***vd-HSQC***

Virtual decoupling of 2D HSQC NMR experiments - version 1.1

<https://github.com/NMRTeamTBI/VirtualDecoupling>

**User Manual**

Version 1.1

Authors: Cyril Charlier (charlier@insa-toulouse.fr)

Guy Lippens (glippens@insa-toulouse.fr)

NMR Team (EAD17)

Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRA, INSA, Toulouse, France

**Table of contents**

[I. Introduction 1](#_Toc69131897)

[1. Software description 1](#_Toc69131898)

[2. Licensing 1](#_Toc69131899)

[II. Installation 1](#_Toc69131900)

[1. Dependencies 1](#_Toc69131901)

[2. Installation 1](#_Toc69131902)

[3. Test of installation 1](#_Toc69131903)

[III. Methods 2](#_Toc69131904)

[1. Data preparation 2](#_Toc69131905)

[2. Peak Picking 2](#_Toc69131906)

[3. Reconstruction (Virtual decoupling) 2](#_Toc69131907)

[IV. User Manual 3](#_Toc69131908)

[1. Input data 3](#_Toc69131909)

[2. Vd-hsqc usage 3](#_Toc69131910)

[2.1. Commands and options 3](#_Toc69131911)

[2.2. Peak Selection and Clustering 4](#_Toc69131912)

[2.3. Output data 4](#_Toc69131913)

[3. Error and warning messages 5](#_Toc69131914)

[V. License for vd-hsqc software 6](#_Toc69131915)

# Introduction

## Software description

vd-HSQC is a scientific software designed to provide a virtually decoupled HSQC spectrum from the combination of InPhase and AntiPhase spectra.

## Licensing

The original version of vd-HSQC was developed in the NMR team in the TBI, Toulouse, France.

The software is licensed under the Educational Community License, Version 2.0 (the “License”); you may not use this software and documentation except in compliance with the License. You may obtain a copy of the License [here](#_License_for_IsoCor) or at <http://www.opensource.org/licenses/ECL-2.0>.

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

# Installation

The software was developed on Linux and can be used on Windows, Linux or MacOS platforms. To use vd-HSQC, you’ll need some dependencies listed below.

## Dependencies

To use vd-hsqc you must have TopSpin (4.0 or higher) installed.

Vd-hsqc also requires Python (2.7+, 3.0 or higher) and modules:

* numpy
* matplotlib
* pandas
* nmrglue
* tqdm

If you are not used to install system wide environments like Python, ask some help from your local computer service. We don’t provide support for installation.

## Installation

Unpack the content of *VirtualDecoupling-main.zip* somewhere on your disk, and copy both the file *VD\_HSQC.py* and *VD\_HSQC\_int.py* in the TopSpin Python programs directory (by default: <TopSpin installation directory>/exp/stan/nmr/py/user).

## Test of installation

* To set the system’s Python interpreter path and the path for two python scripts, run the following command in TopSpin:

**vd-HSQC –opt**

The inputs entered by the user will be saved in “vd\_hsqc\_path.txt” and will be used a default. If the system’s Python interpreter is in the path, the python path is “python”

* To check that all required modules are correctly installed, run the following command in TopSpin:

**vd-HSQC --test**

If a required module is not correctly installed a message will pop-up and it’s up to the user to install the required packages.

# Methods

The vd-hsqc package contains two scripts: *VD\_HSQC.py* and *VD\_HSQC\_int.py*. While the first one contains all the interaction with TopSpin, the second one is used for the peak picking, peak selection and peak shifting and will finally save the matrix into the TopSpin format.

## Data preparation

Both in-phase and anti-phase dataset must be placed in the same folder within TopSpin and must be processed prior running the script.

## Peak Picking

The peak picking is performed on the *procno* 3/ and 4/ using nmrglue package. While combining the two peak lists multiple options are possible:

* Two signals are picked on the same row, one positive and one negative, then the right one is automatically shifted to the central position
* Two signals are picked on the same row but both of them have the same sign, then the user will be asked to select or not the signals for reconstruction
* More than two signals are picked on the same row then the user will be asked to select and to cluster the signals pairwise prior to reconstruction

## Reconstruction (Virtual decoupling)

The virtual decoupling is performed by defining a box (for which the number of points in X and Y dimensions are defined by user) around the right component of the doublet. This box is then shifted by the mean value of the two positions (right and left components) and placed into an empty data matrix. The reconstructed is then saved into the *procno* 5/.

# User Manual

## Input data

* vd-HSQC requires as input a set two 2D spectra acquired using InPhase and AntiPhase conditions which must be located in the same folder. Processed 2D spectra must be in the first *procno* of each *expno*.
* Spectra must be pre-processed (i.e. Fourier-transformed, phased, baseline-corrected and aligned) in TopSpin before running vd-HSQC.

*Example: A set of in-phase and anti-phase spectra are available in the git within “Test\_Data”. This folder should be placed in the TopSpin data directory.*

## Vd-hsqc usage

## Commands and options

vd-HSQC can be run using the following TopSpin command while being in the directory containing both 2D experiments:

**vd-HSQC**

which will prompt a window to be filled by the user:

Graphical user interface, table

Description automatically generated with medium confidence

**Data Directory** Will be automatically filled by the current directory of the data open on

the screen

**InPhase ExpNo** *expno* of the InPhase experiment

**AntiPhase ExpNo** *expno* of the AntiPhase experiment

**New ExpNo** *expno* created that will be used to the save the virtually decoupled spectra

**First ProcNo** vd-hsqc will save 5 *procnos*starting from this number

**Chemical shifts in ppm** Chemical shifts values used to define the region of the spectra extracted for the reconstruction. Leaving “0;0;0;0” will reconstruct the complete spectra

**Comment (Title)** Any comments that the user wants to add to the title of the newly created expnos

**nc\_proc** nc\_proc values used in TopSpin for plotting purposes

**Detection threshold** Values used for the detection of peaks

**Points Shifting** Number of points to define the box to shift

## Peak Selection and Clustering

In the case of more than two signals detected in the same row during the peak picking procedure the user will be asked to select and cluster them using the following window:

Graphical user interface

Description automatically generated with medium confidence

If anything, other than “yes” is indicated in the “Peaks\_Selection” then the Peak will be discarded. Then the user can cluster peaks by indicating the same number in the “Cluster” row.

## Output data

The vd-hsqc script will create a new *expno* containing the data of interest (either the full spectra or a user defined region, see Section IV.2) with the *procno* as followed:

1/ InPhase spectra (16k points)

2/ AntiPhase spectra (16k points)

3/ AntiPhase spectra (1k points) with only one component of the doublet shown as positive

4/ AntiPhase spectra (1k points) with the second component of the doublet shown as positive

5/ Virtually decoupled spectra spectra (16k points)

The *procno* 3/ and 4/ are automatically processed using effective TD (*tdeff* in TopSpin) of 1024 points and then are artificially modified by setting the negative values in the data matrix to observe only the positive either on the left or right component of the doublet.

## Error and warning messages

Error messages are explicit. After correcting the problem, rerun vd-hsqc.

# License for vd-hsqc software

Educational Community License, Version 2.0 [OSI Approved License]

Educational Community License Version 2.0, April 2007

The Educational Community License version 2.0 ("ECL") consists of the Apache 2.0 license, modified to change the scope of the patent grant in section 3 to be specific to the needs of the education communities using this license. The original Apache 2.0 license can be found at: http://www.apache.org/licenses/LICENSE-2.0

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

"License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.

"Licensor" shall mean the copyright owner or entity authorized by the copyright owner that is granting the License.

"Legal Entity" shall mean the union of the acting entity and all other entities that control, are controlled by, or are under common control with that entity. For the purposes of this definition, "control" means (i) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity

"You" (or "Your") shall mean an individual or Legal Entity exercising permissions granted by this License.

"Source" form shall mean the preferred form for making modifications, including but not limited to software source code, documentation source, and configuration files.

"Object" form shall mean any form resulting from mechanical transformation or translation of a Source form, including but not limited to compiled object code, generated documentation, and conversions to other media types.

"Work" shall mean the work of authorship, whether in Source or Object form, made available under the License, as indicated by a copyright notice that is included in or attached to the work (an example is provided in the Appendix below).

"Derivative Works" shall mean any work, whether in Source or Object form, that is based on (or derived from) the Work and for which the editorial revisions, annotations, elaborations, or other modifications represent, as a whole, an original work of authorship. For the purposes of this License, Derivative Works shall not include works that remain separable from, or merely link (or bind by name) to the interfaces of, the Work and Derivative Works thereof.

"Contribution" shall mean any work of authorship, including the original version of the Work and any modifications or additions to that Work or Derivative Works thereof, that is intentionally submitted to Licensor for inclusion in the Work by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, "submitted" means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Work, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as "Not a Contribution."

"Contributor" shall mean Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Work.

2. Grant of Copyright License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form.

3. Grant of Patent License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this section) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Work to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Work or a Contribution incorporated within the Work constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for that Work shall terminate as of the date such litigation is filed. Any patent license granted hereby with respect to contributions by an individual employed by an institution or organization is limited to patent claims where the individual that is the author of the Work is also the inventor of the patent claims licensed, and where the organization or institution has the right to grant such license under applicable grant and research funding agreements. No other express or implied licenses are granted.

4. Redistribution.

You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

You must give any other recipients of the Work or Derivative Works a copy of this License; and

You must cause any modified files to carry prominent notices stating that You changed the files; and

You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and

If the Work includes a "NOTICE" text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works, in at least one of the following places: within a NOTICE text file distributed as part of the Derivative Works; within the Source form or documentation, if provided along with the Derivative Works; or, within a display generated by the Derivative Works, if and wherever such third-party notices normally appear. The contents of the NOTICE file are for informational purposes only and do not modify the License. You may add Your own attribution notices within Derivative Works that You distribute, alongside or as an addendum to the NOTICE text from the Work, provided that such additional attribution notices cannot be construed as modifying the License.

You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions for use, reproduction, or distribution of Your modifications, or for any such Derivative Works as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.

5. Submission of Contributions. Unless You explicitly state otherwise, any Contribution intentionally submitted for inclusion in the Work by You to the Licensor shall be under the terms and conditions of this License, without any additional terms or conditions. Notwithstanding the above, nothing herein shall supersede or modify the terms of any separate license agreement you may have executed with Licensor regarding such Contributions.

6. Trademarks. This License does not grant permission to use the trade names, trademarks, service marks, or product names of the Licensor, except as required for reasonable and customary use in describing the origin of the Work and reproducing the content of the NOTICE file.

7. Disclaimer of Warranty. Unless required by applicable law or agreed to in writing, Licensor provides the Work (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Work and assume any risks associated with Your exercise of permissions under this License.

8. Limitation of Liability. In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Work (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

9. Accepting Warranty or Additional Liability. While redistributing the Work of Derivative Works thereof, You may choose to offer, and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

END OF TERMS AND CONDITIONS

APPENDIX: How to apply the Educational Community License to your work

To apply the Educational Community License to your work, attach the following boilerplate notice, with the fields enclosed by brackets "[]" replaced with your own identifying information. (Don’t include the brackets!) The text should be enclosed in the appropriate comment syntax for the file format. We also recommend that a file or class name and description of purpose be included on the same "printed page" as the copyright notice for easier identification within third-party archives.

Copyright [yyyy] [name of copyright owner] Licensed under the Educational Community License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at:

http://www.osedu.org/licenses/ECL-2.0

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

Vd-hsqc (hereafter "the program") is a scientific software designed to correct MS data for natural abundance.

All its parts are freely distributed according to GPL2 terms http://www.gnu.org/licenses/gpl-2.0.txt (permission to execute, to copy, to modify and to redistribute; obligation to preserve the original copyright notice and to keep your modifications OpenSource only if you redistribute them alone or as part of vd-hsqc)

WARRANTY

BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION. COPYRIGHT 2017, INRA