text{exact}}$$

(excited states = saddle points on $\mathcal{M}^{\text{exact}}$)

Approximate Levels of Theory in Quantum Chemistry (QC)

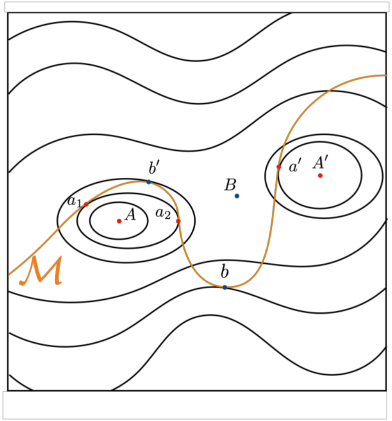

Fig. 6. Saddle points of approximations
(cr. Laura)

Consequence of approximations – Critical points

$$\mathcal{E}(\Gamma) \quad \text{with} \quad \Gamma \in \pi(\mathcal{M})$$

(excited states $\approx$ saddle points on $\pi(\mathcal{M})$)

↳ Index mismatch, "spurious" saddle points

Approximate Levels of Theory in Quantum Chemistry (QC)

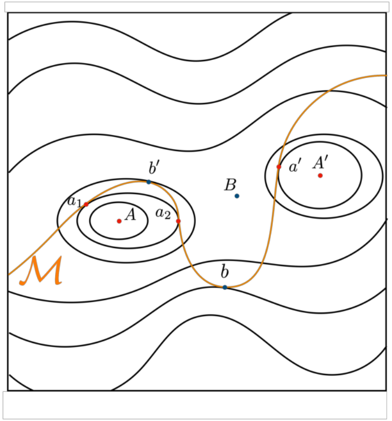

Fig. 6. Saddle points of approximations (cr. Laura)

Consequence of approximations – Critical points

$$\mathcal{E}(\Gamma) \quad \text{with} \quad \Gamma \in \pi(\mathcal{M})$$

(excited states $\approx$ saddle points on $\pi(\mathcal{M})$)

↳ Index mismatch, "spurious" saddle points

- Exhaust the saddle points on $\mathcal{M}$
- Analyze the associated states by QC techniques

Difficulties in finding saddle points

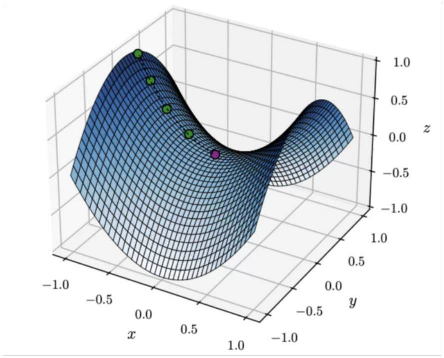

Fig. 7. Instability of saddle points

Example. $f(x, y) = x^2 - y^2$, $\mathcal{M} = \mathbb{R}^2$

$\hookrightarrow$ $(0, 0)$ is an index-1 saddle point

$$(x(0), y(0)) = (-1, 0)$$

$\Downarrow$ gradient flow

$$x(t) = -e^{-2t} \to 0, \ y(t) \equiv 0$$

Difficulties in finding saddle points

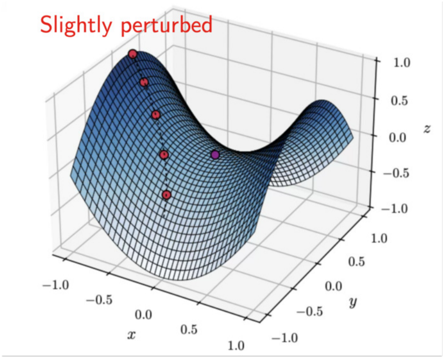

Fig. 7. Instability of saddle points

Example. $f(x, y) = x^2 - y^2$, $\mathcal{M} = \mathbb{R}^2$

$\hookrightarrow$ $(0, 0)$ is an index-1 saddle point

$$(x(0), y(0)) = (-1, -\varepsilon)$$

$\Downarrow$ gradient flow

$$x(t) = -e^{-2t} \to 0, \; y(t) = -\varepsilon \cdot e^{2t} \to -\infty$$

Existing Methods for Computing Excited States

Difficulties in finding saddle points

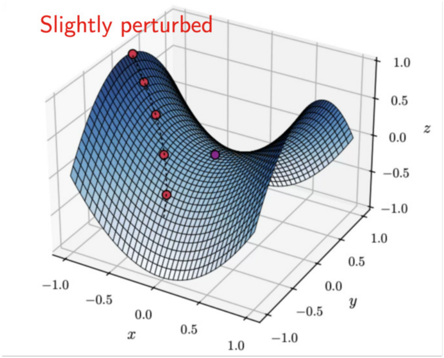

Fig. 7. Instability of saddle points

Example. $f(x,y) = x^2 - y^2$, $\mathcal{M} = \mathbb{R}^2$

↳ $(0,0)$ is an index-1 saddle point

$$(x(0), y(0)) = (-1, -\varepsilon)$$

$\Downarrow$ gradient flow

$$x(t) = -e^{-2t} \to 0, \; y(t) = -\varepsilon \cdot e^{2t} \to -\infty$$

😢 Instability of optimization methods

😭 Lack of a global merit function

Existing Methods for Computing Excited States (QC)

State-specific methods

- (Quasi-)Newton direct optimization Cerjan, Miller '81; Jensen, Jørgensen '84; ...

  ↳ Instability outside attractive regions, expensive calculations

Existing Methods for Computing Excited States (QC)

State-specific methods

- (Quasi-)Newton direct optimization Cerjan, Miller '81; Jensen, Jørgensen '84; ...
- ΔSCF (for HF & DFT), simple root selection (for CASSCF) Jones, Gunnarsson '89; ...
  - Maximum overlap method Gilbert, Besley, Gill '08; Tran, Shea, Neuscamman '19; ...
  - State-targeted energy projection (for HF & DFT) Carter-Fenk, Herbert '20; ...

↳ Convergence issue, the tracked “index” $\neq$ saddle point index

Existing Methods for Computing Excited States (QC)

State-specific methods

- (Quasi-)Newton direct optimization Cerjan, Miller '81; Jensen, Jørgensen '84; ...
- ΔSCF (for HF & DFT), simple root selection (for CASSCF) Jones, Gunnarsson '89; ...
  - Maximum overlap method Gilbert, Besley, Gill '08; Tran, Shea, Neuscamman '19; ...
  - State-targeted energy projection (for HF & DFT) Carter-Fenk, Herbert '20; ...
- Square gradient minimization Hait, Head-Gordon '20; ...
- Generalized variational principles Shea, Gwin, Neuscamman '20; Hanscam, Neuscamman '22; ...

↳ Spurious solutions, expensive calculations, lack of *a priori* knowledge

Existing Methods for Computing Excited States (QC)

State-specific methods

- (Quasi-)Newton direct optimization Cerjan, Miller '81; Jensen, Jørgensen '84; ...
- ΔSCF (for HF & DFT), simple root selection (for CASSCF) Jones, Gunnarsson '89; ...
  - Maximum overlap method Gilbert, Besley, Gill '08; Tran, Shea, Neuscamman '19; ...
  - State-targeted energy projection (for HF & DFT) Carter-Fenk, Herbert '20; ...
- Square gradient minimization Hait, Head-Gordon '20; ...
- Generalized variational principles Shea, Gwin, Neuscamman '20; Hanscam, Neuscamman '22; ...

Other methods in QC

- Orthogonality-constrained methods Bustard, Jaffe '70; Levy, Nagy '99; Surján '00; Yalouz, Robert '23; ...
- Linear response theory Olsen, Jørgensen '85; Casida '95; ...
- State-averaged methods Werner, Meyer '81; ... $\hookrightarrow$ Not aiming for saddle points on $\mathcal{M}$

Existing Methods for Computing Excited States (Math)

Zhang, Du '12; Gao, Leng, Zhou '15; Li, Lu, Yang '15; Yin, Huang, Zhang '22; Liu, Xie, Yuan '23

Manifolds with global regular level-set representations

$$\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}(\boldsymbol{x}) = 0\}, \quad \boldsymbol{c} := (c_1, \ldots, c_p)^\top$$

where $\operatorname{grad} \boldsymbol{c}(\boldsymbol{x}) \in \mathbb{R}^{p \times n}$ is of full row rank for any $\boldsymbol{x} \in \mathcal{M}$

Existing Methods for Computing Excited States (Math)

Zhang, Du '12; Gao, Leng, Zhou '15; Li, Lu, Yang '15; Yin, Huang, Zhang '22; Liu, Xie, Yuan '23

Manifolds with global regular level-set representations

$$\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}(\boldsymbol{x}) = 0\}, \quad \boldsymbol{c} := (c_1, \ldots, c_p)^\top$$

Constrained saddle dynamics: for index-k saddle points

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\left(I_n - \sum_{i=1}^{k} \boldsymbol{v}_i \boldsymbol{v}_i^\top\right) \operatorname{grad}_{\mathcal{M}} f(\boldsymbol{x}) + \sum_{i=1}^{k} \boldsymbol{v}_i \left(\boldsymbol{v}_i^\top \operatorname{grad}_{\mathcal{M}} f(\boldsymbol{x})\right)$$

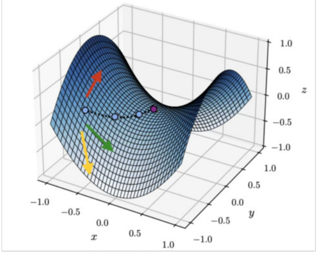

Fig. 8. Trajectory of dynamics

Existing Methods for Computing Excited States (Math)

Zhang, Du '12; Gao, Leng, Zhou '15; Li, Lu, Yang '15; Yin, Huang, Zhang '22; Liu, Xie, Yuan '23

Manifolds with global regular level-set representations

$$\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}(\boldsymbol{x}) = 0\}, \quad \boldsymbol{c} := (c_1, \ldots, c_p)^\top$$

Constrained saddle dynamics: for index-k saddle points

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\left(I_n - \sum_{i=1}^{k} \boldsymbol{v}_i \boldsymbol{v}_i^\top\right) \mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}) + \sum_{i=1}^{k} \boldsymbol{v}_i \left(\boldsymbol{v}_i^\top \mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x})\right)$$

$$\frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = -(I_n - \boldsymbol{v}_i \boldsymbol{v}_i^\top) \mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x})[\boldsymbol{v}_i] + 2 \sum_{j=1}^{i-1} \boldsymbol{v}_j \left(\boldsymbol{v}_j^\top \mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x})[\boldsymbol{v}_j]\right)$$
$$- \mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x})^\top \left(\mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x}) \cdot \mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x})^\top\right)^{-1} \left(\mathrm{Hess}\, \boldsymbol{c}(\boldsymbol{x})\left[\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, \boldsymbol{v}_i\right]\right)$$

Rayleigh-Ritz minimization

Operator splitting

Zhang, Du '12; Gao, Leng, Zhou '15; Li, Lu, Yang '15; Yin, Huang, Zhang '22; Liu, Xie, Yuan '23

Manifolds with global regular level-set representations

$$\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}(\boldsymbol{x}) = 0\}, \quad \boldsymbol{c} := (c_1, \ldots, c_p)^\top$$

Constrained saddle dynamics: for index-k saddle points

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\left(I_n - \sum_{i=1}^{k} \boldsymbol{v}_i \boldsymbol{v}_i^\top\right) \operatorname{grad}_{\mathcal{M}} f(\boldsymbol{x}) + \sum_{i=1}^{k} \boldsymbol{v}_i \left(\boldsymbol{v}_i^\top \operatorname{grad}_{\mathcal{M}} f(\boldsymbol{x})\right)$$

$$\frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = -(I_n - \boldsymbol{v}_i \boldsymbol{v}_i^\top) \operatorname{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\boldsymbol{v}_i] + 2 \sum_{j=1}^{i-1} \boldsymbol{v}_j \left(\boldsymbol{v}_j^\top \operatorname{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\boldsymbol{v}_j]\right)$$

$$- \operatorname{grad} \boldsymbol{c}(\boldsymbol{x})^\top \left(\operatorname{grad} \boldsymbol{c}(\boldsymbol{x}) \cdot \operatorname{grad} \boldsymbol{c}(\boldsymbol{x})^\top\right)^{-1} \left(\operatorname{Hess} \boldsymbol{c}(\boldsymbol{x}) \left[\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, \boldsymbol{v}_i\right]\right)$$

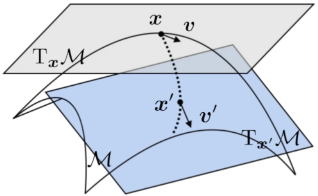

Fig. 9. Varying tangent space

Existing Methods for Computing Excited States (Math)

Zhang, Du '12; Gao, Leng, Zhou '15; Li, Lu, Yang '15; Yin, Huang, Zhang '22; Liu, Xie, Yuan '23

Manifolds with global regular level-set representations

$$\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}(\boldsymbol{x}) = 0\}, \quad \boldsymbol{c} := (c_1, \ldots, c_p)^\top$$

Constrained saddle dynamics: for index-k saddle points

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\left(I_n - \sum_{i=1}^{k} \boldsymbol{v}_i \boldsymbol{v}_i^\top\right) \mathrm{grad}_{\mathcal{M}} f(\boldsymbol{x}) + \sum_{i=1}^{k} \boldsymbol{v}_i \left(\boldsymbol{v}_i^\top \mathrm{grad}_{\mathcal{M}} f(\boldsymbol{x})\right)$$

$$\frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = -(I_n - \boldsymbol{v}_i \boldsymbol{v}_i^\top) \mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\boldsymbol{v}_i] + 2 \sum_{j=1}^{i-1} \boldsymbol{v}_j \left(\boldsymbol{v}_j^\top \mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\boldsymbol{v}_j]\right)$$
$$- \mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x})^\top \left(\mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x}) \cdot \mathrm{grad}\, \boldsymbol{c}(\boldsymbol{x})^\top\right)^{-1} \left(\mathrm{Hess}\, \boldsymbol{c}(\boldsymbol{x}) \left[\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, \boldsymbol{v}_i\right]\right)$$

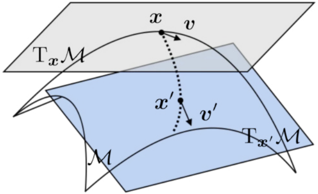

Fig. 9. Varying tangent space

Restriction to special cases (e.g., flat space, sphere, Stiefel)

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

- $\boldsymbol{x}$-dynamics

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}(t)) \quad \rightsquigarrow \quad \boldsymbol{v}(t) \in \mathrm{T}_{\boldsymbol{x}(t)}\mathcal{M}$$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

- $\boldsymbol{x}$-dynamics: $\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}} f(\boldsymbol{x}(t))$
- $\boldsymbol{v}$-dynamics:

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{Proj}_{\boldsymbol{x}(t)}(\boldsymbol{v}(t))\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{Proj}_{\boldsymbol{x}(t)}\big)(\boldsymbol{v}(t)) + \mathrm{Proj}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t)\right) \\
&= \mathrm{II}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}(t)\right) + \mathrm{Proj}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t)\right)
\end{aligned}$$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

- $\boldsymbol{x}$-dynamics: $\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}} f(\boldsymbol{x}(t))$
- $\boldsymbol{v}$-dynamics:

$$
\begin{aligned}
\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{Proj}_{\boldsymbol{x}(t)}(\boldsymbol{v}(t))\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{Proj}_{\boldsymbol{x}(t)}\right)(\boldsymbol{v}(t)) + \mathrm{Proj}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t)\right) \\
&= \mathrm{II}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}(t)\right) + \mathrm{Proj}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t)\right)
\end{aligned}
$$

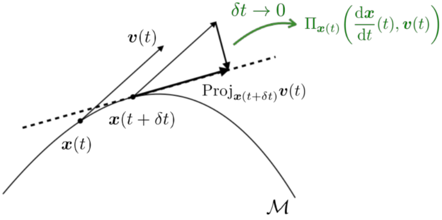

Fig. 10. Second fundamental form

$\mathrm{II}_{\boldsymbol{x}} : \mathrm{T}_{\boldsymbol{x}}\mathcal{M} \times \mathrm{T}_{\boldsymbol{x}}\mathcal{M} \to \mathrm{N}_{\boldsymbol{x}}\mathcal{M}$: second fundamental form of $\mathcal{M}$ at $\boldsymbol{x}$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

- $\boldsymbol{x}$-dynamics: $\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}} f(\boldsymbol{x}(t))$
- $\boldsymbol{v}$-dynamics:

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) = \underbrace{\mathrm{II}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}(t)\right)}_{\text{varying tangent space}} + \underbrace{\mathrm{Proj}_{\boldsymbol{x}(t)}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t)\right)}_{\text{tracking the lowest eigvec.}} \in \mathrm{N}_{\boldsymbol{x}(t)}\mathcal{M} \oplus \mathrm{T}_{\boldsymbol{x}(t)}\mathcal{M}$$

Following the Euler-Lagrange equation of the Rayleigh-Ritz minimization

$$\begin{aligned} \min_{\tilde{\boldsymbol{v}}} \quad & \langle \tilde{\boldsymbol{v}}, \mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\tilde{\boldsymbol{v}}]\rangle \\ \text{s. t.} \quad & \tilde{\boldsymbol{v}} \in \mathrm{T}_{\boldsymbol{x}}\mathcal{M}, \ \|\tilde{\boldsymbol{v}}\| = 1 \end{aligned} \quad \rightsquigarrow \quad \mathrm{Proj}_{\boldsymbol{x}}\left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}\right) = -(\mathrm{Id}_{\boldsymbol{x}} - \boldsymbol{v}\boldsymbol{v}^\top)\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x})[\boldsymbol{v}]$$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}(t)) \\ \dfrac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) = -\big(\mathrm{Id}_{\boldsymbol{x}(t)} - \boldsymbol{v}(t)\boldsymbol{v}(t)^\top\big)\mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x}(t))[\boldsymbol{v}(t)] + \mathrm{II}_{\boldsymbol{x}(t)}\left(\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}(t)\right) \end{cases}$$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Index-1 case

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2\boldsymbol{v}(t)\boldsymbol{v}(t)^\top)\mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}(t)) \\ \dfrac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) = -\big(\mathrm{Id}_{\boldsymbol{x}(t)} - \boldsymbol{v}(t)\boldsymbol{v}(t)^\top\big)\mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x}(t))[\boldsymbol{v}(t)] + \mathrm{II}_{\boldsymbol{x}(t)}\left(\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}(t)\right) \end{cases}$$

Index-k case: let $V := (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \in \mathrm{St}_k(\mathrm{T}_{\boldsymbol{x}}\mathcal{M})$

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = -(\mathrm{Id}_{\boldsymbol{x}(t)} - 2V(t)V(t)^\top)\mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}(t)) \\ \dfrac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t}(t) = -\big(\mathrm{Id}_{\boldsymbol{x}(t)} - V(t)V(t)^\top\big)\mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x}(t))[\boldsymbol{v}_i(t)] + \mathrm{II}_{\boldsymbol{x}(t)}\left(\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \boldsymbol{v}_i(t)\right) \end{cases}$$

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Another perspective: dynamics on the Grassmann bundle of $\mathrm{T}\mathcal{M}$

Stiefel and Grassmann bundles of $\mathrm{T}\mathcal{M}$ Lee '12

$$\mathrm{St}_k(\mathrm{T}\mathcal{M}) := \left\{ \left(\boldsymbol{x}, V\right) \mid \boldsymbol{x} \in \mathcal{M},\ V \in \mathrm{St}_k(\mathrm{T}_{\boldsymbol{x}}\mathcal{M}) \right\}$$

$$\Downarrow \ \sim \ \text{by } \mathcal{O}(k)$$

$$\mathrm{Gr}_k(\mathrm{T}\mathcal{M}) := \left\{ \left(\boldsymbol{x}, [V]\right) \mid \boldsymbol{x} \in \mathcal{M},\ [V] \in \mathrm{Gr}_k(\mathrm{T}_{\boldsymbol{x}}\mathcal{M}) \right\}$$

are themselves embedded submanifolds of proper ambient spaces

Constrained Saddle Dynamics (CSD) on General Riemannian Manifolds

Another perspective: dynamics on the Grassmann bundle of $\mathrm{T}\mathcal{M}$

Stiefel and Grassmann bundles of $\mathrm{T}\mathcal{M}$ Lee '12

$$\mathrm{St}_k(\mathrm{T}\mathcal{M}) := \left\{ \left(\boldsymbol{x}, V\right) \mid \boldsymbol{x} \in \mathcal{M},\ V \in \mathrm{St}_k(\mathrm{T}_{\boldsymbol{x}}\mathcal{M}) \right\}$$

$$\Downarrow \ \sim \ \text{by } \mathcal{O}(k)$$

$$\mathrm{Gr}_k(\mathrm{T}\mathcal{M}) := \left\{ \left(\boldsymbol{x}, [V]\right) \mid \boldsymbol{x} \in \mathcal{M},\ [V] \in \mathrm{Gr}_k(\mathrm{T}_{\boldsymbol{x}}\mathcal{M}) \right\}$$

CSD can be rewritten as

$$\frac{\mathrm{d}(\boldsymbol{x}, [V])}{\mathrm{d}t}(t) = \boldsymbol{h}(\boldsymbol{x}(t), [V(t)]) \in \mathrm{T}_{(\boldsymbol{x}(t),[V(t)])}\mathrm{Gr}_k(\mathrm{T}\mathcal{M})$$

for some vector field $\boldsymbol{h}$ on $\mathrm{Gr}_k(\mathrm{T}\mathcal{M})$

Linear Stability of the CSD

Theorem 1

Suppose that f is C^3, $(\boldsymbol{x}^\star, [V^\star]) \in \mathrm{Gr}_k(\mathrm{T}\mathcal{M})$, and $\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)$ is nondegenerate.

Then $(\boldsymbol{x}^\star, [V^\star])$ is a linearly steady state of the CSD if and only if

1. $\boldsymbol{x}^\star$ is an index-k saddle point of f on $\mathcal{M}$; and
2. $V^\star$ spans the lowest k-dimensional invariant subspace of $\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)$.

🤩 Weaker assumptions: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \dots, \lambda_n^\star\}$ Yin, Huang, Zhang '22

$$\lambda_1^\star < \lambda_2^\star < \cdots < \lambda_k^\star < 0 < \lambda_{k+1}^\star \leq \cdots \leq \lambda_n^\star \quad \rightsquigarrow \quad \lambda_1^\star \leq \cdots \leq \lambda_k^\star < 0 < \cdots$$

↳ Treat $(\boldsymbol{x}, [V]) \in \mathrm{Gr}_k(\mathrm{T}\mathcal{M})$ rather than $(\boldsymbol{x}, V) \in \mathrm{St}_k(\mathrm{T}\mathcal{M})$

Discretization of the CSD

Forward Euler discretization: with $(\boldsymbol{x}^{(0)}, [V^{(0)}]) \in \mathrm{Gr}_k(\mathrm{T}\mathcal{M})$ and $\eta > 0$

$$\boldsymbol{d}_{\boldsymbol{x}}^{(t)} := -(\mathrm{Id}_{\boldsymbol{x}^{(t)}} - 2V^{(t)}(V^{(t)})^\top)\mathrm{grad}_{\mathcal{M}}\, f(\boldsymbol{x}^{(t)})$$

$$\boldsymbol{d}_{\boldsymbol{v}_i}^{(t)} := -(\mathrm{Id}_{\boldsymbol{x}^{(t)}} - V^{(t)}(V^{(t)})^\top)\mathrm{Hess}_{\mathcal{M}}\, f(\boldsymbol{x}^{(t)})[\boldsymbol{v}_i^{(t)}] + \text{normal part}$$ ❌

$$(\boldsymbol{x}^{(t+1)}, [V^{(t+1)}]) := \mathrm{Retr}^{\mathrm{Gr}_k(\mathrm{T}\mathcal{M})}_{(\boldsymbol{x}^{(t)}, [V^{(t)}])}\left((\eta \boldsymbol{d}_{\boldsymbol{x}}^{(t)}, \eta \boldsymbol{d}_V^{(t)})\right)$$ ✅

A retraction on $\mathrm{Gr}_k(\mathrm{T}\mathcal{M})$: for any $\big((\boldsymbol{x}, [V]), (\boldsymbol{d}_{\boldsymbol{x}}, \boldsymbol{d}_V)\big) \in \mathrm{T}\big(\mathrm{Gr}_k(\mathrm{T}\mathcal{M})\big)$ Edelman, Arias, Smith '98

$$\mathrm{Retr}^{\mathrm{Gr}_k(\mathrm{T}\mathcal{M})}_{(\boldsymbol{x}, [V])}\big((\boldsymbol{d}_{\boldsymbol{x}}, \boldsymbol{d}_V)\big) := \left(\mathrm{Retr}^{\mathcal{M}}_{\boldsymbol{x}}(\boldsymbol{d}_{\boldsymbol{x}}), \mathrm{Proj}_{\mathrm{Retr}^{\mathcal{M}}_{\boldsymbol{x}}(\boldsymbol{d}_{\boldsymbol{x}})}(VU\cos(\Sigma)U^\top + Q\sin(\Sigma)U^\top)\right)$$

$\mathrm{Proj}_{\boldsymbol{x}}(\boldsymbol{d}_V) = Q\Sigma U^\top$, $Q \in \mathrm{St}_k(\mathbb{R}^n)$, $\Sigma := \mathrm{diag}(\sigma_1, \ldots, \sigma_k) \in \mathbb{R}_+^{k\times k}$, $U \in \mathcal{O}(k)$

Local Convergence Analysis of the Discretized CSD

Theorem 2

Suppose that f is C^3, $\boldsymbol{x}^\star \in \mathcal{M}$ is an index-k saddle point of f on $\mathcal{M}$, $\operatorname{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)$ is nondegenerate, and $V^\star \in \mathrm{St}_k(\mathrm{T}_{\boldsymbol{x}^\star}\mathcal{M})$ spans the lowest k-dimensional invariant subspace of $\operatorname{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)$. Let $\{(\boldsymbol{x}^{(t)}, [V^{(t)}])\}$ be the iterate sequence by the discretized CSD.

Then $\{(\boldsymbol{x}^{(t)}, [V^{(t)}])\}$ linearly converges to $(\boldsymbol{x}^\star, [V^\star])$ if

1. the step size η is sufficiently small; and
2. the initial point $(\boldsymbol{x}^{(0)}, [V^{(0)}])$ is sufficiently close to $(\boldsymbol{x}^\star, [V^\star])$.

- The first local convergence results in constrained settings
- Weaker assumptions by working on $\mathrm{Gr}_k(\mathrm{T}\mathcal{M})$: eigenvalues ... +

$$\|V^{(t)}(V^{(t)})^\top - V(\cancel{\boldsymbol{x}}^{(t)})V(\boldsymbol{x}^{(t)})^\top\| \le \alpha < 1, \quad \forall\, t$$

with $V(\boldsymbol{x})$ spanning the exact subspace Luo et al. '22

Local Convergence Rates of the Discretized CSD

Condition numbers: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \ldots, \lambda_n^\star\}$, $\lambda^\star_{\max\,/\,\min} := \max_{1\le i\le n} / \min |\lambda_i^\star|$

- $\boldsymbol{x}$-residual: $\max\left\{\left|1-\eta\lambda^\star_{\min}\right|, \left|1-\eta\lambda^\star_{\max}\right|\right\}$
- V-residual: $\max\left\{\left|1-\eta(\lambda^\star_{k+1}-\lambda^\star_k)\right|, \left|1-\eta(\lambda^\star_n-\lambda^\star_1)\right|\right\}$

$$\kappa_{\boldsymbol{x}} \approx \frac{\lambda^\star_{\max}}{\lambda^\star_{\min}} = \frac{\max\{-\lambda_1^\star, \lambda_n^\star\}}{\min\{-\lambda_k^\star, \lambda_{k+1}^\star\}},$$

$$\kappa_V \approx \frac{\lambda_n^\star - \lambda_1^\star}{\lambda_{k+1}^\star - \lambda_k^\star},$$

Local Convergence Rates of the Discretized CSD

Condition numbers: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \ldots, \lambda_n^\star\}$, $\lambda^\star_{\max / \min} := \max_{1 \le i \le n} / \min |\lambda_i^\star|$

- $\boldsymbol{x}$-residual: $\max\left\{\left|1 - \eta\lambda^\star_{\min}\right|, \left|1 - \eta\lambda^\star_{\max}\right|\right\}$
- V-residual: $\max\left\{\left|1 - \eta(\lambda^\star_{k+1} - \lambda^\star_k)\right|, \left|1 - \eta(\lambda^\star_n - \lambda^\star_1)\right|\right\}$

$$\kappa_{\boldsymbol{x}} \approx \frac{\lambda^\star_{\max}}{\lambda^\star_{\min}} = \frac{\max\{-\lambda_1^\star, \lambda_n^\star\}}{\min\{-\lambda_k^\star, \lambda_{k+1}^\star\}}, \qquad \eta_{\boldsymbol{x}}^{\text{best}} \approx \frac{2}{\min\{-\lambda_k^\star, \lambda_{k+1}^\star\} + \max\{-\lambda_1^\star, \lambda_n^\star\}}$$

$$\kappa_V \approx \frac{\lambda_n^\star - \lambda_1^\star}{\lambda_{k+1}^\star - \lambda_k^\star}, \qquad \eta_V^{\text{best}} \approx \frac{2}{\lambda_n^\star + \lambda_{k+1}^\star - \lambda_k^\star - \lambda_1^\star}$$

Local Convergence Rates of the Discretized CSD

Condition numbers: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \ldots, \lambda_n^\star\}$, $\lambda^\star_{\max\,/\,\min} := \max\,/\,\min_{1\le i\le n} |\lambda_i^\star|$

- $\boldsymbol{x}$-residual: $\max\left\{\left|1-\eta\lambda^\star_{\min}\right|, \left|1-\eta\lambda^\star_{\max}\right|\right\} \rightsquigarrow \kappa_{\boldsymbol{x}} \approx \frac{\max\{-\lambda_1^\star, \lambda_n^\star\}}{\min\{-\lambda_k^\star, \lambda_{k+1}^\star\}}$
- V-residual: $\max\left\{\left|1-\eta(\lambda^\star_{k+1}-\lambda^\star_k)\right|, \left|1-\eta(\lambda^\star_n-\lambda^\star_1)\right|\right\} \rightsquigarrow \kappa_V \approx \frac{\lambda_n^\star-\lambda_1^\star}{\lambda_{k+1}^\star-\lambda_k^\star}$

Example. $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, $f(P) = \mathrm{Tr}(PA)$ on $\mathrm{Gr}_p(\mathbb{R}^d)$ with $p > 1$

- Condition number at the global minimizer: $\frac{\Delta\lambda_{d,1}(A)}{\Delta\lambda_{p+1,p}(A)}$

$$\lambda_{ij}(\mathrm{Hess}_{\mathcal{M}} f(P^\star_{\min})) = \Delta\lambda_{ij}(A) := \lambda_i(A) - \lambda_j(A) > 0, \qquad \begin{array}{l} i = p+1, p+2, \ldots, d \\ j = 1, \ldots, p-1, p \end{array}$$

↪ $\lambda^\star_{\max} = \Delta\lambda_{d,1}(A)$, $\lambda^\star_{\min} = \Delta\lambda_{p+1,p}(A)$ (a.k.a. eigengap)

Local Convergence Rates of the Discretized CSD

Condition numbers: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \ldots, \lambda_n^\star\}$, $\lambda^\star_{\max\,/\,\min} := \max_{1\le i\le n} / \min |\lambda_i^\star|$

- $\boldsymbol{x}$-residual: $\max\left\{\left|1-\eta\lambda^\star_{\min}\right|, \left|1-\eta\lambda^\star_{\max}\right|\right\} \rightsquigarrow \kappa_{\boldsymbol{x}} \approx \frac{\max\{-\lambda_1^\star, \lambda_n^\star\}}{\min\{-\lambda_k^\star, \lambda_{k+1}^\star\}}$
- V-residual: $\max\left\{\left|1-\eta(\lambda^\star_{k+1}-\lambda^\star_k)\right|, \left|1-\eta(\lambda^\star_n-\lambda^\star_1)\right|\right\} \rightsquigarrow \kappa_V \approx \frac{\lambda_n^\star-\lambda_1^\star}{\lambda_{k+1}^\star-\lambda_k^\star}$

Example. $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, $f(P) = \mathrm{Tr}(PA)$ on $\mathrm{Gr}_p(\mathbb{R}^d)$ with $p > 1$

- Condition number at the global minimizer: $\frac{\Delta\lambda_{d,1}(A)}{\Delta\lambda_{p+1,p}(A)}$
- Condition number at the index-1 saddle point: $\frac{\Delta\lambda_{d,1}(A)}{\min\{\Delta\lambda_{p,p-1}(A), \Delta\lambda_{p+1,p}(A), \Delta\lambda_{p+2,p+1}(A)\}}$

$$\lambda_{ij}(\mathrm{Hess}_{\mathcal{M}} f(P^\star_{\mathrm{saddle}})) = \Delta\lambda_{ij}(A), \qquad \begin{array}{l} i = \cancel{p+1}\,p, p+2, \ldots, d \\ j = 1, \ldots, p-1, p+1\,\cancel{p} \end{array}$$

$\hookrightarrow \lambda^\star_{\max} = \Delta\lambda_{d,1}(A)$, $\lambda^\star_{\min} = \min\{\Delta\lambda_{p,p-1}(A), \Delta\lambda_{p+1,p}(A), \Delta\lambda_{p+2,p+1}(A)\}$

Local Convergence Rates of the Discretized CSD

Condition numbers: $\sigma(\mathrm{Hess}_{\mathcal{M}} f(\boldsymbol{x}^\star)) = \{\lambda_1^\star, \dots, \lambda_n^\star\}$, $\lambda^\star_{\max\,/\,\min} := \max\limits_{1\le i\le n} / \min |\lambda_i^\star|$

- $\boldsymbol{x}$-residual: $\max\left\{\left|1-\eta\lambda^\star_{\min}\right|, \left|1-\eta\lambda^\star_{\max}\right|\right\} \rightsquigarrow \kappa_{\boldsymbol{x}} \approx \frac{\max\{-\lambda_1^\star,\lambda_n^\star\}}{\min\{-\lambda_k^\star,\lambda_{k+1}^\star\}}$
- V-residual: $\max\left\{\left|1-\eta(\lambda^\star_{k+1}-\lambda^\star_k)\right|, \left|1-\eta(\lambda^\star_n-\lambda^\star_1)\right|\right\} \rightsquigarrow \kappa_V \approx \frac{\lambda_n^\star-\lambda_1^\star}{\lambda_{k+1}^\star-\lambda_k^\star}$

Example. $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, $f(P) = \mathrm{Tr}(PA)$ on $\mathrm{Gr}_p(\mathbb{R}^d)$ with $p > 1$

- Condition number at the global minimizer: $\frac{\Delta\lambda_{d,1}(A)}{\Delta\lambda_{p+1,p}(A)}$
- Condition number at the index-1 saddle point: $\frac{\Delta\lambda_{d,1}(A)}{\min\{\Delta\lambda_{p,p-1}(A),\Delta\lambda_{p+1,p}(A),\Delta\lambda_{p+2,p+1}(A)\}}$

More nondegeneracies needed & Worse conditioned!

Numerical Results — Linear Eigenvalue Problems

Problem description: For any $A \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$, find the saddle points of

$$f(X) := \frac{1}{2}\mathrm{Tr}(X^\top A X) \quad \text{with} \quad X \in \mathrm{St}_p(\mathbb{R}^n)$$

or

$$\tilde{f}(P) := \frac{1}{2}\mathrm{Tr}(PA) \quad \text{with} \quad P \in \mathrm{Gr}_p(\mathbb{R}^n)$$

Explicit solutions: $A = U\mathrm{diag}(\lambda_1, \dots, \lambda_n)U^\top$, $U \in \mathcal{O}(n)$, for the index-1 saddle points

$$X^\star = [U_{\cdot,1:p-1}, U_{\cdot,p+1}],\ P^\star = X^\star(X^\star)^\top,\ f(X^\star) = \tilde{f}(P^\star) = \frac{1}{2}\Big(\sum_{i=1}^{p-1}\lambda_i + \lambda_{p+1}\Big)$$

Importance of non-redundant parameterization

- $n = 64$, $p = 8$, $\xi = 1.01$, $\eta_{\mathrm{St}} = 2.0$, $\eta_{\mathrm{Gr}} = 4.0$
- Searching for index-1 saddle points

Numerical Results — Linear Eigenvalue Problems

Importance of non-redundant parameterization

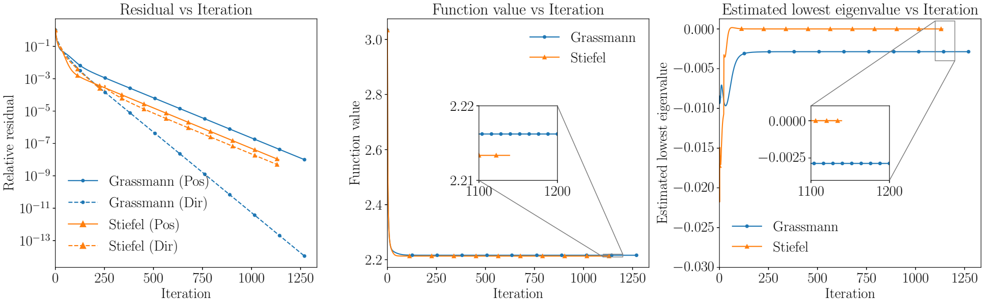

Fig. 11. Convergence curves of algorithms running on the Grassmann and Stiefel manifolds

Numerical Results — Linear Eigenvalue Problems

Importance of non-redundant parameterization

Table 1. Empirical success rates (100 trials) around the global minimizer and index-1 saddle point

Algorithms	Perturbation from global minimizer					Perturbation from saddle point				
	10^{-3}	10^{-2}	10^{-1}	10^{0}	10^{1}	10^{-3}	10^{-2}	10^{-1}	10^{0}	10^{1}
Grassmann	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Stiefel	0%	0%	0%	29%	33%	100%	98%	59%	25%	30%

The Stiefel version fails often due to parameterization redundancy

Influence of problem data on performance

- $n \in \{10, 20, \dots, 80\}$, $p \in \{2, 3, \dots, 8\}$, $\xi \in \{1.0001, 1.001, 1.01, 1.05\}$
- Using the estimated best step sizes $\eta_{\boldsymbol{x}}^{\text{best}}$, η_V^{best}
- Searching for index-1 saddle points

$$\text{Estimated condition number: } \kappa = \frac{\xi^{n-1} - 1}{\xi^{p-2}(\xi - 1)}$$

Numerical Results — Linear Eigenvalue Problems

Influence of problem data on performance

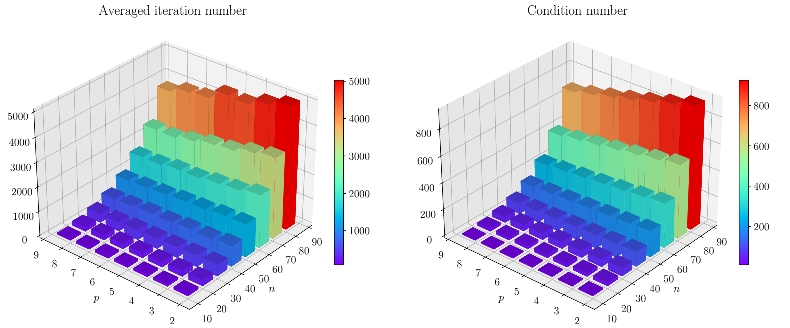

Fig. 12. Average iteration numbers (10 trials) and estimated condition numbers ($\xi = 1.05$)

🤩 Good agreement between iteration numbers and estimated condition numbers

Results on the H_2 molecule

- Restricted HF/6-31G (RHF): $\mathrm{Gr}_1(\mathbb{R}^4)$, dim $= 3$
- Restricted CASSCF (RCAS) (2,2)/6-31G: $\big(\mathbb{S}^3 \times \mathcal{O}(4)\big)/\big(\mathcal{O}(2) \times \mathcal{O}(2)\big)$, dim $= 7$
- FCI (reference): $\mathbb{S}^{15}$, dim $= 15$

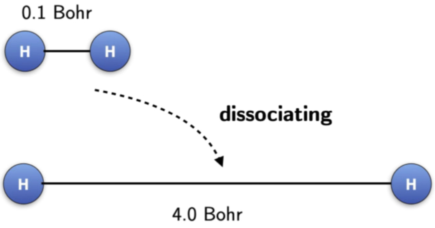

Fig. 13. H_2 molecule with a varying bond length

Numerical Results — Electronic Excited-State Calculations

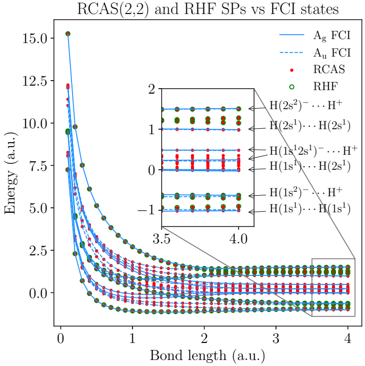

Fig. 14. Found saddle points of RCAS and RHF (1,000 trials) vs FCI states

- 🔍 RHF (dim $= 3$) captures $3 \sim 4$ FCI states
- 🔍 RCAS (dim $= 7$) captures all the FCI states
- 🔍 Higher-index saddle points $\rightarrow$ Higher excitations

Numerical Results — Electronic Excited-State Calculations

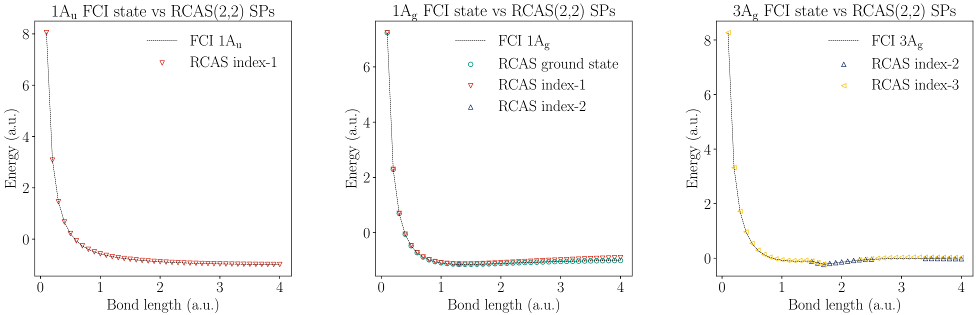

Fig. 15. Projection of RCAS saddle points onto FCI states

- 🔍 Non-one-to-one correspondences: one RCAS $\to$ many FCI; many RCAS $\to$ one FCI
- 🔍 Swaps in RCAS indices with the bond length varying

Take-Home Messages

- Electronics excited states as saddle points of the exact energy functional
- Variational approximate levels of theory on Riemannian manifolds
- **Constrained saddle points of energy functionals in approximate theories**
- Existing works: instability at saddle points, restriction to special manifolds
- A constrained saddle dynamics built on the Grassmann bundle geometry
  - Local theoretical properties under weaker assumptions
  - Non-redundant parameterization, ill condition
  - Applications to standard benchmark molecules

👉 Reference for this talk: arXiv 2601.03931

Acknowledgements

Eric Cancès
(ENPC, IPP)

Laura Grazioli
(ENPC, IPP)

Tony Lelièvre
(ENPC, IPP)

European Research Council
Established by the European Commission

Filippo Lipparini
(UniPi)

Tommaso Nottoli
(UniPi)

Panos Parpas
(Imperial)

EMC^2

Thanks for your attention!

Email: yukuan.hu@enpc.fr

Homepage: https://huyukuan.github.io