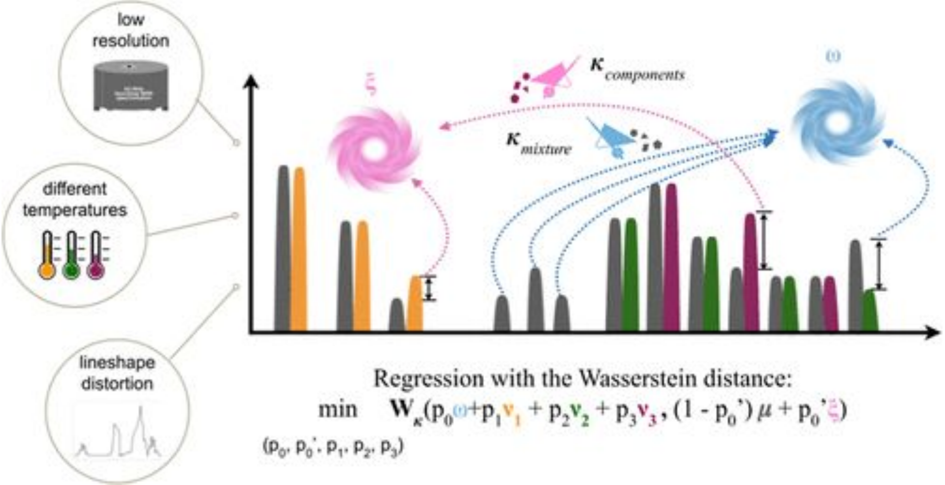
Visualization package for *Magnetstein*

Zofia Kozaryna, Ko Natoński

Introduction to the project

Magnetstein: An Open-Source Tool for Quantitative NMR Mixture Analysis Robust to Low Resolution, Distorted Lineshapes, and Peak Shifts



Regression with the Wasserstein distance:

$$\min_{(p_0,\, p_0',\, p_1,\, p_2,\, p_3)} \mathbf{W}_\kappa(p_0\omega + p_1\mathbf{v}_1 + p_2\mathbf{v}_2 + p_3\mathbf{v}_3,\ (1 - p_0')\,\mu + p_0'\xi)$$

The Magnetstein project addresses a key challenge in NMR-based mixture analysis: how to decompose complex spectra into their molecular components. While algorithms based on the optimal transport theory can estimate proportions with removing noise detected in spectra, interpreting and validating these results visually remains difficult—especially for non-specialist users.

Thus, **we developed a dedicated visualization package** to bridge this gap by turning computational outputs into clear, informative, and interactive spectral plots.

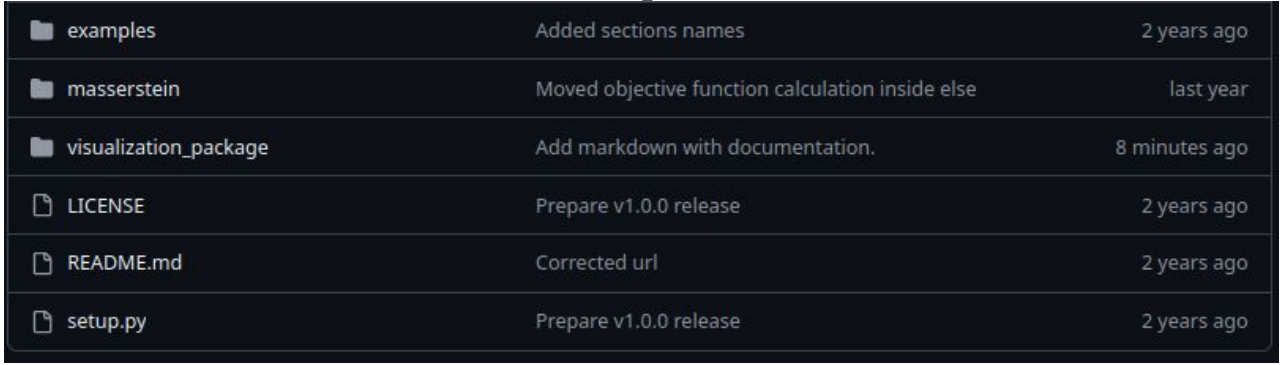
Objectives

The primary goals of the visualization module were:

- **Visual Clarity**: To provide intuitive spectral plots for mixtures, components, and decompositions.
- **Decomposition Insight**: To illustrate how a mixture is decomposed into components and how much each contributes.
- **Noise Interpretation**: To show what signal parts were considered "noise" and removed during normalization.
- **Transport Interpretation**: To visualize the transport plan — the optimal way intensities are shifted between mixture and components.

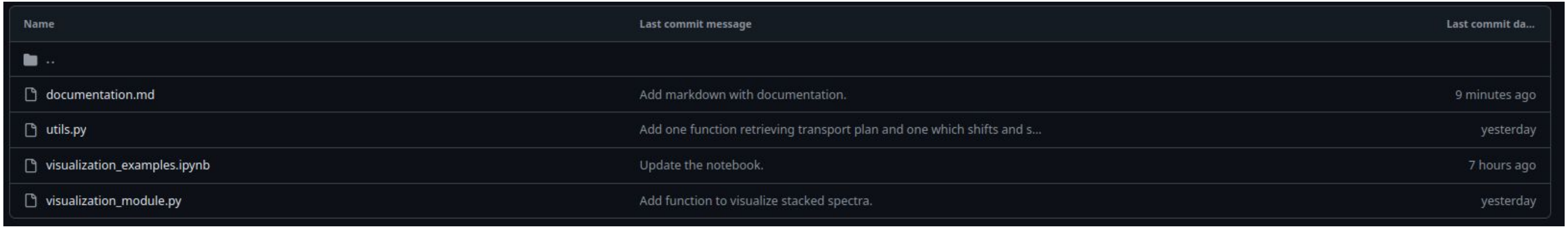
Github repository

official Magnetstein repository



examples	Added sections names	2 years ago
masserstein	Moved objective function calculation inside else	last year
visualization_package	Add markdown with documentation.	8 minutes ago
LICENSE	Prepare v1.0.0 release	2 years ago
README.md	Corrected url	2 years ago
setup.py	Prepare v1.0.0 release	2 years ago

Branch: visualization_package



Name	Last commit message	Last commit da...
..		
documentation.md	Add markdown with documentation.	9 minutes ago
utils.py	Add one function retrieving transport plan and one which shifts and s...	yesterday
visualization_examples.ipynb	Update the notebook.	7 hours ago
visualization_module.py	Add function to visualize stacked spectra.	yesterday

Results

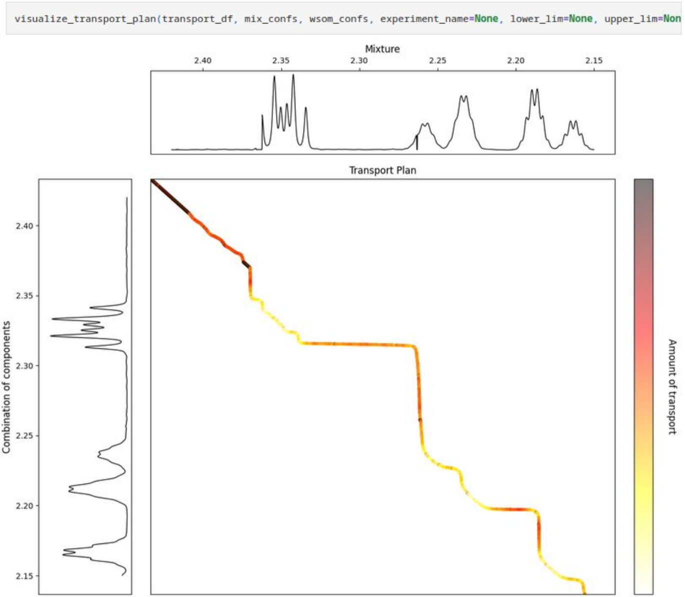
The visualization package now supports the following core features:

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

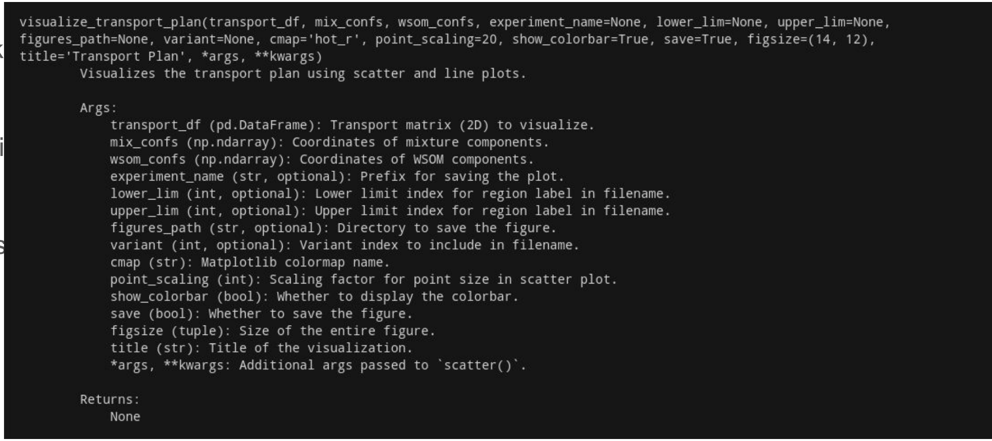
Results

The visualization package now support

- **Transport Plan Matrix**
  - Visualizes how mass (in this case ombination of components in the spectral domain, optionally scaled logarith
- Distance Distribution Plot
  - Generates a histogram showing h stances (in ppm). This helps assess the kappas required for decompositio ɔr need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconst ısing fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stack interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as noi decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .confs



```
visualize_transport_plan(transport_df, mix_confs, wsom_confs, experiment_name=None, lower_lim=None, upper_lim=Non
```



```
visualize_transport_plan(transport_df, mix_confs, wsom_confs, experiment_name=None, lower_lim=None, upper_lim=None,
figures_path=None, variant=None, cmap='hot_r', point_scaling=20, show_colorbar=True, save=True, figsize=(14, 12),
title='Transport Plan', *args, **kwargs)
        Visualizes the transport plan using scatter and line plots.

        Args:
            transport_df (pd.DataFrame): Transport matrix (2D) to visualize.
            mix_confs (np.ndarray): Coordinates of mixture components.
            wsom_confs (np.ndarray): Coordinates of WSOM components.
            experiment_name (str, optional): Prefix for saving the plot.
            lower_lim (int, optional): Lower limit index for region label in filename.
            upper_lim (int, optional): Upper limit index for region label in filename.
            figures_path (str, optional): Directory to save the figure.
            variant (int, optional): Variant index to include in filename.
            cmap (str): Matplotlib colormap name.
            point_scaling (int): Scaling factor for point size in scatter plot.
            show_colorbar (bool): Whether to display the colorbar.
            save (bool): Whether to save the figure.
            figsize (tuple): Size of the entire figure.
            title (str): Title of the visualization.
            *args, **kwargs: Additional args passed to `scatter()`.

        Returns:
            None
```

Results

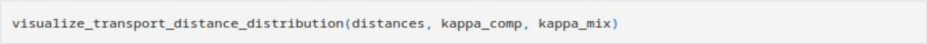
The visualization package now supports the following core features:

- Transport Plan Matrix
    - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
    - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
    - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
    - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- Noise Visualization
    - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- Integration with Internal Data Models
    - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

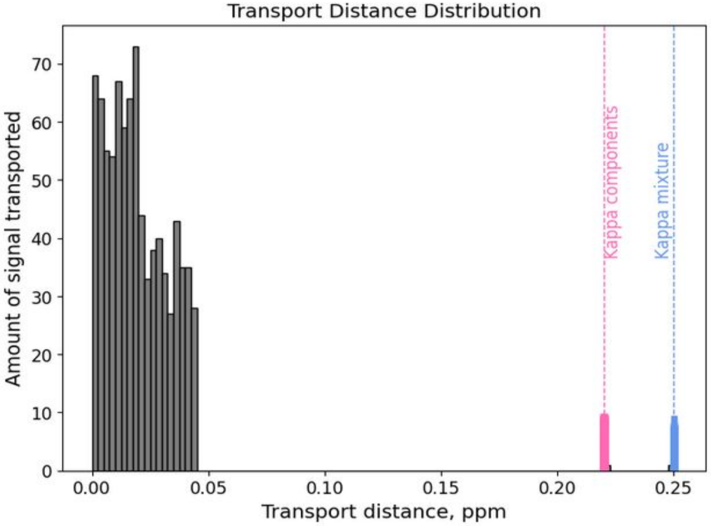
Results

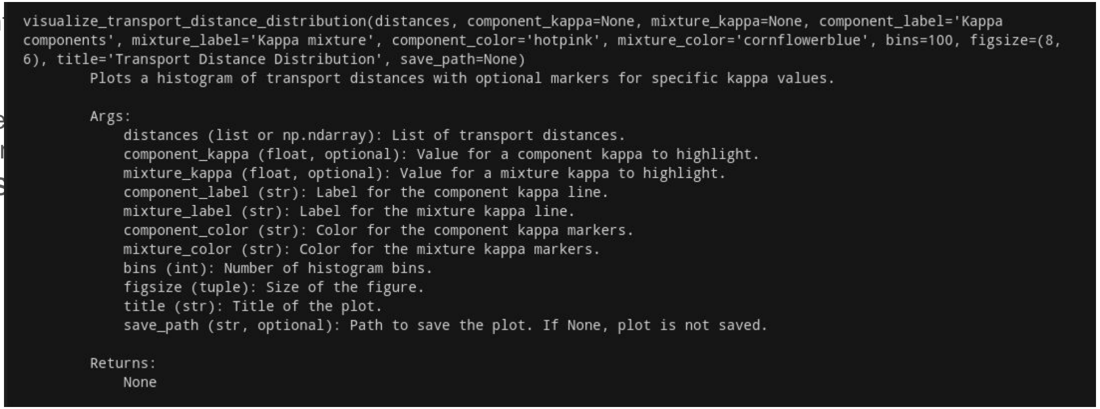
The visualization package now supports the following

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transp domain, optionally scaled logarithmically for clarity.
- **Distance Distribution Plot**
  - Generates a histogram showing how much spectral r kappas required for decomposition, and whether the
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighte nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., componen interest.
- Noise Visualization
  - Highlights what parts of the spectra we decomposition quality and residual sig
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum



```
visualize_transport_distance_distribution(distances, kappa_comp, kappa_mix)
```





```
visualize_transport_distance_distribution(distances, component_kappa=None, mixture_kappa=None, component_label='Kappa
components', mixture_label='Kappa mixture', component_color='hotpink', mixture_color='cornflowerblue', bins=100, figsize=(8,
6), title='Transport Distance Distribution', save_path=None)
        Plots a histogram of transport distances with optional markers for specific kappa values.

        Args:
            distances (list or np.ndarray): List of transport distances.
            component_kappa (float, optional): Value for a component kappa to highlight.
            mixture_kappa (float, optional): Value for a mixture kappa to highlight.
            component_label (str): Label for the component kappa line.
            mixture_label (str): Label for the mixture kappa line.
            component_color (str): Color for the component kappa markers.
            mixture_color (str): Color for the mixture kappa markers.
            bins (int): Number of histogram bins.
            figsize (tuple): Size of the figure.
            title (str): Title of the plot.
            save_path (str, optional): Path to save the plot. If None, plot is not saved.

        Returns:
            None
```

Results

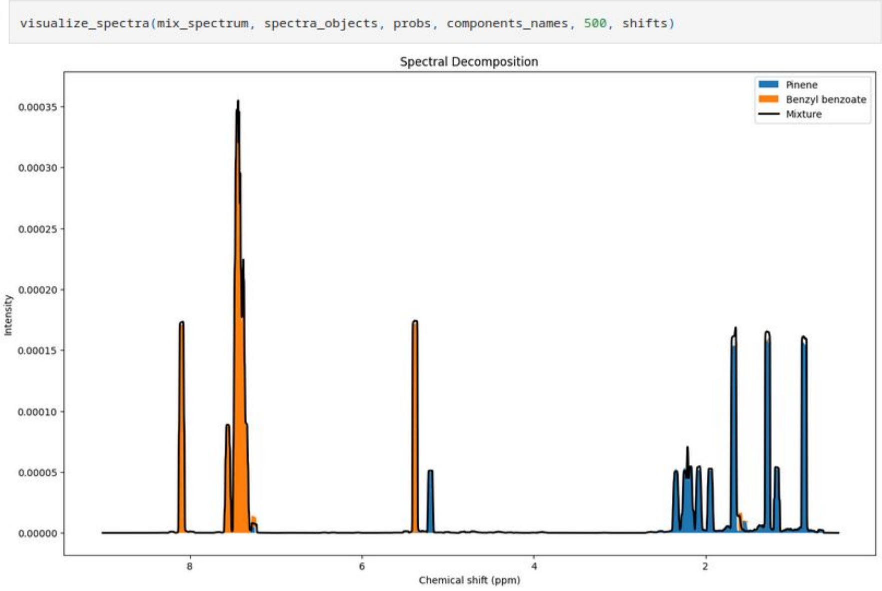
The visualization package now supports the following core features:

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

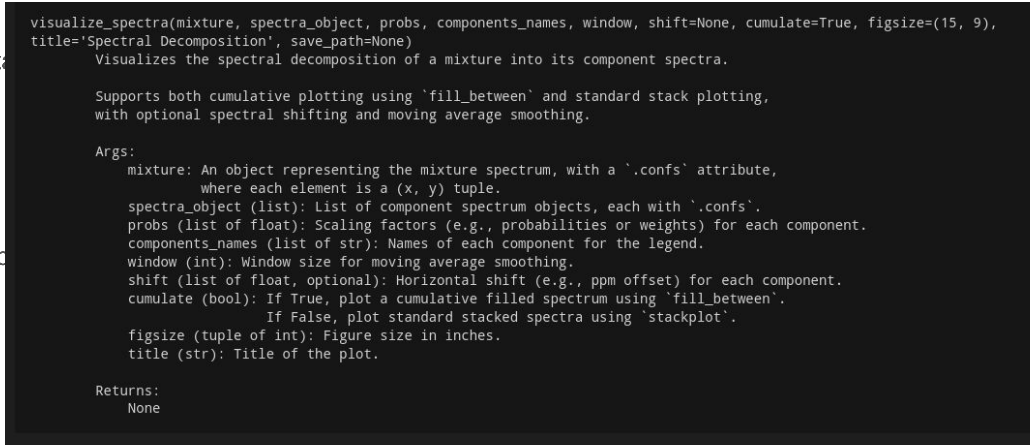
Results

The visualization package now supports the following core

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported
    domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass
    kappas required for decomposition, and whether the comp
- **Cumulative Decomposition Plot**
  - Displays how a mixture is reconstructed from weighted, sh
    nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) sta
    interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as
    decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .co



```
visualize_spectra(mix_spectrum, spectra_objects, probs, components_names, 500, shifts)
```



```
visualize_spectra(mixture, spectra_object, probs, components_names, window, shift=None, cumulate=True, figsize=(15, 9),
title='Spectral Decomposition', save_path=None)
    Visualizes the spectral decomposition of a mixture into its component spectra.

    Supports both cumulative plotting using `fill_between` and standard stack plotting,
    with optional spectral shifting and moving average smoothing.

    Args:
        mixture: An object representing the mixture spectrum, with a `.confs` attribute,
                 where each element is a (x, y) tuple.
        spectra_object (list): List of component spectrum objects, each with `.confs`.
        probs (list of float): Scaling factors (e.g., probabilities or weights) for each component.
        components_names (list of str): Names of each component for the legend.
        window (int): Window size for moving average smoothing.
        shift (list of float, optional): Horizontal shift (e.g., ppm offset) for each component.
        cumulate (bool): If True, plot a cumulative filled spectrum using `fill_between`.
                         If False, plot standard stacked spectra using `stackplot`.
        figsize (tuple of int): Figure size in inches.
        title (str): Title of the plot.

    Returns:
        None
```

Results

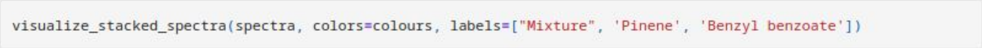
The visualization package now supports the following core features:

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

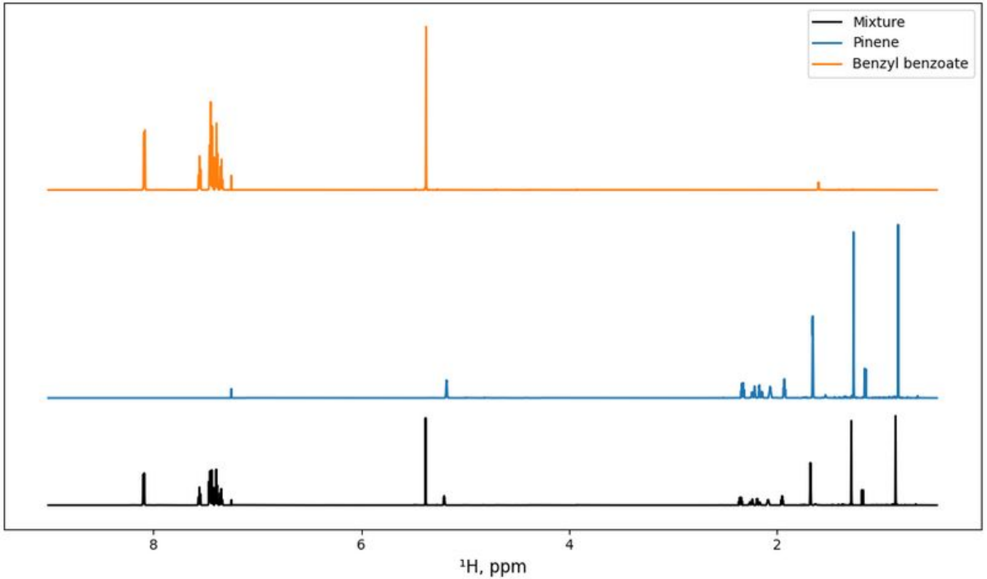
Results

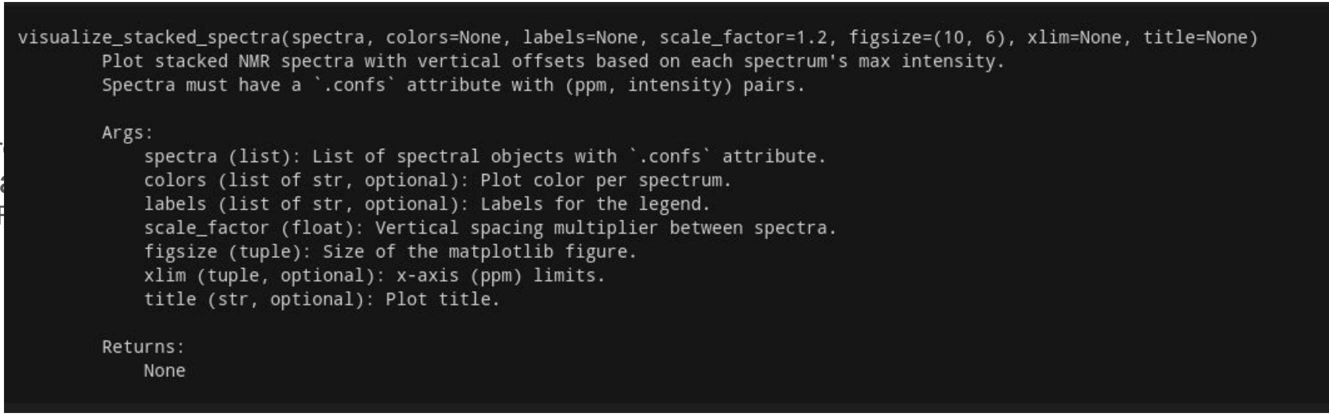
The visualization package now supports the fo

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) ectral domain, optionally scaled logarithmically fo
- Distance Distribution Plot
  - Generates a histogram showing how much the kappas required for decomposition, and wh
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed fro itive nature of the decomposition.
- **Stacked Spectra Plotting**
  - Plots multiple spectra (e.g., interest.
- Noise Visualization
  - Highlights what parts of the decomposition quality and r
- Integration with Internal Data
  - Works seamlessly with NMR



```
visualize_stacked_spectra(spectra, colors=colours, labels=["Mixture", 'Pinene', 'Benzyl benzoate'])
```





```
visualize_stacked_spectra(spectra, colors=None, labels=None, scale_factor=1.2, figsize=(10, 6), xlim=None, title=None)
        Plot stacked NMR spectra with vertical offsets based on each spectrum's max intensity.
        Spectra must have a `.confs` attribute with (ppm, intensity) pairs.

        Args:
            spectra (list): List of spectral objects with `.confs` attribute.
            colors (list of str, optional): Plot color per spectrum.
            labels (list of str, optional): Labels for the legend.
            scale_factor (float): Vertical spacing multiplier between spectra.
            figsize (tuple): Size of the matplotlib figure.
            xlim (tuple, optional): x-axis (ppm) limits.
            title (str, optional): Plot title.

        Returns:
            None
```

Results

The visualization package now supports the following core features:

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- **Noise Visualization**
  - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- Integration with Internal Data Models
  - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

Results

The visualization package now supports the following core features:

- Transport Plan Matrix
  - Visualizes how mass (in this case intensity) is transported between the mixture and optimal combination of components in the spectral domain, optionally scaled logarithmically for clarity.
- Distance Distribution Plot
  - Generates a histogram showing how much spectral mass was transported across different distances (in ppm). This helps assess the kappas required for decomposition, and whether the components align well with the mixture or need significant shifting.
- Cumulative Decomposition Plot
  - Displays how a mixture is reconstructed from weighted, shifted, and smoothed components using fill_between—revealing the additive nature of the decomposition.
- Stacked Spectra Plotting
  - Plots multiple spectra (e.g., components + mixture) stacked with proper vertical offsets, customizable labels, colors, and regions of interest.
- Noise Visualization
  - Highlights what parts of the spectra were removed as noise from both mixture and component sides. This helps evaluate the decomposition quality and residual signal.
- **Integration with Internal Data Models**
  - Works seamlessly with NMRSpectrum objects and .confs structure used internally in Magnetstein.

Conclusions & possible directions for the future

Conclusions

The visualization module has proven essential for understanding and communicating the results of spectral decomposition in Magnetstein. It allows users to:

- Validate decomposition quality visually.
- Identify noise patterns and transport mismatches.
- Generate high-quality figures for publication and presentations.

Possible directions for the future

- Interactive Plots (e.g., via Plotly or Bokeh) for zooming, highlighting transport lines, or toggling components on/off.

Thank you for your attention!