Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Spawning Semiclassical Wavepackets

O. Rietmann and V. Gradinaru

Research Report No. 2022-49
December 2022

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Spawning Semiclassical Wavepackets

Vasile Gradinaru, Oliver Rietmann

December 31, 2022

Abstract

The semiclassical (or Hagedorn) wavepackets depending on a fixed set of parameters are an orthonormal L^2-basis of generalized coherent states. They have been used to solve numerically the time-dependent Schrödinger equation in its semiclassical formulation, yet their localization property makes them inefficient in case of non-local phenomena such as quantum tunneling. In order to overcome this difficulty, we use simultaneously a few members of several bases with different parameters. We propose an algorithm to expand a given wavefunction in terms of multiple families of Hagedorn wavepackets; each family can then be accurately and efficiently propagated using modern semiclassical time-splittings.

1 Introduction

The semiclassical formulation of the time-dependent Schrödinger equation involves a small parameter (denoted ε^2 below), see e.g. [26, II.2.3], [24]. An L^2-orthonormal basis that is suitable for such models are the parametrized semiclassical (recently called Hagedorn [17, 18]) wavepackets. The numerical methods based upon them combine time-splitting methods with spectral grid free methods in space that work best for localized semiclassical solutions [11, 25, 14, 5, 15, 24]. However, they are very inefficient when non-local phenomena have to be represented [13]. A short outline of the main ingredients of these methods is given in the next section; the idea is that the parameters of such a basis (family) evolve classically (i.e. are governed by some specific ordinary differential equations), while the corrections of the coefficients of the members of the basis expansion are computed via a Galerkin method. The evolution of the coefficients does not influence the evolution of the parameters. Hence, only short-time semiclassical propagation can be accurate, while long-time propagation or non-localized phenomena cannot be reproduced easily [13].
We attempt to overcome these two drawbacks by the following idea. At each time-step we write the solution as a superposition of several well localized families of semiclassical wavepackets. By linearity of the Schrödinger equation, we can use the known very efficient semiclassical propagation algorithms for each of the localized families for a short time and reassemble the solution when needed. The essential part of this procedure is the spawning of the families of Hagedorn wavepackets. We call spawning the creation of a superposition of well localized wavepackets that approximates a given possibly non-localized wavefunctions: a set of different parameters responsible for the different families is proposed. Since each family is a basis, such a representation of the wavefunction is not unique and it is not obvious how to choose the involved families. Spawning Gaussians has already been done in several contexts, e.g. in [7, 23, 4]. Based on the original idea of Bargmann

[3], adaptive phase space lattice propagation methods using standard linear algebra techniques were addressed [6]; since the algorithms are too expensive, modern linear algebra techniques as tensor trains are used [16]. We propose now a different spawning technique. To respect the symmetry between space and momentum, we use phase space representations of wavefunctions. Explicit analytical formulas for the Fourier–Bros–Iagolnitzer (FBI) transform and for the Husimi transform of the Hagedorn wavepackets are known and summarized in Section 3. To achieve a well localized expansion we use so-called localization operators [8]. We introduce them in Section 4. The spawning process is explained in Section 5 and consists of two parts. First, we determine a new set of parameters by fitting a Gaussian mixture to the Husimi transform of the target wavefunction. This is done by a modified expectation maximization algorithm [10] that seeks to minimize the cross-entropy or Kullback-Leibler divergence to the target wavefunction [1]. In a second step, the localization operators, the orthogonality of the members of a family, and the analytical formulas for the FBI transforms of the involved Hagedorn wavepackets are used to compute the coefficients in the new expansions. Section 5 ends with a didactic example that illustrates the spawning process and the fast decay of the coefficients in the expansions. In Section 6, we combine this new spawning technique with modern semiclassical propagation algorithms to solve numerically the time-dependent Schrödinger equation for two benchmark problems. We investigate models that exhibit tunneling, a non-local phenomenon that is notoriously difficult in the semiclassical limit of quantum dynamics: the problem of the Eckart potential [27, 13], which leads to scattering-like tunneling and the doublewell potential, which evolves into delocalized confined solutions.

2 Hagedorn Wavepackets

The Hagedorn wavepackets were developed in [17] and [18]. However, we use the notation of [11]. Let $q, p \in \mathbb{R}^d$ and $Q, P \in \mathbb{C}^{d\times d}$ with

$$Q^T P - P^T Q = 0 \quad \text{and} \quad Q^* P - P^* Q = 2i\,\mathrm{id}, \tag{2.1}$$

where Q^* denotes conjugate-transpose and Q^T denotes transpose without complex conjugation. Condition (2.1) holds if and only if the matrix

$$F := \begin{pmatrix} \mathrm{Re}\,(Q_0) & \mathrm{Im}\,(Q_0) \\ \mathrm{Re}\,(P_0) & \mathrm{Im}\,(P_0) \end{pmatrix}$$

is symplectic. Thus we also call these conditions the *symplecticity* condition. Finally, we abbreviate $\Pi := (q, p, Q, P)$. Using the vector-valued raising and lowering operators

$$\begin{aligned} \mathcal{R}^\varepsilon[\Pi] &:= \frac{i}{\sqrt{2\varepsilon^2}} \left(P^* (x-q) - Q^* \left(-i\varepsilon^2 \nabla_x - p \right) \right) \\ \mathcal{L}^\varepsilon[\Pi] &:= -\frac{i}{\sqrt{2\varepsilon^2}} \left(P^T (x-q) - Q^T \left(-i\varepsilon^2 \nabla_x - p \right) \right) \end{aligned} \tag{2.2}$$

we define for $k \in \mathbb{N}_0^d$ the Hagedorn wavepackets $\varphi_k^\varepsilon[\Pi]$ recursively by

$$\varphi_0^\varepsilon[\Pi](x) := \left(\pi\varepsilon^2\right)^{-\frac{d}{4}} (\det Q)^{-\frac{1}{2}} \exp\left(\tfrac{i}{2\varepsilon^2} (x-q)^T P Q^{-1} (x-q) + \tfrac{i}{\varepsilon^2} p^T (x-q) \right)$$

and

$$\varphi_{k+e_j}^\varepsilon[\Pi] := \tfrac{1}{\sqrt{k_j+1}} \mathcal{R}_j^\varepsilon[\Pi]\, \varphi_k^\varepsilon[\Pi] \tag{2.3}$$

for all $k \in \mathbb{N}_0^d$ and all $j \in \{1, \ldots, d\}$. We denote e_j the j-th member of the canonical basis in $\mathbb{R}^d$. Then, the three-term recurrence [26, (2.7)]

$$Q\left(\sqrt{k_j+1}\varphi_{k+e_j}^{\varepsilon}(x)\right)_{j=1}^{d} = \sqrt{\tfrac{2}{\varepsilon^2}}(x-q)\,\varphi_k^{\varepsilon}(x) - \overline{Q}\left(\sqrt{k_j}\varphi_{k-e_j}^{\varepsilon}(x)\right)_{j=1}^{d} \tag{2.4}$$

holds (the terms with negative entries in the multi-indices are treated as zero). We see here the difference to the simple tensor product of one-dimensional Hermite functions: the matrix Q mixess the coordinates in a non-trivial way. The obtained functions form an orthonormal basis of $L^2\left(\mathbb{R}^d\right)$, see [18]. All members of the basis share the same Π, so we call such a basis a family of Hagedorn wavepackets. The parameters q and p are respectively the means in the physical and in the frequency space, while the he matrices Q and P are related to the localization of the wavepackets in these spaces [18]. A wavefunction $\psi \in L^2\left(\mathbb{R}^d\right)$ can thus be approximated by the linear combination

$$\psi(x) = \sum_{k\in\mathcal{K}} c_k \varphi_k^{\varepsilon}[\Pi](x), \tag{2.5}$$

where $\mathcal{K} \subset \mathbb{N}_0$ is a multi-index set and $c_k \in \mathbb{C}$ for all $k \in \mathcal{K}$. A common choice for $\mathcal{K}$ is [11, Sec.5.4]

$$\mathcal{K} = \left\{ k \in \mathbb{N}_0^d \,|\, \prod_{j=1}^{d}(1+k_j) \leq K \right\} \tag{2.6}$$

with some truncation constant $K \in \mathbb{N}$. As in the case of Hermite functions in one dimension, we are faced with a spectral method for the space approximation that does not need any artificial boundary conditions and that is in principle grid free. Constructed with the application to semiclassical quantum-dynamics in mind, the approximation is very efficient for localized wavefunctions, but it is very expensive for non-localized ones.

3 Phase Space Representations

Throughout this chapter, we fix parameters $\Pi_0 = (q_0, p_0, Q_0, P_0)$ such that Q_0 and P_0 satisfy the symplecticity condition (2.1). Our goal is to find analytic expressions for two phase space representations of the Hagedorn functions $\varphi_k^{\varepsilon}[\Pi_0]$ that can be evaluated efficiently.

3.1 FBI Transform

The FBI transform of a wavefunction yields a complex valued phase space representation. It was developed in [21] and the abbreviation stands for Fourier–Bros–Iagolnitzer [12, Ch. 3.3]. There are different conventions for the transform and we will use the one from [22]. Let $q, p \in \mathbb{R}^d$ and consider a multi-index $k \in \mathbb{N}^d$. Let us shorten the notation

$$\varphi_k^{\varepsilon}[q,p] := \varphi_0^{\varepsilon}[q,p,\mathrm{Id},i\mathrm{Id}]$$

for a Hagedorn function with standard parameters $Q = \mathrm{Id}$ and $P = i\mathrm{Id}$. Moreover, let

$$\begin{aligned} g[q,p](q,p) &:= e^{\frac{i}{2\varepsilon^2}p^T q}\varphi_0^{\varepsilon}[q,p](x) \\ &= \left(\pi\varepsilon^2\right)^{-\frac{d}{4}} \exp\left(-\tfrac{1}{2\varepsilon^2}\|x-q\|^2 + \tfrac{i}{\varepsilon^2}p^T\left(x-\tfrac{q}{2}\right)\right). \end{aligned} \tag{3.1}$$

In this notation, the FBI transform $\mathcal{B}^\varepsilon[\psi] : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ of a wavefunction $\psi \in L^2\left(\mathbb{R}^d\right)$ is defined by the L^2 inner product

$$\begin{aligned}\mathcal{B}^\varepsilon[\psi](q,p) &:= (2\pi\varepsilon)^{-\frac{d}{2}} \langle g[q,p], \psi\rangle \\ &= (2\pi\varepsilon)^{-\frac{d}{2}} e^{-\frac{i}{2\varepsilon^2}p^T q} \langle \varphi_0^\varepsilon[q,p], \psi\rangle.\end{aligned}$$

Note the different convention in [25], where the authors omit the factor $e^{\frac{i}{2\varepsilon^2}p^T q}$. The FBI transform is a complex valued phase space representation of ψ. It is worth noting that for all $\psi_1, \psi_2 \in L^2(\mathbb{R})$, we have

$$\langle \psi_1, \psi_2\rangle_{L^2(\mathbb{R}^d)} = \langle \mathcal{B}^\varepsilon[\psi_1], \mathcal{B}^\varepsilon[\psi_2]\rangle_{L^2(\mathbb{R}^{2d})},$$

which follows from the well-known resolution of identity property of coherent states

$$\mathrm{Id}_{L^2(\mathbb{R}^d)} = (2\pi\varepsilon)^{-d} \int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi_0^\varepsilon[q,p] \langle \varphi_0^\varepsilon[q,p], \cdot\rangle \,\mathrm{d}q\,\mathrm{d}p. \tag{3.2}$$

This will be important because it means that the FBI transform $\mathcal{B}^\varepsilon$ preserves the angles between and hence the orthonormality of the members of a single family of semiclassical wavepackets. The FBI transform of the Hagedorn function $\varphi_k^\varepsilon[\Pi_0]$ can be computed analytically [25, Prop 5], where $\Pi_0 = (q_0, p_0, Q_0, P_0)$ is sucht that Q_0 and P_0 satisfy symplecticity condition (2.1). We only state the result for $k = 0$, namely

$$\mathcal{B}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]](q,p) = \left(\pi\varepsilon^2\right)^{-\frac{d}{2}} \left(\det(Z_0)\right)^{-\frac{1}{2}} e^{-\frac{1}{2\varepsilon^2}(q-q_0)^2 + \frac{i}{\varepsilon^2}p^T\left(\frac{q}{2}-q_0\right)} e^{\frac{1}{2\varepsilon^2}z_0^T(q,p)Q_0Z_0^{-1}z_0(q,p)},$$

where

$$z_0(q,p) := (q-q_0) - i(p-p_0) \qquad \text{and} \qquad Z_0 := Q_0 - iP_0. \tag{3.3}$$

In the special case of $Q = \mathrm{Id}$ and $P = i\,\mathrm{Id}$ we have

$$\mathcal{B}^\varepsilon[\varphi_0^\varepsilon[q_0,p_0]](q,p) = \left(2\pi\varepsilon^2\right)^{-\frac{d}{2}} e^{-\frac{i}{2\varepsilon^2}p_0^T q_0} e^{-\frac{1}{4\varepsilon^2}|z_0(q,p)|^2}.$$

In contrast to [25, Prop 5], the multiplicative complex phase does not depend on (q,p) and this is the reason why we use this convention for the FBI transform. Getting rid of these oscillations makes numerical evaluation more stable and the phase space plots readable. In order to evaluate $\mathcal{B}^\varepsilon[\varphi_k^\varepsilon[\Pi_0]]$ for $k \neq 0$, we use the following three-term recurrence.

Proposition 3.1. Let $\varepsilon > 0$ and $d \in \mathbb{N}$. Moreover, let $\Pi_0 = (q_0, p_0, Q_0, P_0)$ be parameters where Q_0 and P_0 satisfy the symplecticity condition (2.1). The FBI transforms

$$B_k := \begin{cases} \mathcal{B}^\varepsilon[\varphi_k^\varepsilon[\Pi_0]] & \text{, if } k_1,\dots,k_d \geq 0 \\ 0 & \text{, else} \end{cases}$$

of the Hagedorn functions satisfy the three-term recurrence

$$(Q_0 - iP_0)\begin{pmatrix} \sqrt{k_1+1}\, B_{k+e_1}(q,p) \\ \vdots \\ \sqrt{k_d+1}\, B_{k+e_d}(q,p)\end{pmatrix} = \sqrt{\tfrac{2}{\varepsilon^2}}\, z_0(q,p)\, B_k(q,p) - \left(\overline{Q_0} - i\overline{P_0}\right)\begin{pmatrix} \sqrt{k_1}\, B_{k-e_1}(q,p) \\ \vdots \\ \sqrt{k_d}\, B_{k-e_d}(q,p)\end{pmatrix}$$

for all $k \in \mathbb{N}_0^d$ and all $q,p \in \mathbb{R}^d$, where $z_0(q,p) := (q-q_0) - i(p-p_0)$.

Proof. Consider the lowering operator, see (2.2):

$$\mathcal{L}^{\varepsilon}\left[\Pi_0\right] = -\tfrac{i}{\sqrt{2\varepsilon^2}}\left(P_0^T\left(x-q_0\right) - Q_0^T\left(-i\varepsilon^2\nabla_x - p_0\right)\right),$$

which is the adjoint of the raising operator and hence

$$\sqrt{k_j+1}\, B^{\varepsilon}_{k+e_j}\left(q,p\right) = \langle \mathcal{L}^{\varepsilon}_j\left[\Pi_0\right] g^{\varepsilon}\left[q,p\right], \varphi^{\varepsilon}_k\left[\Pi_0\right]\rangle, \tag{3.4}$$

where $g^{\varepsilon}\left[q,p\right]$ was defined in (3.1). We compute

$$\begin{aligned}\mathcal{L}^{\varepsilon}\left[\Pi_0\right] g^{\varepsilon}\left[q,p\right] &= \tfrac{1}{\sqrt{2\varepsilon^2}}\left(-iP_0^T\left(x-q_0\right) + iQ_0^T\left(p-p_0\right) - Q_0^T\left(x-q\right)\right) g^{\varepsilon}\left[q,p\right] \\ &= \tfrac{1}{\sqrt{2\varepsilon^2}}\left(Q_0^T\left(\left(q-q_0\right) + i\left(p-p_0\right)\right) - \left(Q_0+iP_0\right)^T\left(x-q_0\right)\right) g^{\varepsilon}\left[q,p\right].\end{aligned}$$

In the notation (3.3) and with $\tilde{Z}_0 := Q_0 + iP_0$, this can be rewritten as

$$\mathcal{L}^{\varepsilon}\left[\Pi_0\right] g^{\varepsilon}\left[q,p\right] = \tfrac{1}{\sqrt{2\varepsilon^2}} Q_0^T\, \overline{z_0\left(q,p\right)}\, g^{\varepsilon}\left[q,p\right] - \tfrac{1}{\sqrt{2\varepsilon^2}} \tilde{Z}_0^T\left(x-q_0\right) g^{\varepsilon}\left[q,p\right].$$

We substitute this into (3.4) and get

$$\begin{aligned}&\begin{pmatrix}\sqrt{k_1+1}\, B_{k+e_1}\left(q,p\right)\\ \vdots \\ \sqrt{k_d+1}\, B_{k+e_d}\left(q,p\right)\end{pmatrix} \\ &\qquad = \tfrac{1}{\sqrt{2\varepsilon^2}} Q_0^*\, z_0\left(q,p\right) B_k\left(q,p\right) - \tfrac{1}{\sqrt{2\varepsilon^2}} \tilde{Z}_0^* \begin{pmatrix}\langle g^{\varepsilon}\left[q,p\right], \left(x-q_0\right)_1 \varphi^{\varepsilon}_k\left[\Pi_0\right]\rangle \\ \vdots \\ \langle g^{\varepsilon}\left[q,p\right], \left(x-q_0\right)_d \varphi^{\varepsilon}_k\left[\Pi_0\right]\rangle\end{pmatrix}.\end{aligned}$$

We replace the term $\left(x-q_0\right)$ using the three-term recurrence for Hagedorn wavepackets (2.4)

$$\left(x-q_0\right)\, \varphi^{\varepsilon}_k\left[\Pi_0\right] = \sqrt{\tfrac{\varepsilon^2}{2}}\, Q_0 \begin{pmatrix}\sqrt{k_1+1}\, \varphi^{\varepsilon}_{k+e_1}\left[\Pi_0\right] \\ \vdots \\ \sqrt{k_d+1}\, \varphi^{\varepsilon}_{k+e_d}\left[\Pi_0\right]\end{pmatrix} + \sqrt{\tfrac{\varepsilon^2}{2}}\, \overline{Q_0} \begin{pmatrix}\sqrt{k_1}\, \varphi^{\varepsilon}_{k-e_1}\left[\Pi_0\right] \\ \vdots \\ \sqrt{k_d}\, \varphi^{\varepsilon}_{k-e_d}\left[\Pi_0\right]\end{pmatrix}.$$

This yields the three-term recurrence

$$\begin{aligned}&\left(\mathrm{Id} + \tfrac{1}{2}\tilde{Z}_0^* Q_0\right) \begin{pmatrix}\sqrt{k_1+1}\, B_{k+e_1}\left(q,p\right)\\ \vdots \\ \sqrt{k_d+1}\, B_{k+e_d}\left(q,p\right)\end{pmatrix} \\ &\qquad = \tfrac{1}{\sqrt{2\varepsilon^2}} Q_0^*\, z_0\left(q,p\right) B_k\left(q,p\right) - \tfrac{1}{2}\tilde{Z}_0^* \overline{Q_0} \begin{pmatrix}\sqrt{k_1}\, B_{k-e_1}\left(q,p\right)\\ \vdots \\ \sqrt{k_d}\, B_{k-e_d}\left(q,p\right)\end{pmatrix}\end{aligned}$$

that can be simplified using the symplecticity property (2.1)

$$Q_0^T P_0 - P_0^T Q_0 = 0 \quad \text{and} \quad Q_0^* P_0 - P_0^* Q_0 = 2i\,\mathrm{Id}.$$

The first equation implies

$$\tilde{Z}_0^*\, \overline{Q_0} = Q_0^*\, \left(\overline{Q_0} - i\overline{P_0}\right)$$

and the second equation implies

$$\mathrm{Id} + \tfrac{1}{2}\tilde{Z}_0^*\, Q_0 = \tfrac{1}{2} Q_0^*\, Z_0.$$

Because $Q_0 Q_0^*$ is real and positive definite [26, Ch. 5,Lem. 1.1], Q_0^* is invertible. This concludes the proof. □

Remark 3.1. The matrix $Z_0 = Q_0 - iP_0$ is invertible and well-conditioned since [25, Page 19]

$$(Q_0 - iP_0)^* (Q_0 - iP_0) = (Q_0 + iP_0)^* (Q_0 + iP_0) + 4\,\mathrm{Id}.$$

3.2 Husimi Transform

The Husimi transform was first introduced in [20]. Let $\psi \in L^2\left(\mathbb{R}^d\right)$ be a wavefunction. The Husimi transform of ψ is the function on the phase space $\mathcal{H}^\varepsilon[\psi] : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ defined as the absolute value squared of the FBI transform:

$$\mathcal{H}^\varepsilon[\psi](q,p) := |\mathcal{B}^\varepsilon[\psi](q,p)|^2.$$

An equivalent definition can be obtained from the Wigner transform [25, Page 3]

$$\mathcal{W}^\varepsilon[\psi](q,p) := \left(2\pi\varepsilon^2\right)^{-d} \int_{\mathbb{R}^d} \overline{\psi\left(q + \tfrac{x}{2}\right)} \psi\left(q - \tfrac{x}{2}\right) e^{ix^T p/\varepsilon^2}\,\mathrm{d}x.$$

The Husimi transform is just a mollified Wigner transform [25, Page 3]

$$\mathcal{H}^\varepsilon[\psi] = \mathcal{G} * W^\varepsilon[\psi], \qquad \mathcal{G}(q,p) = \left(\pi\varepsilon^2\right)^{-d} \exp\left(-\tfrac{1}{\varepsilon^2}\left(q^2+p^2\right)\right),$$

where $*$ denotes convolution. It follows from the resolution of identity property (3.2) that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathcal{H}^\varepsilon[\psi](q,p)\,\mathrm{d}q\,\mathrm{d}p = \|\psi\|^2_{L^2(\mathbb{R}^d)}.$$

The Husimi transform is thus a quasi probability density if $\|\psi\|_{L^2} = 1$. It is not an honest probability density because the marginal densities w.r.t. q and p are not equal to $|\widehat{\psi}|^2$ and $|\psi|^2$, respectively. The Husimi transform of a Hagedorn function $\varphi_k^\varepsilon[\Pi_0]$ can be computed analytically, see [25, Prop 5]. For subsequent use, we only need the Husimi transform of $\varphi_0^\varepsilon[\Pi_0]$. However, we require a different expression than the one in [25], see Remark 5.2.

Proposition 3.2. Let $\varepsilon > 0$ and $d \in \mathbb{N}$. Moreover, let $\Pi_0 = (q_0, p_0, Q_0, P_0)$ be parameters such that the $2d \times 2d$ matrix

$$F_0 := \begin{pmatrix} \mathrm{Re}\,(Q_0) & \mathrm{Im}\,(Q_0) \\ \mathrm{Re}\,(P_0) & \mathrm{Im}\,(P_0) \end{pmatrix}$$

is symplectic. The Husimi transform of $\varphi_0^\varepsilon[\Pi_0]$ is then given by

$$\mathcal{H}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]](q,p) = (2\pi)^{-\frac{d}{2}} \sqrt{\det(\Sigma)} \exp\left(-\frac{1}{2}\left\langle \begin{pmatrix} q-q_0 \\ p-p_0 \end{pmatrix}, \Sigma^{-1} \begin{pmatrix} q-q_0 \\ p-p_0 \end{pmatrix} \right\rangle\right),$$

with covariance matrix

$$\Sigma := \tfrac{\varepsilon^2}{2}\left(F_0 F_0^T + \mathrm{Id}\right).$$

Proof. We use the fact, that the Husimi transform is just a mollified Wigner transform

$$\mathcal{H}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]] = \mathcal{G}^\varepsilon * \mathcal{W}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]], \qquad \mathcal{G}^\varepsilon(q,p) = \left(\pi\varepsilon^2\right)^{-d} \exp\left(-\tfrac{1}{\varepsilon^2}\left(q^2+p^2\right)\right).$$

The Wigner transform $\mathcal{W}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]]$ of φ_0^ε is given by [25, Thm 1]

$$\mathcal{W}^\varepsilon[\varphi_0^\varepsilon[\Pi_0]](q,p) = \left(\pi\varepsilon^2\right)^{-d} \exp\left(-\frac{1}{\varepsilon^2}\left\| F_0^{-1} \begin{pmatrix} q-q_0 \\ p-p_0 \end{pmatrix} \right\|^2\right).$$

This is a Gaussian with mean (q_0, p_0) and covariance matrix $\frac{\varepsilon^2}{2} F_0 F_0^T$. The covariance matrix of $\mathcal{G}^\varepsilon$ is $\frac{\varepsilon^2}{2}\mathrm{Id}$. The convolution is thus a Gaussian with mean (q_0, p_0) and covariance matrix $\Sigma = \frac{\varepsilon^2}{2} F_0 F_0^T + \frac{\varepsilon^2}{2}\mathrm{Id}$. □

4 Localization Operators

The localization operators have been introduced in [8] and are closely related to the FBI transform. Let $F : \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$ be a function on the phase space. The localization operator with symbol F is defined as

$$P_F := \left(2\pi\varepsilon^2\right)^{-d} \int_{\mathbb{R}^d \times \mathbb{R}^d} F(q,p)\ \varphi_0^\varepsilon[q,p]\ \langle \varphi_0^\varepsilon[q,p],\ \cdot\ \rangle_{L^2}\ \mathrm{d}q\,\mathrm{d}p,$$

where $\varphi_0^\varepsilon[q,p] := \varphi_0^\varepsilon[q,p,\mathrm{Id},i\mathrm{Id}]$. It has the following properties [8].

(i) If $F \equiv 1$, the resolution of identity (3.2) implies $P_F = \mathrm{Id}_{L^2}$.

(ii) Given a partition of unity $F_1 + \ldots + F_n \equiv 1$ on the phase space, we obtain a partition of unity on $L^2\left(\mathbb{R}^d\right)$ by $P_{F_1} + \ldots + P_{F_n} = \mathrm{Id}_{L^2}$.

(iii) If $F = \mathrm{Id}_\Omega$ is the characteristic function of a subset $\Omega \subseteq \mathbb{R}^2$ of the phase space, then $P_F\psi$ decays quickly outside of Ω, i.e. is localized around Ω.

The Hagedorn functions $\varphi_k^\varepsilon[q_0,p_0]$ happen to be the eigenfunctions of P_F whenever F is spherically symmetric around (q_0,p_0). Moreover, if F is a spherically symmetric Gaussian with mean (q_0,p_0), then the eigenvalues decay exponentially [8].

Example 4.1. Let $d = 1$ for simplicity. Consider the phase space Gaussian

$$F(q,p) = \exp\left(-\frac{(q-q_0)^2 + (p-p_0)^2}{2\varepsilon^2\sigma^2}\right) \tag{4.1}$$

with mean $(q_0,p_0) \in \mathbb{R}^2$ and $\sigma^2 > 0$. The Galerkin matrix of P_F with respect to the Hagedorn basis $\varphi_k^\varepsilon[q_0,p_0]$ can be computed explicitly, namely (see [8] for the derivation)

$$M_{kj} := \langle \varphi_k^\varepsilon[q_0,p_0], P_F\,\varphi_j^\varepsilon[q_0,p_0]\rangle_{L^2} = \left(1+\tfrac{1}{\sigma^2}\right)^{-(k+1)}\ \delta_{kj}, \tag{4.2}$$

where δ_{kj} denotes the Kronecker delta, i.e. M is diagonal with exponentially decaying entries and

$$P_F \sum_{k=0}^{\infty} c_k \varphi_k^\varepsilon[q_0,p_0] = \sum_{k=0}^{\infty} \left(1+\tfrac{1}{\sigma^2}\right)^{-(k+1)} c_k \varphi_k^\varepsilon[q_0,p_0], \qquad c_k \in \mathbb{C}.$$

This encodes the localization property of P_F. Recall that the state $\varphi_k^\varepsilon[q_0,p_0]$ for $k = 0$ minimizes the Heisenberg uncertainty, i.e. wavefunctions cannot be more localized in phase space. The larger we choose k, the less localized the state becomes. Then P_F suppresses the higher excited states (large k, less localized) exponentially in k.

Example 4.1 can be levaraged to an approximate partition of unity. Consider an equally spaced grid of Gaussians like (4.1) in phase space. If the grid spacing is small enough, the pointwise sum of Gaussians will be approximately constant [2] and yield a partition of unity after multiplication by a suitable factor, see Figure 1 for an example in one dimension (instead of $2d$). By Item (ii) above, this gives rise to an (approximate) partition of unity on $L^2\left(\mathbb{R}^d\right)$. Suppose we want to represent a given wavefunction $\psi \in L^2(\mathbb{R})$ by a sum of members of different families of Hagedorn wavepackets

$$\psi = \sum_{i,j,k} c_k^{i,j} \varphi_k^\varepsilon[q_i,p_j], \tag{4.3}$$

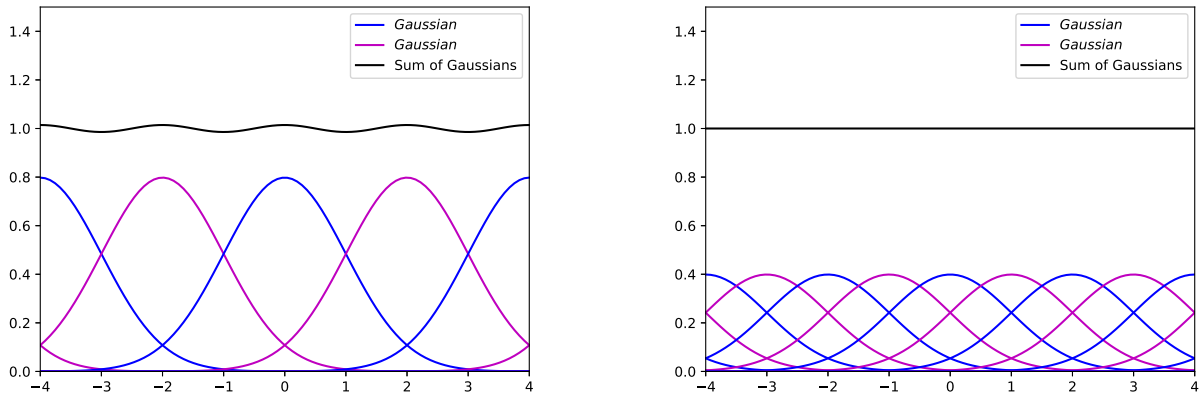

Figure 1: Gaussian approximate partition of unity using equally spaced normal densities of unit variance. The smaller the spacing, the better the approximation of the constant one function. See [2] for details.

where $(q_i, p_j) \in \mathbb{R}^2$ lie on an equally spaced grid in phase space with sufficiently small spacing, so that approximately $\sum_{i,j} F_{i,j} \equiv 1$, where

$$F_{i,j}(q,p) := \alpha\, F(q - q_i, p - p_j)$$

for a suitable normalization factor α depending only on ε and σ^2. By Item (ii), we then have

$$\psi \approx \sum_{i,j} P_{i,j}\psi$$

and we can project every $P_{i,j}\psi$ onto a Hagedorn family to get the coefficients in (4.3)

$$c_k^{i,j} = \langle \varphi_k^\varepsilon [q_i, p_j], P_{F_{i,j}}\psi \rangle_{L^2(\mathbb{R}^d)} = \langle \mathcal{B}^\varepsilon [\varphi_k^\varepsilon [q_i, p_j]], F_{i,j}\, \mathcal{B}^\varepsilon [\psi] \rangle_{L^2(\mathbb{R}^{2d})}.$$

This is closely related to an idea from [23], where the authors also use a Gaussian approximate partition of unity in order to expand with respect to an equally spaced grid of coherent states. Our work differs in two essential ways.

1. We use the higher excited states $\varphi_k^\varepsilon [q,p]$, while [23] only considers $k = 0$ in (4.3).
2. We do *not* use a grid, but we choose (automatically) just a few suitable points (q_m, p_m) in the phase space.

The details on how to obtain the (q_m, p_m) for a given ψ are explained in the next section. This also requires a different symbol F_m associated with (q_m, p_m). We will choose the symbol

$$F_m(q,p) = \frac{r_m H_0^\varepsilon (q - q_m, p - p_m)}{\sum_\ell r_\ell H_0^\varepsilon (q - q_\ell, p - p_\ell)},$$

where $r_1, r_2, \ldots$ are positive and sum to one, and

$$H_0^\varepsilon (q,p) := \mathcal{H}^\varepsilon [\varphi_0^\varepsilon [0,0]] (q,p) = \left(2\pi\varepsilon^2\right)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2\varepsilon^2}\left(q^2 + p^2\right)\right)$$

is the Husimi function of the Hermite function of order zero. Observe that $\sum_m F_m \equiv 1$ so that $\sum_m P_{F_m} = \mathrm{Id}_{L^2}$.

5 Spawning Semiclassical Wavepackets

Our goal is to approximate a given wavefunction $\psi \in L^2\left(\mathbb{R}^d\right)$ as a linear combination of $n \in \mathbb{N}$ Hagedorn families, that is

$$\psi \approx \sum_{m=0}^{n} \sum_{k \in \mathcal{K}} c_k^{(m)} \varphi_k^{\varepsilon}\left[q_m, p_m\right] \tag{5.1}$$

from some coefficient vectors $c^{(m)} \in \mathbb{C}^{|\mathcal{K}|}$ and a truncated multi-index $\mathcal{K}$ set as in (2.6). In principle, we could use parameter sets different from $\Pi_m = (q_m, p_m, \mathrm{Id}, i\mathrm{Id})$, but this did not improve the results, see Remark 5.2. The following is worth noting.

(i) Since the Hagedorn wavepackets are an orthonormal basis, a single family would be enough to approximate ψ given that $\mathcal{K}$ is sufficiently large.

(ii) The two functions $\varphi_j^{\varepsilon}\left[q_m, p_m\right]$ and $\varphi_k^{\varepsilon}\left[q_m, p_m\right]$ of the same family are orthogonal if $j \neq k$. Functions of different parameters $(q_\ell, p_\ell) \neq (q_m, p_m)$ are never orthogonal.

Item (i) means that we seek points (q_m, p_m) in phase space that yield sparse coefficients meaning

$$\frac{|c_k^{(m)}|}{|c_0^{(m)}|} \text{ decays fast as } |k| \to \infty.$$

Item (ii) means that for given points (q_m, p_m), we want to determine $c^{(m)}$ by a method that can use the orthogonality within the families, but cares also about the non-orthogonality across the families. The localization operators achieve both at the same time as we will see below. We will now explain our expansion scheme that consists of two steps.

5.1 Fit the Parameters (Expectation Maximization)

The parameters $(q_m, p_m) \in \mathbb{R}^{2d}$ are computed by fitting a Gaussian mixture to the Husimi transform $H^{\varepsilon} := \mathcal{H}^{\varepsilon}[\psi]$ of our target function ψ, that is

$$H^{\varepsilon}(q,p) \approx f(q,p) := \sum_{m=1}^{n} r_m \, H_0^{\varepsilon}\left[q_m, p_m\right](q,p) \tag{5.2}$$

for some means (q_m, p_m) and some mixture probabilities $r_m \in [0,1]$ with $r_1 + \ldots + r_n = 1$. To this end, we will use the associated partition of unity on phase space

$$F_m := \frac{r_m H_0^{\varepsilon}\left[q_m, p_m\right]}{\sum_{\ell=1}^{n} r_\ell H_0^{\varepsilon}\left[q_\ell, p_\ell\right]}.$$

In the context of Bayesian inference, $F_m(q,p)$ is usually denoted by $\pi(m|q,p)$ and is called the *posterior* probability given (q,p). In order to determine good parameters r_m, q_m, p_m of the fitting problem (5.2), we use a modification of the well-known *expectation maximization* (EM) algorithm from [10]. We start with an initial guess for the parameters r_m, q_m, p_m, where $1 \leq m \leq n$. The algorithm then alternates N-times between the following two steps.

E-Step: Compute the normalization constants

$$\widehat{r}_m = \int_{\mathbb{R}^d \times \mathbb{R}^d} F_m\,(q,p)\; H^\varepsilon\,(q,p)\;\mathrm{d}q\mathrm{d}p,$$

where $1 \leq m \leq n$.

M-Step: Obtain a new candidate $(\widehat{q}_m, \widehat{p}_m)$ for the mean of index m by computing its expectation under the density

$$f_m\,(q,p) := \widehat{r}_m^{-1}\, F_m\,(q,p)\;\; H^\varepsilon\,(q,p)\,.$$

This means, we compute for all $1 \leq m \leq n$ the integral

$$\begin{pmatrix} \widehat{q}_m \\ \widehat{p}_m \end{pmatrix} = \frac{1}{\widehat{r}_m} \int_{\mathbb{R}^d \times \mathbb{R}^d} \begin{pmatrix} q \\ p \end{pmatrix} F_m\,(q,p)\; H^\varepsilon\,(q,p)\;\mathrm{d}q\mathrm{d}p. \tag{5.3}$$

Then we replace (q_m, p_m) by $(\widehat{q}_m, \widehat{p}_m)$ and r_m by $\widehat{r}_m$ for all $1 \leq m \leq n$. This changes also the partition of unity $F_1, \ldots, F_n$.

While we use H^ε, the classical EM works on samples thereof. This variation apeared in [1, Ch. 3.3]. In the discrete case, this algorithm seeks to maximize the log-likelihood of the observed data. In our (continuous) case, it seeks to minimize [1, Ch. 3.3] the *cross-entropy* or *Kullback-Leibler divergence*

$$D\,(H^\varepsilon \| f) := \int_{\mathbb{R}^d \times \mathbb{R}^d} H^\varepsilon\,(q,p) \log\left(\frac{H^\varepsilon\,(q,p)}{f\,(q,p)}\right) \mathrm{d}q\mathrm{d}p. \tag{5.4}$$

It is known that $D\,(H^\varepsilon \| f) \geq 0$. Moreover, $D\,(H^\varepsilon \| f) = 0$ if and only if $H^\varepsilon = f$.

Remark 5.1. The classical EM-algorithm learns the mixture distribution from samples $(x_1, k_1), \ldots, (x_N, k_N) \in \mathbb{R}^{2d}$ according to the (unknown) Husimi density H^ε. We obtain the classical EM from our algorithm if we replace H^ε by the empirical density [1, A.2]

$$H^\varepsilon_{\mathrm{emp}} := \frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i,k_i)},$$

where $\delta_{(x,k)}$ is the delta distribution at (x,k). For example, the Kullback-Leibler divergence turns into the negative log-likelihood if we replace H^ε by $H^\varepsilon_{\mathrm{emp}}$

$$D\left(H^\varepsilon_{\mathrm{emp}} \| f\right) = -\frac{1}{N} \sum_{i=1}^{N} \log f\,(x_i, k_i)\,.$$

Likewise we can recover the classical EM updates for $\widehat{r}_m$ and $(\widehat{q}_m, \widehat{p}_m)$, see [1, A.2].

Since $\sum_m F_m \equiv 1$, we have

$$D\,(H^\varepsilon \| f) = \sum_{m=1}^{n} \widehat{r}_m \log\left(\frac{\widehat{r}_m}{r_m}\right) + \sum_{m=1}^{n} \widehat{r}_m\, D\left(\widehat{r}_m^{-1}\, F_m\, H^\varepsilon \parallel H^\varepsilon_0\,[q_m, p_m]\right). \tag{5.5}$$

Another reason for the choice of $(\widehat{q}_m, \widehat{p}_m)$ in (5.3) is due to the following proposition, which also proves that it can only decrease the Kullback-Leibler divergence.

Proposition 5.1. Fix r_m and F_m for all $1 \leq m \leq n$, the pair $(\widehat{q}_m, \widehat{p}_m)$ in (5.3) satisfies

$$\begin{pmatrix} \widehat{q}_m \\ \widehat{p}_m \end{pmatrix} = \underset{(\tilde{q},\tilde{p})}{\arg\min}\, D\left(\widehat{r}_m^{-1}\, F_m\, H^\varepsilon \parallel H_0^\varepsilon\,[\tilde{q},\tilde{p}]\right)$$

Proof. The gradient w.r.t. $(\tilde{q},\tilde{p})$ is

$$\tilde{\nabla}\, D\left(\widehat{r}_m^{-1}\, F_m\, H^\varepsilon \parallel H_0^\varepsilon\,[\tilde{q},\tilde{p}]\right) = -\frac{1}{\varepsilon^2}\begin{pmatrix} \widehat{q}_m - \tilde{q} \\ \widehat{p}_m - \tilde{p} \end{pmatrix}$$

The Hessian equals the positive definite matrix $\frac{1}{\varepsilon^2}\mathrm{Id}$. Setting the gradient to zero finishes the proof. □

Example 5.1. In order to be able to see the algorithm at work, we consider an example in dimension $d = 1$ with $n = 6$ families:

$$\psi\,(x) = \tilde{c}_0^{(1)}\varphi_0^\varepsilon\,[q_1,p_1] + \cdots + \tilde{c}_0^{(6)}\varphi_0^\varepsilon\,[q_6,p_6]\,, \tag{5.6}$$

where the parameters and coefficients are in Table 1. The Husimi transform $\mathcal{H}^\varepsilon\,[\psi]$ of

m	1	2	3	4	5	6
$\tilde{c}_0^{(m)}$	0.5	$0.2+0.2i$	$-0.5+0.4i$	-0.5	$0.25+0.25i$	$0.75+0.1i$
(q_m,p_m)	$(0,0)$	$(4,-1)$	$(5,-1)$	$\left(\frac{11}{2},\frac{11}{2}\right)$	$\left(-\frac{5}{2},3\right)$	$(-6,-2)$

Table 1: Parameters and coefficients of the linear combination (5.6).

ψ along with the centers of the initial and final Gaussians mixtures are plotted in Figure 2. We have used equal mixture probabilities $r_m = \frac{1}{5}$ for the initial Gaussian mixture;

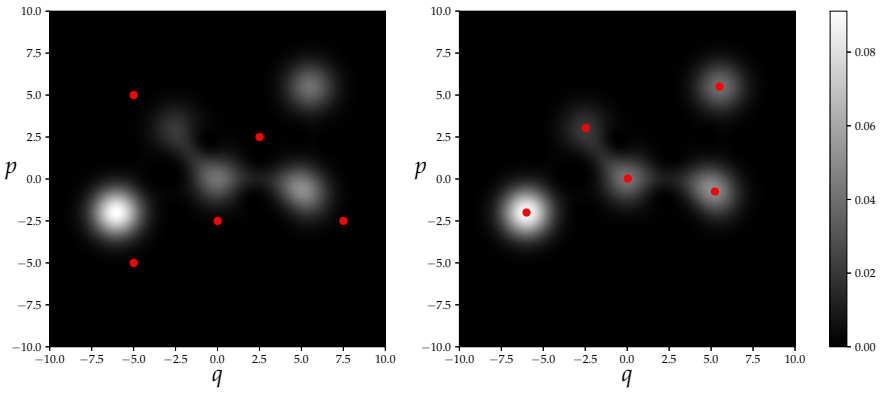

Figure 2: The Husimi transform of ψ and the centers of the means before the first iteration (left) and after the last iteration (right).

the initial centers were intentionally selected far from the solution in order to give the optimization algorithm a harder task. Then 30 iterations of the EM-algorithm were performed. The Kullback-Leibler divergence $D\,(H^\varepsilon \| f)$ vs. the number of iterations is shown in Figure 3. All phase space integrals are computed using 64 Gauss-Hermite quadrature points translated to the corresponding previous guess (q_m, p_m) for the mean of the corresponding mixture component. Note that exponential convergence is only due to the very simple setting we are using here. In general the convergence of EM can be very slow [9].

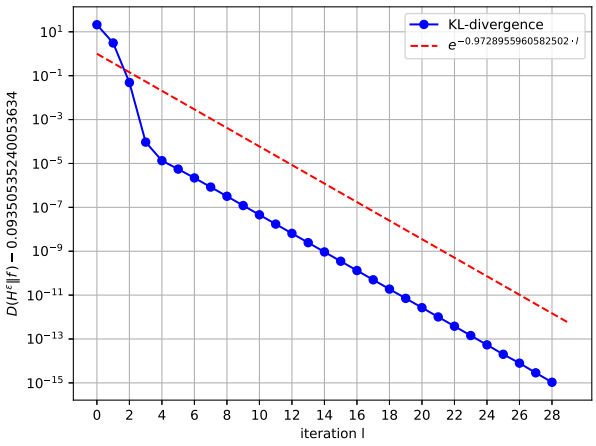

Figure 3: Kullback-Leibler divergence $D\left(H^{\varepsilon} \| f\right)$ after each iteration. At the last iteration, the Kullback-Leibler divergence equals 0.09350535240053634.

Remark 5.2. So far we have only fitted a mixture of Gaussians with identity covariance matrix. But we could also fit the covariance matrices in (5.2). Doing so, we have to respect the constraint on the covariance matrix of Husimi groundstates

$$\Sigma = \frac{\varepsilon^2}{2}\left(F_0 F_0^T + \mathrm{Id}\right), \tag{5.7}$$

where F_0 is symplectic, see Proposition 3.1. A simple sub-family of symplectic $2d \times 2d$ matrices is given by

$$F_0(\alpha) = \begin{pmatrix} e^{\alpha} & 0 \\ 0 & e^{-\alpha} \end{pmatrix}, \qquad \alpha \in \mathbb{R}.$$

In this case, we can explicitly compute the maximum likelihood estimator for α and thus an update in the EM-algorithm for the covariance matrices

$$\Sigma(\alpha) = \frac{\varepsilon^2}{2}\begin{pmatrix} e^{2\alpha}+1 & 0 \\ 0 & e^{-2\alpha}+1 \end{pmatrix} = \varepsilon^2 \cosh(\alpha) \exp\left(\alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right)$$

By direct computation, we find

$$\begin{aligned} \det\left(\Sigma(\alpha)\right) &= \left(\varepsilon^2 \cosh(\alpha)\right)^{2d} \\ \frac{\mathrm{d}}{\mathrm{d}\alpha} \log\left(\det\left(\Sigma(\alpha)\right)\right) &= 2d \tanh(\alpha) \\ \frac{\mathrm{d}}{\mathrm{d}\alpha} \Sigma^{-1} &= \frac{\varepsilon^2}{\cosh^2(\alpha)} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Using these expressions we can compute the derivative w.r.t. α of the Kullback-Leibler divergence like in the proof of Proposition 5.1. Setting it to zero, we obtain

$$\widehat{\alpha} = \tanh^{-1}(\beta), \qquad \beta := \frac{-d + \sqrt{d^2 + S_\varepsilon^2}}{S_\varepsilon},$$

where

$$S_\varepsilon = \frac{1}{\varepsilon^2 N} \sum_{i=1}^{N} \left(\|\widehat{q}_m - x_i\|^2 - \|\widehat{p}_m - k_i\|^2\right)$$

in the descrete case (see Remark 5.1) and

$$S_\varepsilon = \frac{1}{\varepsilon^2 \widehat{r}_m} \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\|\widehat{q}_m - q\|^2 - \|\widehat{p}_m - p\|^2 \right) F_m(q,p)\ H^\varepsilon(q,p)\ \mathrm{d}q\mathrm{d}p$$

in the continuous case. Here, $(\widehat{q}_m, \widehat{p}_m)$ are updates for the means, which remain the same for any parametrization of Σ. Note that $\widehat{\alpha}$ is well defined since $\beta \in (-1,1)$ and $\beta \to 0$ as $S_\varepsilon \to 0$. However, optimizing for Σ did not gave better results in our simulations, except for very specific cases. Moreover, not respecting the constraint (5.7) at all also leads to worse results. Hence we stick to $\Sigma = \mathrm{Id}$.

5.2 Compute the Coefficients (Localization Operators)

Recall that we want to expand a given wavefunction $\psi \in L^2\left(\mathbb{R}^d\right)$ as in Equation (5.1), that is

$$\psi \approx \sum_{m=0}^{n} \sum_{k \in \mathcal{K}} c_k^{(m)} \varphi_k^\varepsilon [q_m, p_m], \tag{5.1}$$

We have already obtained the parameters $(q_m, p_m) \in \mathbb{R}^{2d}$ by fitting a Gaussian mixture

$$\mathcal{H}^\varepsilon[\psi](q,p) \approx \sum_{m=1}^{n} r_m H_0^\varepsilon [q_m, p_m](q,p)$$

to the Husimi transform of ψ. We use the corresponding posteriors

$$F_m := \frac{r_m H_0^\varepsilon [q_m, p_m]}{\sum_{\ell=1}^{n} r_\ell H_0^\varepsilon [q_\ell, p_\ell]}$$

as symbols for the localization operators from Section 4, that is

$$P_{F_m} := \left(2\pi\varepsilon^2\right)^{-d} \int_{\mathbb{R}^d \times \mathbb{R}^d} F_m(q,p)\ \varphi_0^\varepsilon[q,p]\ \langle \varphi_0^\varepsilon[q,p],\ \cdot\ \rangle_{L^2}\ \mathrm{d}q\,\mathrm{d}p.$$

Due to the partition of unity property $F_1 + \ldots F_n \equiv 1$, we have $P_{F_1} + \ldots + P_{F_n} = \mathrm{Id}_{L^2}$ and thus

$$\psi = P_{F_1}\psi + \ldots + P_{F_n}\psi.$$

The function $P_{F_m}\psi$ is concentrated in the phase space near (q_m, p_m). In order to obtain the coefficients in (5.1) we can simply project it onto the corresponding Hagedorn wavepacket

$$c_k^{(m)} := \langle \varphi_k^\varepsilon [q_m, p_m], P_{F_m}\psi \rangle_{L^2(\mathbb{R}^d)} = \langle \mathcal{B}^\varepsilon [\varphi_k^\varepsilon [q_m, p_m]], F_m\, \mathcal{B}^\varepsilon[\psi] \rangle_{L^2(\mathbb{R}^{2d})}, \tag{5.8}$$

where $\mathcal{B}^\varepsilon$ denotes the Bargmann transform from Section 3.1. We use the three-term recurrence of Proposition 3.1 and Gauss-Hermite quadrature translated to (q_m, p_m) to compute the phase space integral on the right. Due to the localization property of P_{F_m}, this will typically yield quickly decaying coefficients.

Example 5.2. We consider the same wave function ψ as in Example 5.1, that is

$$\psi(x) = \tilde{c}_0^{(1)} \varphi_0^\varepsilon [q_1, p_1] + \cdots + \tilde{c}_0^{(6)} \varphi_0^\varepsilon [q_6, p_6],$$

where the means and coefficients are in Table 1. We take the exact means $(q_1, p_1), \ldots, (q_6, p_6)$ and project onto the corresponding truncated Hagedorn wavepackets

$$\psi \approx \sum_{m=1}^{6} \sum_{k=0}^{63} c_k^{(m)} \varphi_k^{\varepsilon} [q_m, p_m],$$

where the coefficients $c_k^{(m)}$ have been computed by (5.8). The L^2-error between ψ and our approximation is $1.12 \cdot 10^{-5}$. Each of the six phase space integrals was computed by a grid of 64×64 Gauss-Hermite quadrature points translated to (q_m, p_m). In Figure 4 we can see the FBI transform of ψ and the strongly localized functions $P_{F_m}\psi$. In Figure 5 we

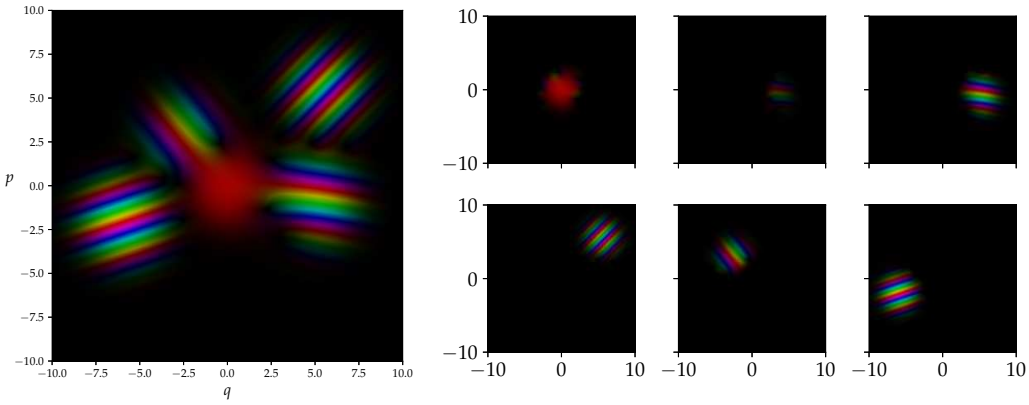

Figure 4: The FBI transforms of ψ (left) and $P_{F_m}\psi$ for $m = 1, \ldots, 6$ (right). The color indicates the complex phase $e^{i\alpha}$ for $\alpha \in [0, 2\pi)$ at each point.

have assigned a color to each Hagedorn family. We see that the corresponding posterior F_m really cuts off the other families, so that the projection yields exponentially decaying coefficients as shown in the same figure.

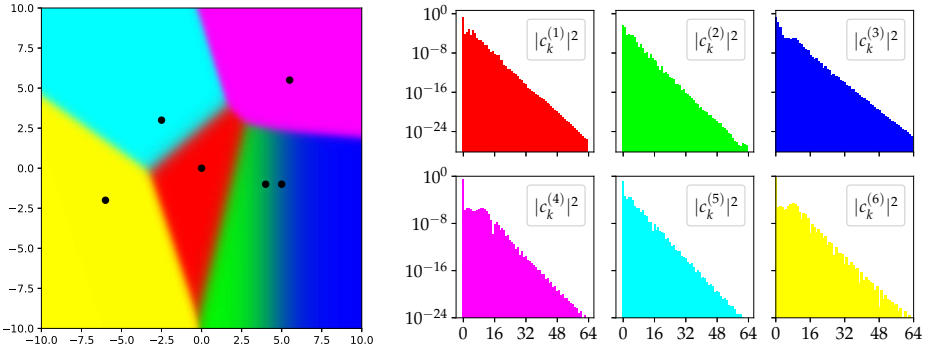

Figure 5: For $m = 1, \ldots, 6$ and $k = 0, \ldots, 63$, we show the partition of unity F_m (left) and the coefficients $|c_k^{(m)}|^2$ (right) and in the corresponding colors. The black dots on the left indicate the means (q_m, p_m). The color labels the six different families.

6 Application: Tunneling

The semiclassical formulation of the time-dependent Schrödinger equation with potential $V : \mathbb{R}^d \to \mathbb{R}$ is

$$-i\varepsilon^2 \partial_t \psi(x,t) = \left(-\tfrac{\varepsilon^4}{2}\Delta + V(x,t)\right)\psi(x,t), \qquad \psi(x,0) = \psi_0(x), \tag{6.1}$$

where $x \in \mathbb{R}^d, t \in \mathbb{R}$ and $\varepsilon > 0$. The semiclassical splitting [11] is a possible algorithm to propagate a single Hagedorn family

$$\psi(x) = \sum_{k \in \mathcal{K}} c_k e^{iS(t)/\varepsilon^2} \varphi_k^\varepsilon [q, p, Q, P](x)$$

according to this equation. We refer to [5] for details about the semiclassical splitting. The basic idea for propagation by a time step h is as follows.

1. Write $V = U_q + W_q$, where U is the quadratic Taylor expansion of $V(x)$ at $x = q$.
2. Propagate ψ by $\frac{h}{2}$ according to the Hamiltonian $-\frac{\varepsilon^4}{2}\Delta + U_{q(0)}$. This is achieved simply by propagating the parameters according to the classical equations of motion

   $$\dot{q} = p, \quad \dot{p} = -\nabla V(q), \quad \dot{Q} = P, \quad \dot{P} = -\nabla^2 V(q) Q, \quad \dot{S} = \frac{1}{2} p^T p - U_{q(0)}(q).$$

   The coefficients c_k remain fixed.
3. Propagate the resulting wavefunction by h according to the Hamiltonian $W_{q(h/2)}$ by a Galerkin method that changes the coefficients c_k and keeps the parameters fixed.
4. Propagate the parameters by $\frac{h}{2}$ according to $-\frac{\varepsilon^4}{2}\Delta + U_{q(h/2)}$ similar to Step 2.

This is basically a Strang splitting between $-\frac{\varepsilon^4}{2}\Delta + U_q$ and W. The semiclassical splitting has been improved by more sophisticated splittings that treat W as a perturbation. In the simulations below, we use an enhancement proposed in [5, Chapter 3.4] that has roughly the same computational effort as the semiclassical splitting, but is of improved order $\varepsilon^2 h^6 + \varepsilon^4 h^4$ and thus performs particularly well for small ε, see [5] for details.

6.1 Eckart Potential

As an example of quantum mechanical tunneling with delocalization in space, we consider the Eckart potential

$$V(x) = \frac{V_0}{\cosh^2(ax)},$$

where $V_0 = 0.038008$ and $a = 0.94485808231$. At time $t = 0$ we have a particle described by the wavefunction $\varphi_0^\varepsilon[\Pi]$, where $\Pi := (q, p, Q, P)$ with

$$q := -7.5589045088306, \quad p := 0.2478854736792$$

and

$$Q := 3.5355339059327, \quad P := 0.2828427124746\, i,$$

see Figure 6. The model parameter is $\varepsilon = 0.1530417681822$, see Figure 6. This setting is also treated in [27] and [13] in the context of Hagedorn wavepackets. However, note the different convention $\hbar \mathrel{\widehat{=}} \varepsilon^2$ in the present work and $\hbar \mathrel{\widehat{=}} \varepsilon$ in [13]. This inital data is then projected onto a linear combination of Hagedorn families with $K = 64$ coefficients. All phase space integrals were computed by a grid of 64×64 translated Gauss-Hermite quadrature points. The starting values (at any time step) for the EM-algorithm are plotted in Figure 7. We perform 30 iterations of the EM-algorithm. In every iteration, we discard all Gaussians with weight $\widehat{r}_m < 10^{-6}$. The effective number of Hagedorn families used for propagation is plotted in Figure 8 (left). We use a time step of size $h = 0.01$ and

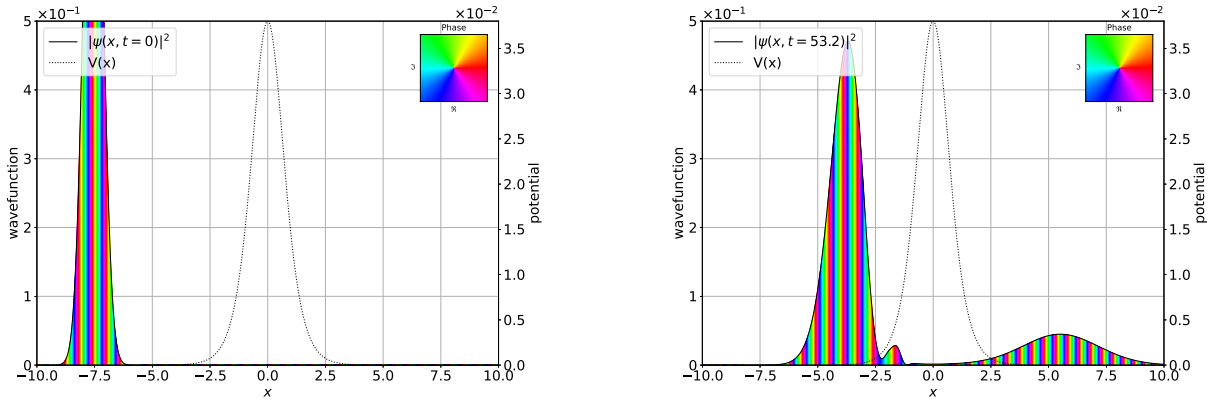

Figure 6: Initial wave function (left) and numerical solution (right) at time $t = 53.2$ or 54.93964 femtoseconds. The dotted line shows the Eckart potential.

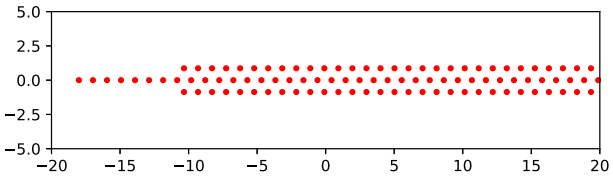

Figure 7: The 103 starting values for the means in the EM-algorithm (Eckart poential).

solve until time $T = 70$ which corresponds to 72.289 femtoseconds [13, IV]. After every 280 steps, we re-expand the propagated wavefunction. This corresponds to the jumps of the blue curve in Figure 8, which is the L^2 difference to the reference solution. The latter is the solution given by a splitting of order 8 in time [19, p. 145] with a time step size $h = 0.01$, together with a Fourier approximation in space on an equidistant grid on $[-40, 40]$ with 16384 points. Note that the error stems mostly from the re-expansions and *not* from the propagation. In the example in Section 6.2 it is exactly the other way around. The reason is that here the exact solution has very heavy tails, which cannot be approximated well by Gaussians with identity covariance matrix. The cumulative execution time of the re-expansion and the propagation is split like 51% (re-expansion) and 49% (propagation). In Figure 8 (right) we show the tunneling probability

$$\int_0^\infty |\psi(x,t)|^2 \, \mathrm{d}x$$

according to the numerical solution. We obtain the same result as in [13, Fig. 3 (b)]. Finally, Figure 9 shows the energy along the numerical solution. This is again in accordance with the results in [13, Fig. 2].

6.2 Doublewell Potential (1D)

We consider a doublewell potential

$$V(x) = x^4 - x^2$$

that is tailored to produce a suitable phase space picture to illustrate our spawning of wavepackets, see Figure 13. To make use of the high-order splitting methods, we choose the small value $\varepsilon = 0.1$ for the model parameter. The initial data is $\varphi_0^\varepsilon[0.9, 0.0]$, see Figure 11 (right). We use a time step of size $h = 0.05$ and simulate until time $T =$

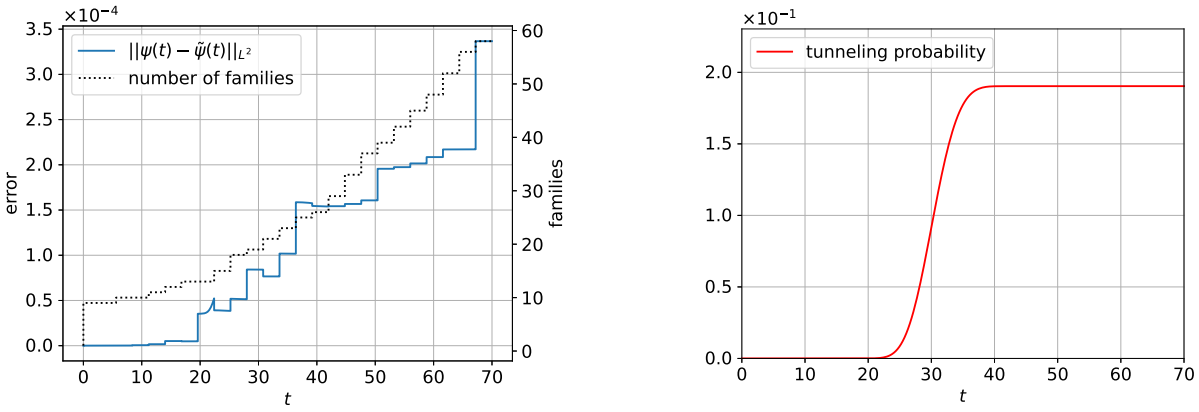

Figure 8: Left: Error and number of families. The end time $T = 70$ corresponds to 72.289 femtoseconds. Right: Tunneling probability in the Eckart potential.

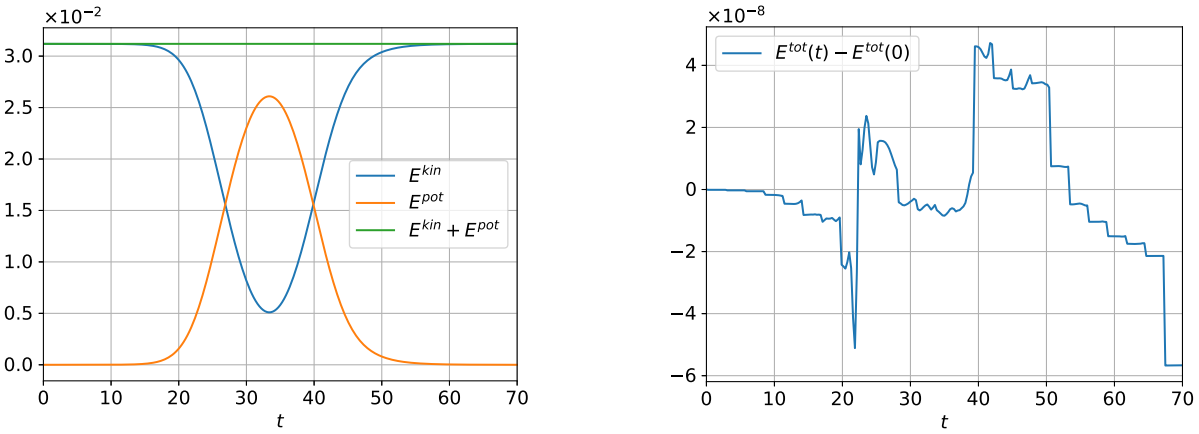

Figure 9: Left: Kinetic energy, potential energy and total energy along the numerical solution of the Eckart potential. Right: Drift of total energy along the numerical solution.

10. Every 5 time steps we re-expand the numerical solution by families with $K = 64$ coefficients. We initialize the EM-algorithm each time with the (q_m, p_m) pairs in Figure 11 (left). We perform 30 iterations of the algorithm and discard all Gaussians with weight $\widehat{r}_m < 10^{-6}$. Figure 10 shows the L^2-error to a reference solution and the effective number of families in use (left), along with the position representation of the wavefunction at time $t = 5.5$. The reference solution is computed as in the previous example, now with a time step size $h = 0.05$ and on an equidistant grid on $[-2, 2]$ with 4096 points. Note that the error is coming mostly from the propagation and *not* from the re-expansion, opposed to our observation in Section 6.1. Unlike in the previous example, the cumulative execution time is split up very unevenly like 97% (re-expansion) and 3% (propagation). In Figure 12 we see the energy along the solution. Figure 13 shows the FBI transform of the numerical solution. We can see how the wavepackets are spawned and propagated in phase space.

6.3 Doublewell Potential (2D)

In this section, we consider the two dimensional doublewell potential

$$V(x) = x_1^4 - x_1^2 + \tfrac{1}{2}\left(x_2 - \tfrac{1}{8}x_1^2\right)^2,$$

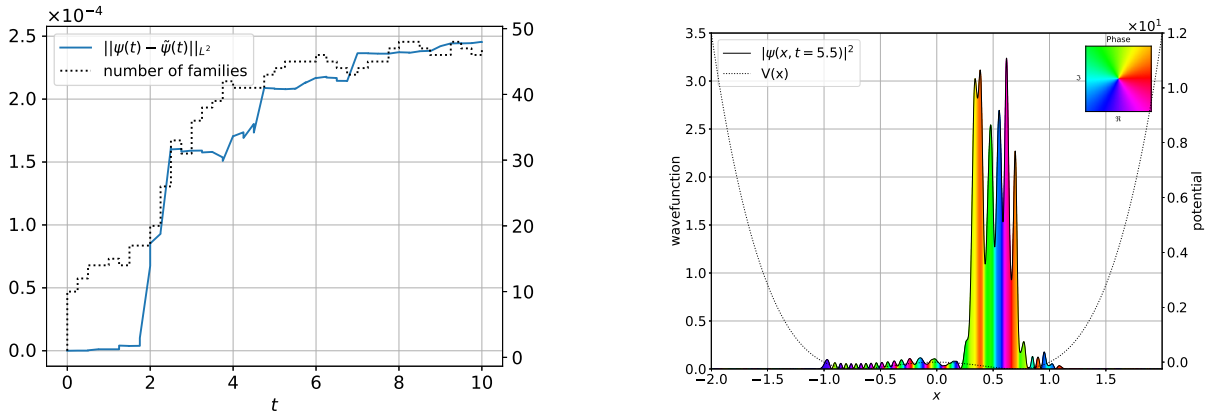

Figure 10: Left: L^2-error and number of families. Right: position representation of the wavefunction at time $t = 5.5$. The dotted line represents the doublewell potential.

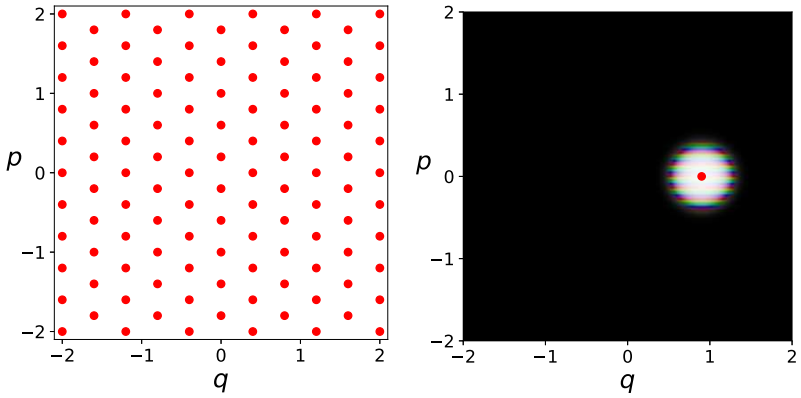

Figure 11: Left: The 116 starting values for the means in the EM-algorithm for the doublewell potential. Right: FBI transform of the initial wave function. The red dot marks the mean.

where $x = (x_1, x_2) \in \mathbb{R}^2$. The model parameter is again $\varepsilon = 0.1$. The initial data is the Hagedorn groundstate $\varphi_0^\varepsilon[q, p]$, where

$$q = \begin{pmatrix} 0.9 \\ 0.405 \end{pmatrix} \qquad \text{and} \qquad q = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We use a time step of size $h = 0.05$ and simulate until time $T = 6$. Every 5 time steps we re-expand the numerical solution by families with index set given by Equation (2.6) for $K = 64$. We use a 4-dimensional Gauss-Hermite quadrature for all phase space integrals. In the EM-algorithm, we use 16^4 quadrature points and for the localization operators we use 32^4 quadrature points. In both cases, the quadrature points are translated to the means of the corresponding Gaussians in phase space. We initialize the EM-algorithm each time with Gaussians of means

$$\left(q^{(1)}, 0, p^{(1)}, 0\right)^T,$$

where the $\left(q^{(1)}, p^{(1)}\right)$ pairs run over the values in Figure 14 (left). Then we perform 10 iterations of the algorithm and discard all Gaussians with weight $\widehat{r}_m < 10^{-5}$. In Figure 14 (right) we see the number of families in use and the error to a reference solution. The latter is again computed as in the two previous examples, with a time step size $h = 0.05$

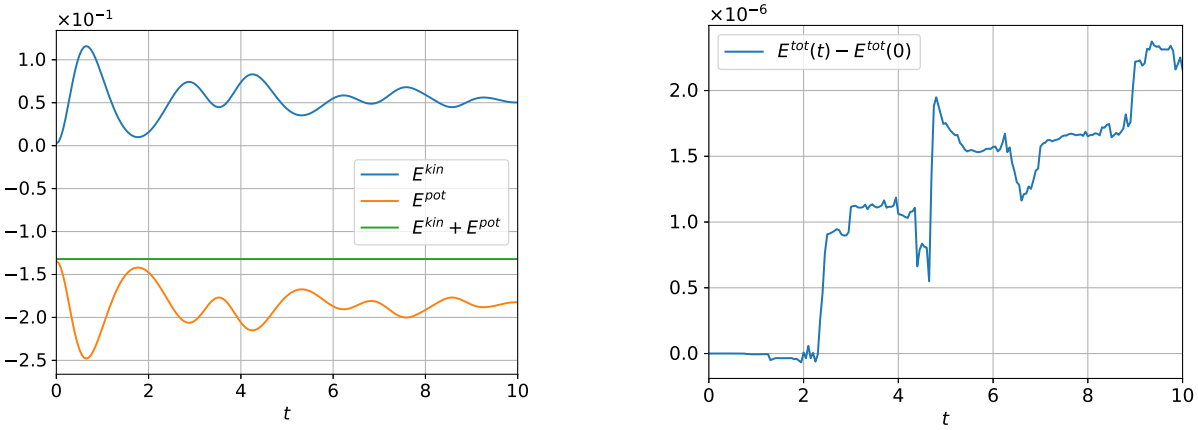

Figure 12: Left: Kinetic energy, potential energy and total energy along the numerical solution of the doublewell potential (1D). Right: Drift of total energy along the numerical solution.

and on an equidistant spatial grid on $[-4, 4] \times [-1, 1]$, with spacing 2^{-7} in both directions. The energy along the solution is plotted in Figure 15. The cumulative execution time is again split up very unevenly like 99% (re-expansion) and 1% (propagation). Finally, we show the position representation of the numerical solution (not the FBI transform) and of the reference solution in Figure 16.

6.4 Computational Effort

The computational effort of the re-expansion is quadratic in the number of families in use. On the other hand, the time evolution is linear in the number of families since every family is propagated independently. The re-expansion is the bottleneck of this time propagation scheme. The most expensive part in the re-expansion is the EM-algorithm and the computation of the involved phase space integrals. Unlike the integrals that typically appear in an ansatz the pure physical space, our integrals in the phase space are *not* highly-oscillatory, hence they are easier to compute. In higher dimensions, we can replace the Gauss-Hermite Quadrature by a Monte-Carlo quadrature with importance sampling. The latter is done by sampling according to the Gaussian mixture associated with the propagated wavefunction. We have applied such a Monte-Carlo quadrature with importance sampling to the examples of the last section. We have observed that this requires also to change the threshold for ignoring mixture components and the number of iterations of the EM-algorithm, which makes the two approaches a bit harder to compare. In general, we have obtained similar results with similar computational effort in these one-dimensional examples. However, in case of a larger number of dimensions d, the use of a (quasi-) Monte-Carlo integration maybe be indispensable. This suggests that the method can possibly be improved by choosing a more efficient optimization algorithm to determine the parameters of the wavepackets and the symbols of the associated localization operators.

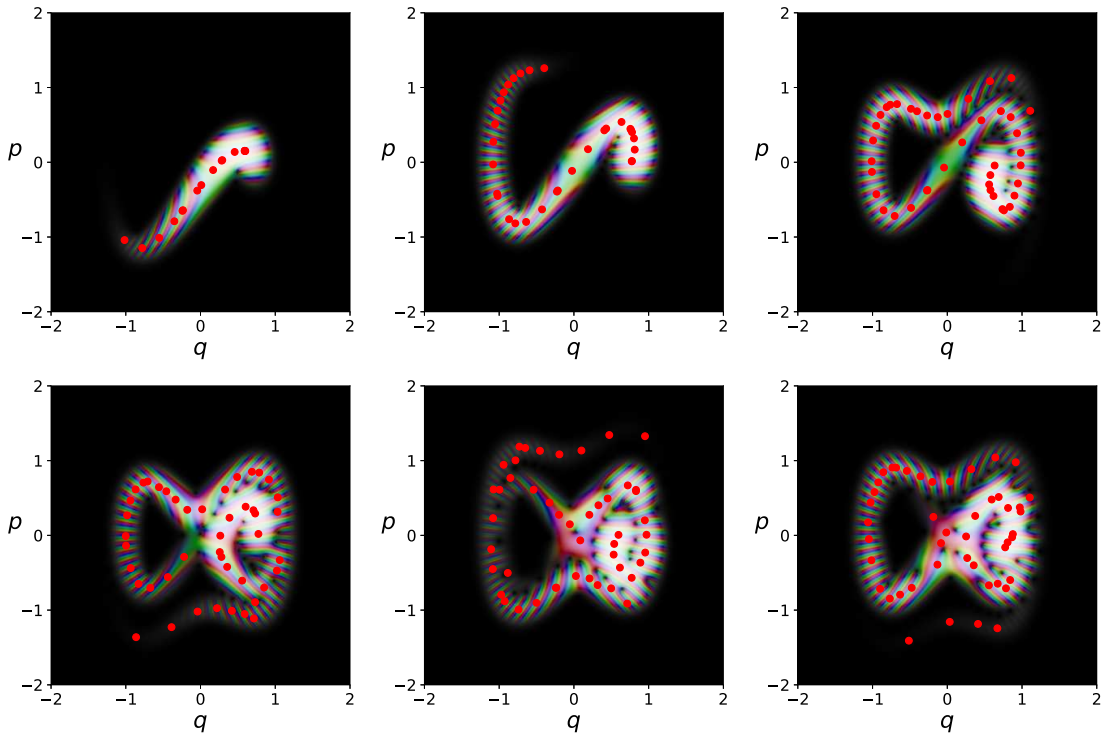

Figure 13: FBI transform of the numerical solution at times $t = 2, 3, 4.5, 6, 8, 10$ (from top-left to bottom-right, before the re-expansion) for the doublewell potential (1D). The red dots mark the means of the Hagedorn families.

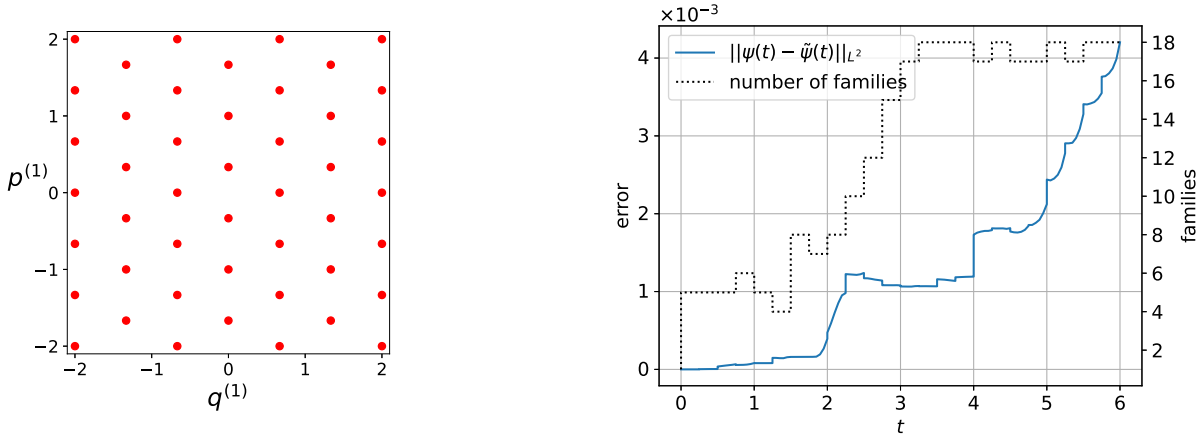

Figure 14: Left: The 46 starting values for the means in the EM-algorithm, projected onto the first coordinates of position and momentum. Right: Error and number of families at different times for the numerical solution of the doublewell potential (2D)

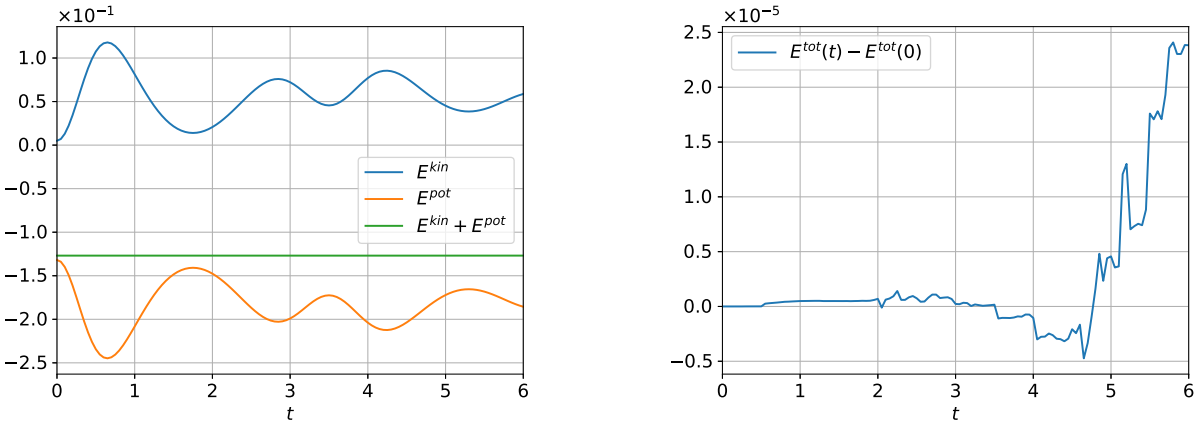

Figure 15: Left: Kinetic energy, potential energy and total energy along the numerical solution of the doublewell potential (2D). Right: Drift of total energy along the numerical solution.

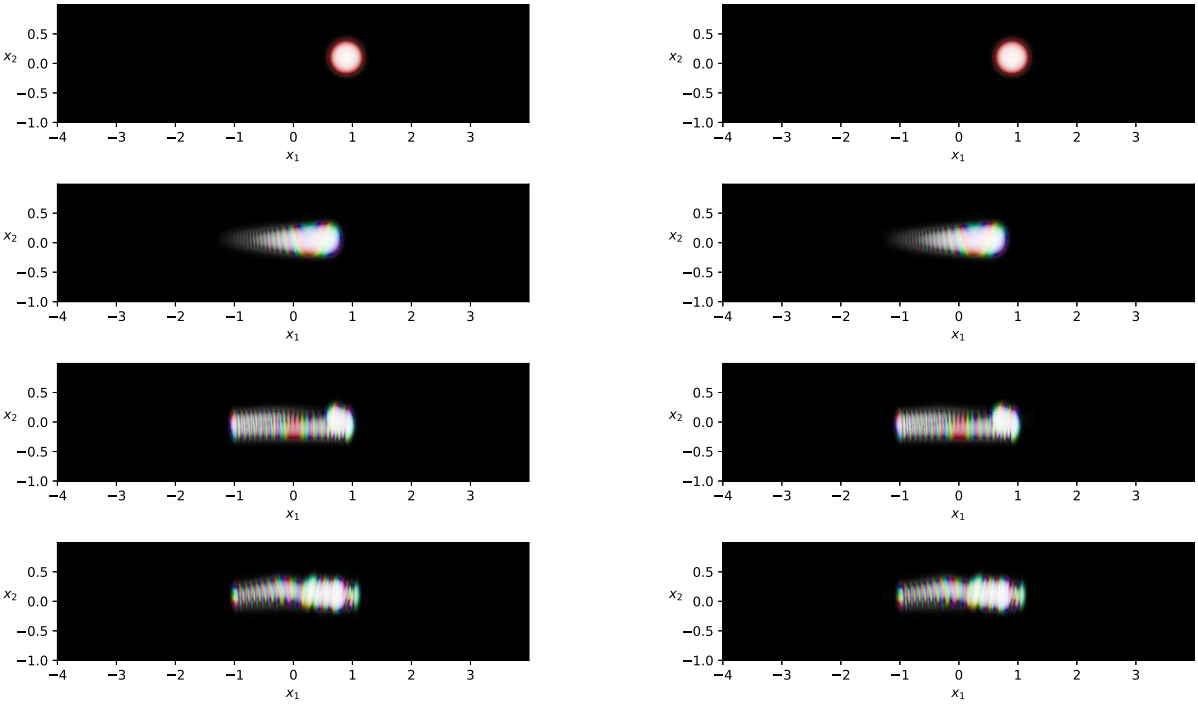

Figure 16: Position representation (not FBI transform) of the numerical solution (left) and of the reference solution (right) of the doublewell potential (2D) at times $t = 0, 2, 4, 6$ (from top to bottom).

References

[1] Søren Asmussen, Olle Nerman, and Marita Olsson. Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics*, 23(4):419–441, 1996.

[2] Richard Bale, Jeff Grossman, Gary Margrave, and Michael Lamoureux. Multidimensional partitions of unity and Gaussian terrains. 10 2002.

[3] V. Bargmann, P. Butera, L. Girardello, and John R. Klauder. On the completeness of the coherent states. *Reports on Mathematical Physics*, 2(4):221–228, 1971.

[4] Paul Bergold and Caroline Lasser. An error bound for the time-sliced thawed Gaussian propagation method. *Numerische Mathematik*, Sep 2022.

[5] Sergio Blanes and Vasile Gradinaru. High order efficient splittings for the semiclassical time–dependent Schrödinger equation. *Journal of Computational Physics*, 405:109157, 2020.

[6] R. Bourquin. *Numerical Algorithms for Semiclassical Wavepackets*. PhD thesis, ETH Zurich, 2016.

[7] Xin Chen and Victor S. Batista. Matching-pursuit/split-operator-fourier-transform simulations of excited-state nonadiabatic quantum dynamics in pyrazine. *The Journal of Chemical Physics*, 125(12):124313, 2006.

[8] Ingrid Daubechies. Time-frequency localization operators: A geometric phase space approach. *IEEE Trans. Inf. Theory*, 34:605–612, 1988.

[9] Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. Cambridge University Press, 2020.

[10] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[11] Erwan Faou, Vasile Gradinaru, and Christian Lubich. Computing semiclassical quantum dynamics with Hagedorn wavepackets. *SIAM J. Scientific Computing*, 31:3027–3041, 01 2009.

[12] Gerald B. Folland. *Harmonic Analysis in Phase Space. (AM-122)*. Princeton University Press, Princeton, 2016.

[13] V. Gradinaru, G. A. Hagedorn, and A. Joye. Tunneling dynamics and spawning with adaptive semiclassical wave packets. *The Journal of Chemical Physics*, 132(18):184108, 2010.

[14] Vasile Gradinaru and George A. Hagedorn. Convergence of a semiclassical wavepacket based time-splitting for the Schrödinger equation. *Numerische Mathematik*, 126(1):53–73, Jan 2014.

[15] Vasile Gradinaru and Oliver Rietmann. A high-order integrator for the Schrödinger equation with time-dependent, homogeneous magnetic field. *The SMAI journal of computational mathematics*, 6:253–271, 2020.

[16] Samuel M. Greene and Victor S. Batista. Tensor-train split-operator fourier transform (tt-soft) method: Multidimensional nonadiabatic quantum dynamics. *Journal of Chemical Theory and Computation*, 13(9):4034–4042, 2017.

[17] George A. Hagedorn. Semiclassical quantum mechanics. i. the $\hbar \to 0$ limit for coherent states. *Comm. Math. Phys.*, 1(1):77–93, 1980.

[18] George A. Hagedorn. Raising and lowering operators for semiclassical wave packets. *Annals of Physics*, 269(1):77 – 104, 1998.

[19] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations*, volume 31 of *Springer series in computational mathematics*. Springer, Berlin, second edition, 2010.

[20] Kôdi Husimi. Some formal properties of the density matrix. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, 22(4):264–314, 1940.

[21] D. Iagolnitzer. Microlocal Essential Support of a Distribution and Decomposition Theorems–An Introduction: Appendix 1. *Lect. Notes Math.*, 449:121–132, 1975.

[22] Johannes Keller, Caroline Lasser, and Tomoki Ohsawa. A new phase space density for quantum expectations. *SIAM Journal on Mathematical Analysis*, 48(1):513–537, 2016.

[23] Xiangmeng Kong, Andreas Markmann, and Victor S. Batista. Time-sliced thawed Gaussian propagation method for simulations of quantum dynamics. *The Journal of Physical Chemistry A*, 120(19):3260–3269, 2016. PMID: 26845486.

[24] Caroline Lasser and Christian Lubich. Computing quantum dynamics in the semiclassical regime. *Acta Numerica*, 29:229–401, 2020.

[25] Caroline Lasser and Stephanie Troppmann. Hagedorn wavepackets in time-frequency and phase space. *Journal of Fourier Analysis and Applications*, 20, 03 2013.

[26] C. Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zurich lectures in advanced mathematics. European Mathematical Society, 2008.

[27] Panchanan Puzari and Satrajit Adhikari. Semi-classical formulation of time-dependent discrete variable representation method. *International Journal of Quantum Chemistry*, 98(5):434–446, 2004.