From many-body physics to financial markets: sparse modeling for inverse problems

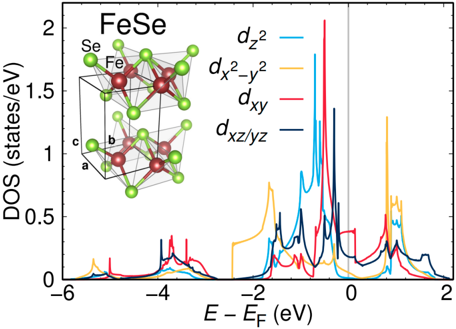
Dr. Daniel Guterding

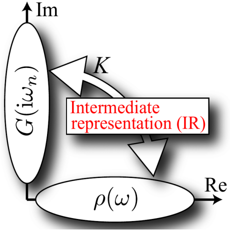
daniel.guterding@gmail.com

December 2nd, 2022

Density of states and Matsubara Green's function

- Electronic density of states describes behavior of material, e.g. spectroscopy
- Start with $\rho(E)$ from effective non-interacting theory
- Matsubara Green's function
  $G(i\omega_n) = \int_{-\infty}^{\infty} dE \, \frac{\rho(E)}{i\omega_n - E}$
  with $\omega_n = (2n+1)\pi/\beta$, $n \in \mathbb{Z}$
- Calculate interacting $G_{\text{int}}(i\omega_n)$ using $G(i\omega_n)$ and Feynman diagrams
- Compare $\rho_{\text{int}}(E)$ to experiments
- How to obtain $\rho_{\text{int}}(E)$ from $G_{\text{int}}(i\omega_n)$?





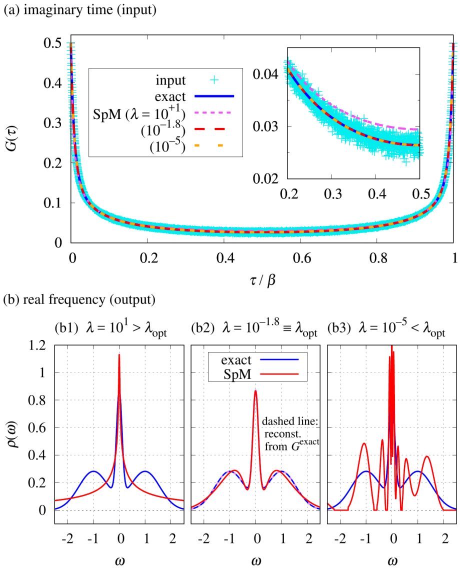
Figures: Comp. Phys. Commun. **231**, 114 (2018), JPSJ **89**, 012001 (2020)

Analytic continuation of continuous-time QMC data

- Imaginary-time Green's function

  $G(\tau) = \int_{-\infty}^{\infty} d\omega\, K(\tau,\omega)\rho(\omega)$

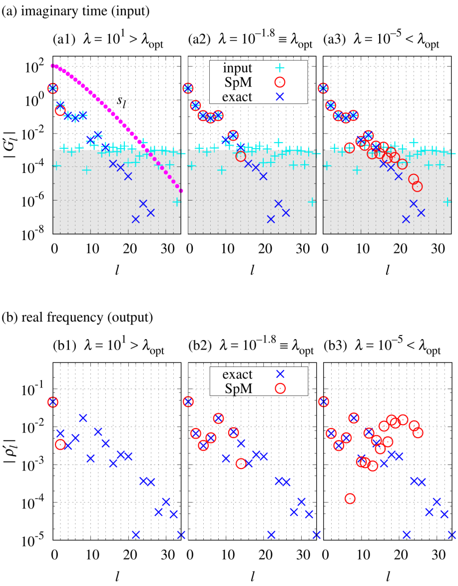
  with $K(\tau,\omega) = \frac{e^{-\tau\omega}}{1+e^{-\beta\omega}}$
- Choose regression basis for $\rho(\omega)$
- Perform optimization with regularization parameter λ
- Overfitting leads to oscillations in spectral function
- Underfitting washes out features of spectral function
- Which regression basis?
- How to regularize in practice?
- Does optimization converge?



Figures: PRE **95**, 061302R (2017)

Discretization, SVD, Optimization

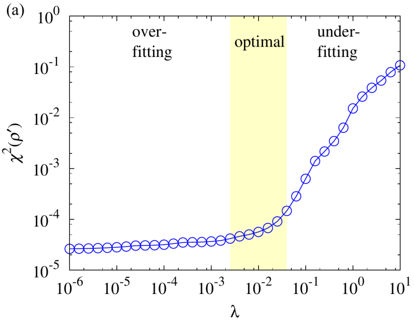
- Imaginary-time Green's function $G(\tau) = \int_{-\infty}^{\infty} d\omega \, K(\tau, \omega) \rho(\omega)$ with $K(\tau, \omega) = \frac{e^{-\tau\omega}}{1+e^{-\beta\omega}}$
- Set energy cutoff $\Lambda \equiv \beta\omega_{max}$, Legendre basis for $\Lambda \to \infty$
- Discretize ω and τ into N vs. M slices
- Perform SVD on kernel matrix $K(\tau_i, \omega_j) = USV^T$
- $G' = U^T G = SV^T \rho = S\rho'$
- Singular values are weights, decay rapidly, apply cutoff
- Minimize squared error $\frac{1}{2}\|G - US\rho'\|_2^2 + \lambda\|\rho'\|_1$

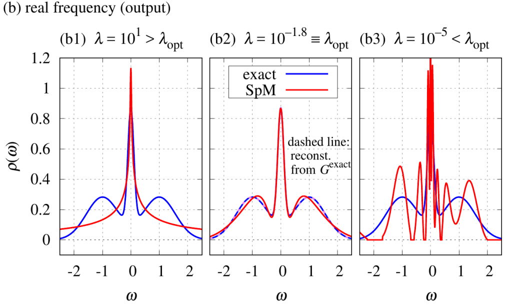


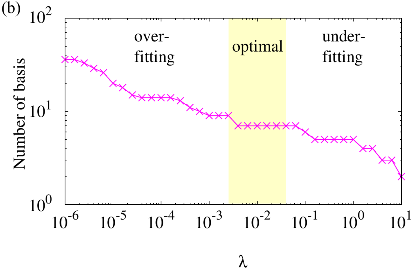
Figures: PRE **95**, 061302R (2017)

Regularization

- Find ρ' that minimizes squared error $\frac{1}{2}\|G - US\rho'\|_2^2 + \lambda\|\rho'\|_1$
- L_1-Regularization via λ (Lasso)
- λ facilitates compromise between accuracy of fit and smoothness of implied density (Elbow method)
- Perfect fit not desirable due to noise



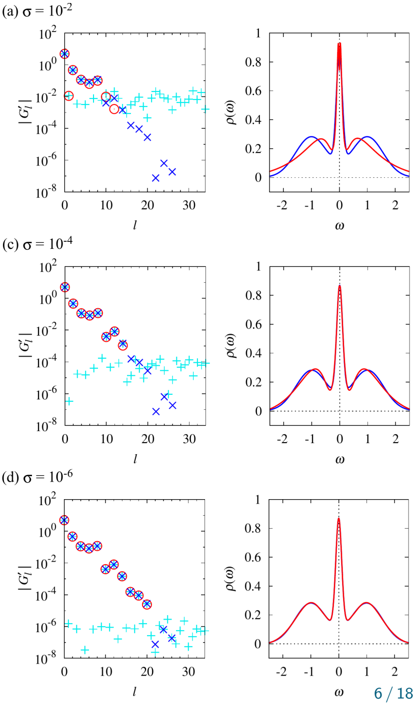




Figures: PRE **95**, 061302R (2017), JPSJ **89**, 012001 (2020)

Preliminary summary

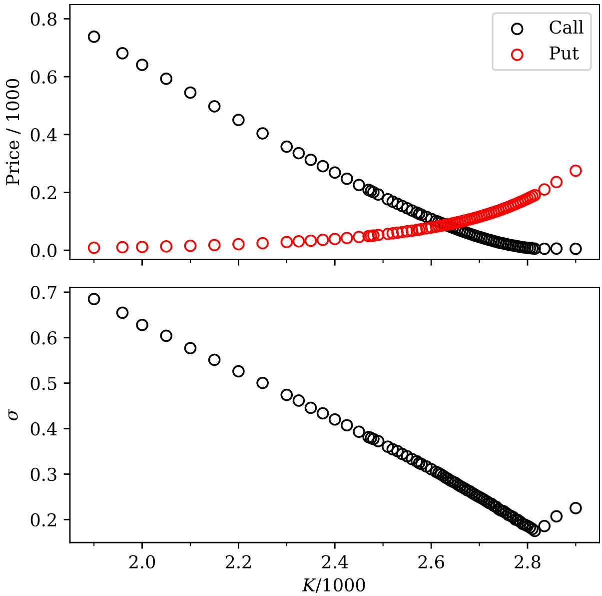
- Simple, fast and reliable method for analytic continuation of noisy data
- Elbow method yields best possible solution given the noise level
- Regularized optimization algorithm widely used in machine learning
- Physical properties of sparse basis well understood (intermediate representation)
- Extensions to two-particle propagators are being studied
- Optimized implementation available, Comp. Phys. Commun. **240**, 181 (2019)



Figures: JPSJ **89**, 012001 (2020)

Option price vs. volatility interpolation and extrapolation

- How to interpolate option prices?
- Interpolation in prices directly?
- Interpolation in implied volatility?
- Which conditions must prices fulfill?
- Which conditions for implied volatility?
- How to extrapolate implied volatility?
- How to get around these issues?

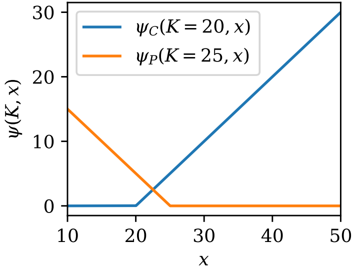


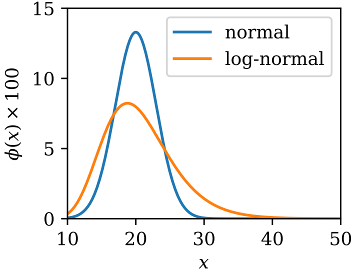
Terminal density and plain-vanilla option price

- Price for option with payoff $\psi(K, x)$ and known terminal density $\phi(x)$:

  $$\Pr(K) = e^{-r\tau} \int_{-\infty}^{\infty} dx\, \psi(K, x)\phi(x)$$

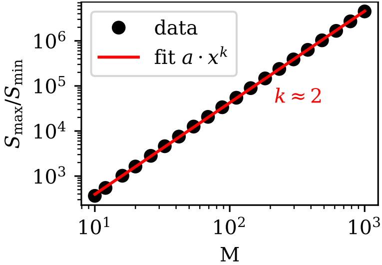
- Call option: $\psi_C(K, x) = \max(0, x - K)$
- Put option: $\psi_P(K, x) = \max(0, K - x)$
- Bachelier model: normal $\phi(x)$
- Black-Scholes model: log-normal $\phi(x)$
- Simple models do not match the market
- How to imply continuous $\phi(x)$ from discrete set of market prices $\Pr(K)$?





Linearization and singular value decomposition

- $\Pr(K) = e^{-r\tau} \int_{-\infty}^{\infty} dx\, \psi(K, x) \phi(x)$
- M known prices
- N equidistant points in x-direction
- Trapezoidal approximation for integral
- $g_i(x) = \Delta x \cdot e^{-r\tau} \psi_{C/P}(K_i, x)$



$$\Pr = \begin{pmatrix} \Pr_1 \\ \vdots \\ \Pr_M \end{pmatrix} = \begin{pmatrix} \frac{1}{2} g_1(x_1) & g_1(x_2) & \dots & g_1(x_{N-1}) & \frac{1}{2} g_1(x_N) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{2} g_M(x_1) & g_M(x_2) & \dots & g_M(x_{N-1}) & \frac{1}{2} g_M(x_N) \end{pmatrix} \begin{pmatrix} \phi(x_1) \\ \vdots \\ \phi(x_N) \end{pmatrix} = G\phi$$

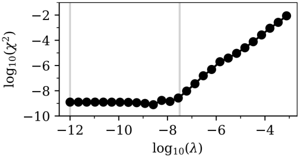
- In general G is $M \times N$ matrix, with $M \leqslant N$
- G is ill-conditioned, cannot invert $\Pr = G\phi$ to get ϕ
- Analyze kernel with singular value decomposition: $G = USV^T$
- Singular values decay quadratically, makes inversion unstable

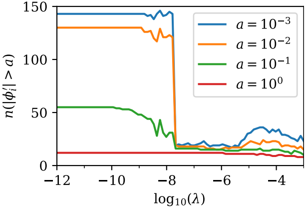
Figures: arXiv:2205.10865

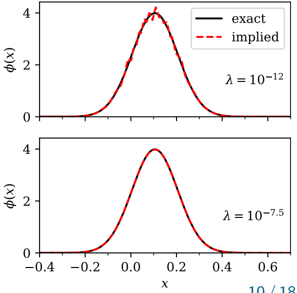
SVD, optimization, regularization

- SVD: $\Pr = G\phi = USV^T\phi$
- $\Pr' = U^T \Pr = SV^T\phi = S\phi'$
- S is diagonal, $\Pr'_i = S_{ii}\phi'_i = s_i\phi'_i$
- Keep Q singular values: $(U, V) \to (\tilde{U}, \tilde{V})$
- $\tilde{U}$ is $M \times Q$ with $Q \leqslant \min(M, N)$,
  approximate $G \approx \tilde{G} = \tilde{U}\tilde{S}\tilde{V}^T$
- L_1-regularized optimization problem
  $F(\phi'|\Pr, \lambda) = \frac{1}{2}\|\Pr - \tilde{U}\tilde{S}\phi'\|_2^2 + \lambda\|\phi'\|_1$
- N discretization points for ϕ, but optimize only $Q \leqslant \min(M, N)$ entries of ϕ'
- additional conditions $\phi_i = \left(\tilde{V}\phi'\right)_i \geqslant 0 \quad \forall i$

$$\text{and } 1 = \left(\frac{1}{2}(\phi_1 + \phi_N) + \sum_{i=2}^{N-1} \phi_i\right)\Delta x$$







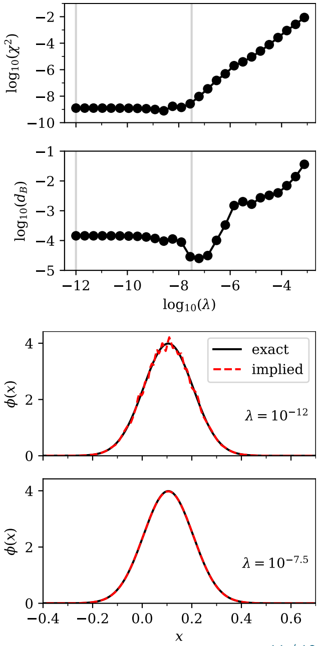
Figures: arXiv:2205.10865

Normal model (Bachelier)

- Use analytical formula for option price
- Density is known analytically
- Get implied density from SVD procedure
- Regularization parameter λ
- Plot RMSE $\chi^2 = \frac{1}{2}\|\mathrm{Pr} - \tilde{U}\tilde{S}\phi'\|_2^2$
- Compare exact and implied density using Bhattacharyya distance

$$d_B(p, q) = -\ln\left[\int_{-\infty}^{\infty} dx\, \sqrt{p(x)q(x)}\right]$$

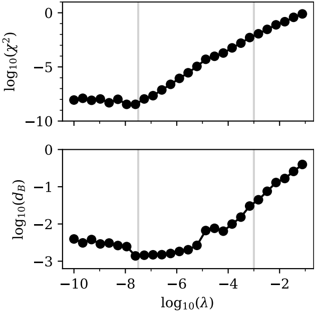
- Plateau in RMSE $\chi^2(\lambda)$
- Too strong regularization increases error
- Regularization makes density smooth
- Elbow method

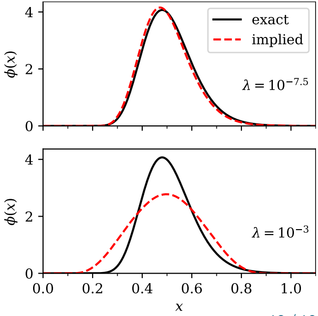


Figures: arXiv:2205.10865

Log-Normal model (Black-Scholes)

- Analytical formula for option price
- Density is known analytically
- Elbow method works as usual
- Accuracy of density reproduction lower than for normal models
- Probably due to `scipy` implementation of log-normal PDF vs. CDF
- Too strong regularization leads to extreme broadening of density

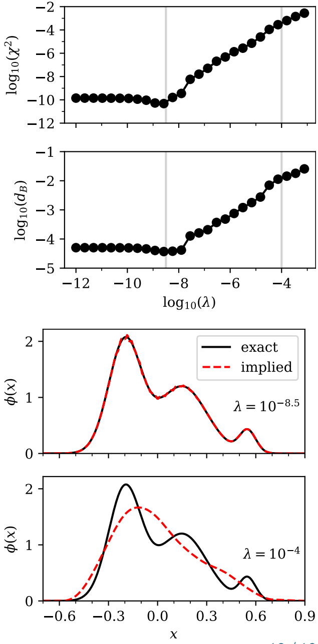




Figures: arXiv:2205.10865

Mixture of normal models

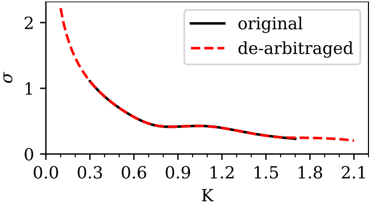
- Linear combination of normal models
- Price is linear combination of prices for component models
- Multimodal models challenging for most other methods
- Reproduced without problems by our SVD method
- Elbow method works in error and Bhattacharyya distance domain
- Too strong regularization removes essential features

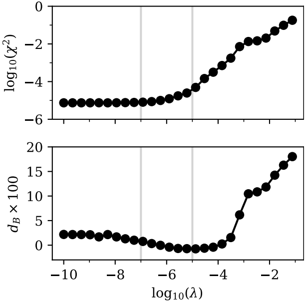


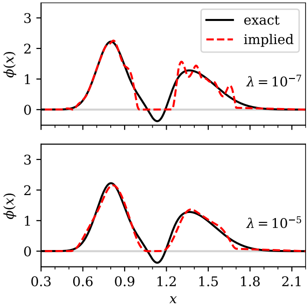
Figures: arXiv:2205.10865

De-arbitraging of a mixture model with arbitrage

- Linear combination of normal models with a negative coefficient
- Analytical formula for option price
- Model 'density' is partly negative
- Bhattacharyya distance is questionable measure, original model not really a density
- Perfect agreement not desirable here
- Obvious de-arbitraging in density domain, but not in implied volatility domain

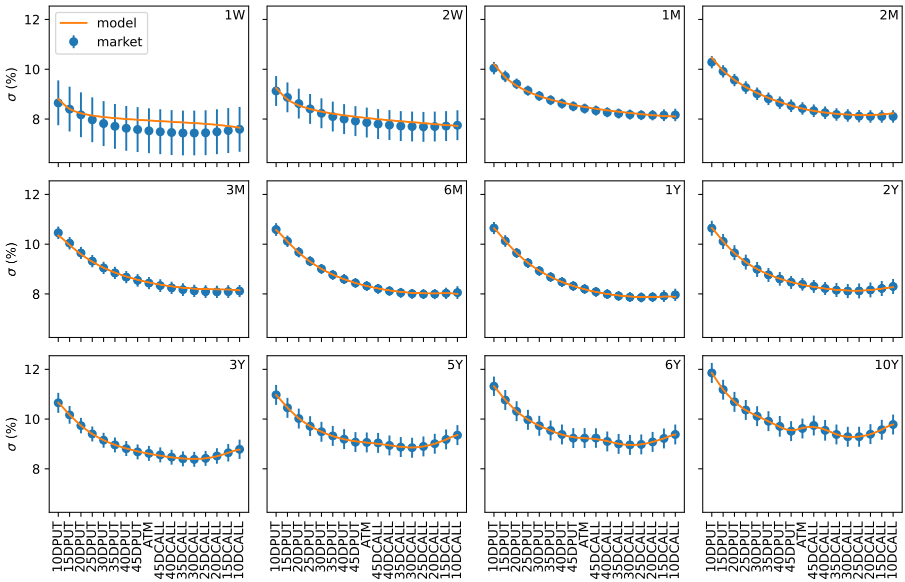




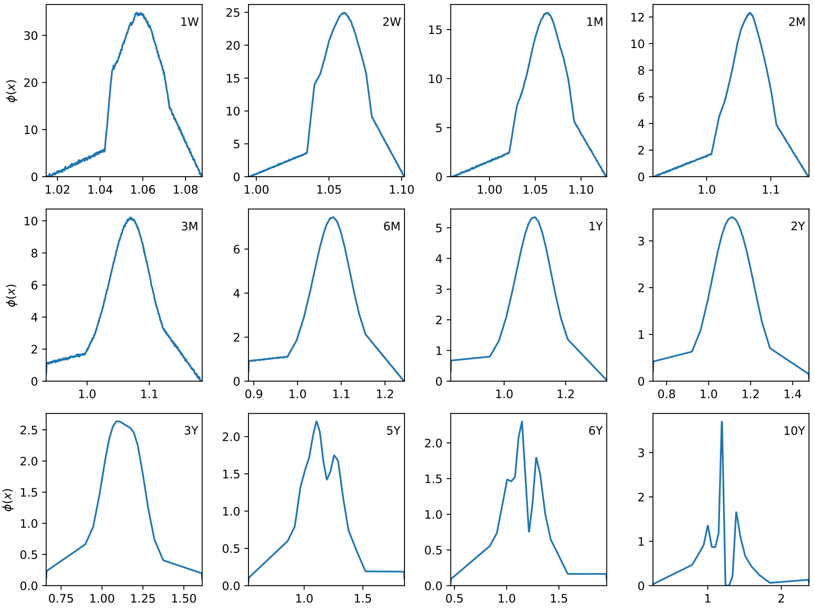


Figures: arXiv:2205.10865

EUR/USD options: Implied Volatility

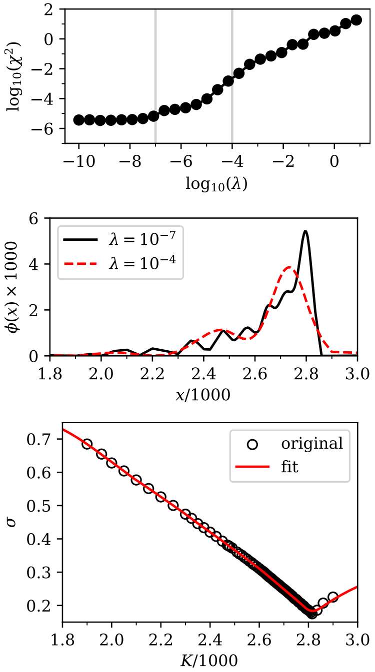


EUR/USD options: Implied Density



S&P 500 index options

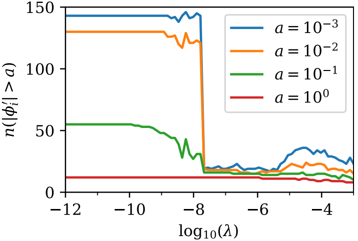
- Perfect reproduction of normal (Bachelier), mixtures of normal and log-normal (Black-Scholes) models
- One-month S&P 500 index options on February 5th, 2018
- Calculate implied future distribution of stock index price
- Automatically correct implied volatility smiles
- Sensible inter- and extrapolation, currently no other general method available
- Performance easily tunable by choice of Q and N, execution time ~ 10 ms

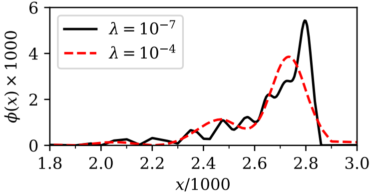


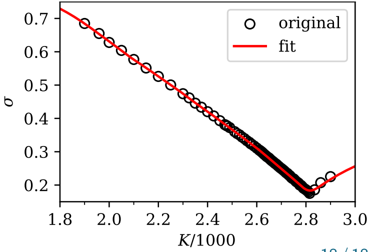
Figures: arXiv:2205.10865

Summary

- Want to find density $\phi(x)$ in integral
  $$\Pr(K) = e^{-r\tau} \int_{-\infty}^{\infty} dx\, \psi(K, x)\phi(x)$$
- Approximate integral using trapezoidal rule, write as matrix equation
- Perform SVD of kernel matrix
- Discard small singular values
- Reformulate as optimization problem in SVD-transformed domain
- Apply L_1-regularization to parameters
- Solution to the problem of volatility smile interpolation and extrapolation
- General recipe for ill-conditioned inverse problems of this form







Figures: arXiv:2205.10865

References

- Guterding & Jeschke, Comp. Phys. Commun. **231**, 114 (2018)
- Otsuki *et al.*, J. Phys. Soc. Japan **89**, 012001 (2020)
- Otsuki *et al.*, Phys. Rev. E **95**, 061302R (2017)
- Chikano *et al.*, Comp. Phys. Commun. **240**, 181 (2019)
- Guterding, arXiv:2205.10865