

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ЭЛЕМЕНТООРГАНИЧЕСКИХ СОЕДИНЕНИЙ
им. А.Н. НЕСМЕЯНОВА
РОССИЙСКОЙ АКАДЕМИИ НАУК

ДОКЛАД ОБ ОСНОВНЫХ РЕЗУЛЬТАТАХ
НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ

**АНАЛИЗ ПРОЧНОСТИ ТОПОЛОГИЧЕСКОГО
СВЯЗЫВАНИЯ КАК МЕТОД ОЦЕНКИ ВКЛАДОВ В
ЭНЕРГИЮ ВЗАИМОДЕЙСТВИЯ КВАНТОВЫХ АТОМОВ**

1.4.4. – Физическая химия

Автор: Анисимов Алексей Альбертович
Научный руководитель: к.х.н. Ананьев Иван Вячеславович

Москва 2024

Концепция переносимости свойств атомов, функциональных групп и целых молекул позволяет отнести эти объекты к базовым строительным блокам химии. Связывающим звеном для таких блоков служат так называемые химические связи – взаимодействия, которые обуславливают устойчивость многоатомных систем, будь то молекулы или надмолекулярные ассоциаты. Однако с точки зрения квантовой механики выделение соответствующих взаимодействий, а также изучение их энергетических свойств – это достаточно серьезная проблема, которая не имеет однозначного решения и по сей день. Так, локализация связывающих взаимодействий может производиться посредством определенных манипуляций с орбиталями системы – получения различных локализованных орбиталей по типу натуральных орбиталей связи (рис. 1А); путем поиска областей локализации электронной плотности – например, посредством выявления максимумов функции ELF (рис. 1Б), а также полагаясь на анализ функции электронной плотности в рамках теории «Атомы в молекулах» (рис. 1В).

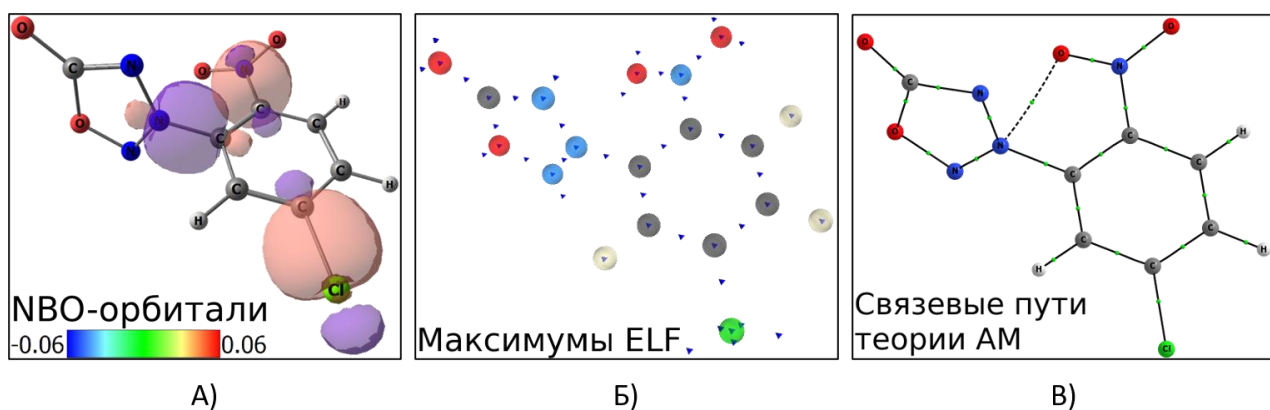


Рисунок 1. Различные способы локализации связывающих взаимодействий.

Именно последний вариант хотелось бы выделить среди остальных. Комбинация теории Р. Бейдера «Атомы в молекулах» (далее – теория АМ) и основанного на ней подхода Взаимодействующие Квантовые Атомы позволяют как получить граф связности, привычный для химика, так и провести анализ энергетических свойств всех взаимодействий без привлечения орбитальной картины. Для целей данной работы необходимо кратко пояснить, каким образом это достигается.

Теория Р. Бейдера построена на анализе двух полей: скалярного – электронной плотности и векторного – градиента электронной плотности. Анализируя их, можно однозначно выделить в реальном пространстве следующие структуры (рис. 2):

- Во-первых, это максимумы электронной плотности, которые чаще всего отвечают за положение ядер (рис. 2А).

- Во-вторых, это межатомные поверхности нулевого потока электронной плотности, которые служат границами атомных бассейнов, и, формально, позволяют выделять атомы в реальном трехмерном пространстве (рис. 2Б).
- В-третьих, это критические точки связи и связевые пути – линии максимальной электронной плотности между двумя максимумами, которые служат дескриптором образования связи с топологической точки зрения (далее я буду называть их просто термином «топологическая связь») (рис. 2В).

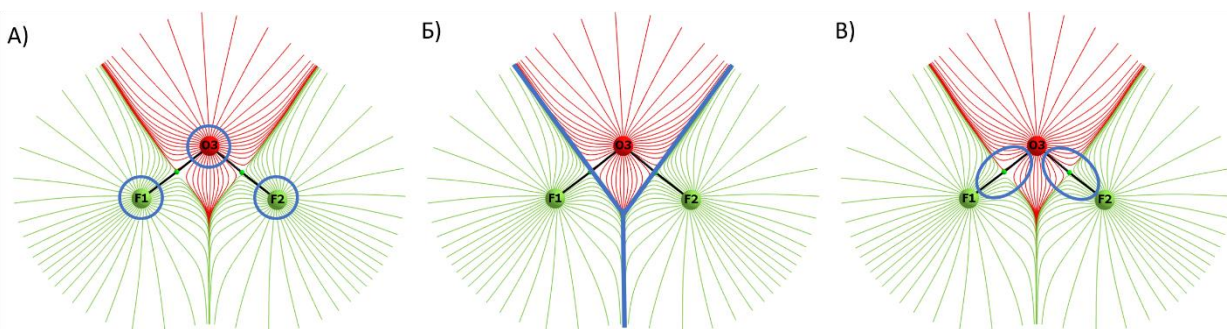


Рисунок 2. Основные топологические объекты, определяемые в рамках теории АМ.

Основываясь на разбиении системы в теории АМ, группа профессора М. А. Пендаса предложила следующий вариант разложения общей энергии системы на внутри- и межатомные вклады (рис. 3). Все вклады получаются путем интегрирования соответствующих операторов либо по отдельным бассейнам (для получения внутриатомных вкладов, для этого необходим трехмерный интеграл), либо же для каждой пары бассейнов (для получения межатомных вкладов, для этого необходим шестимерный интеграл).

$$E = \sum_A E_{net}^A + \frac{1}{2} \sum_A \sum_{B \neq A} E_{int}^{AB}$$

$T^A + V_{cl}^{AA} + V_{xc}^{AA}$

внутриатомные

$V_{cl}^{AB} + V_{xc}^{AB}$

межатомные

Рисунок 3. Получаемые в рамках подхода Взаимодействующие Квантовые Атомы энергетические компоненты.

Однако, несмотря на логичность, простоту и прямолинейность такого подхода, он же является и источником определенных проблем. Большая часть этих проблем связана с «неклассической» обменно-корреляционной межатомной энергией электрон-электронных

взаимодействий $V_{ee,xc}^{AB} = \iint_{\Omega\Omega'} \frac{\rho_2^{xc}(1',1,2',2)}{r_{12}} dx_1 x_2$. Этот вклад получается в результате ресурсозатратного шестимерного интегрирования, а также требует для корректного вычисления матрицу плотности второго порядка – объёмного объекта, получение которого затруднено как для пост-ХФ методов, так и для методов теории функционала плотности. По понятным причинам, сложность растёт степенным образом с увеличением системы, что ограничивает применимость подхода Взаимодействующие Квантовые Атомы для больших систем, в частности, значительно затрудняет его использование для анализа невалентных межмолекулярных взаимодействий.

При этом, обозначенные энергетические компоненты имеют непосредственное отношение к энергетической характеристике связи – энергии связывания (рис. 4).

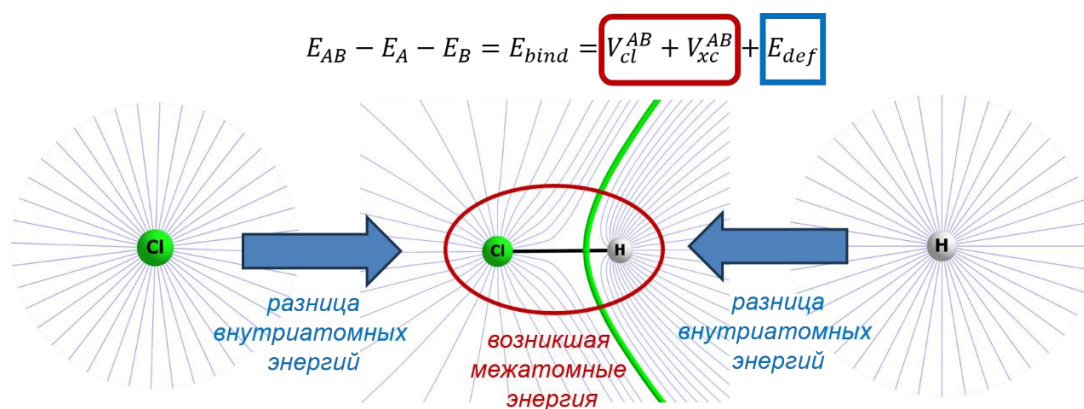


Рисунок 4. Связь энергетических компонент из подхода Взаимодействующие Квантовые Атомы и энергии связывания.

Исходя из наиболее банального – супермолекулярного подхода, энергия связывания E_{bind} определяется как разность энергий связанного комплекса и составляющих его невзаимодействующих подсистем. Межатомная кулоновская V_{cl}^{AB} и обменно-корреляционная V_{xc}^{AB} энергии характеризуют образовавшуюся связь, являясь естественными метриками степени ионности и ковалентности связи. В свою очередь, внутрифрагментные энергетические компоненты позволяют оценить так называемую энергию деформации E_{def} путем сопоставления начального и конечного состояния системы. Этот вклад описывает энергию изменения спинового состояния, переноса заряда, и иные искажения электронной структуры бассейнов, а также, в случае многоатомных фрагментов – изменение во взаимодействиях внутри этих фрагментов. Тем не менее, основным минусом в получении энергии деформации является необходимость расчета начального состояния в дополнении к интересующей нас системе, что также крайне негативно влияет на ресурсозатратность, особенно при исследовании невалентных межмолекулярных взаимодействий. Вдобавок, следует отметить, что нет очевидной и однозначно доказанной

взаимосвязи между озвученными энергетическими вкладами и явлением топологического связывания.

Как следствие, ввиду наличия ряда проблем для комбинации теории АМ и подхода Взаимодействующие Квантовые Атомы, **целью** данной работы являлось развитие представлений о закономерностях между явлением топологического связывания и соответствующими вкладами в энергию межатомного взаимодействия, определенными в рамках подхода Взаимодействующих Квантовых Атомов.

Для достижения обозначенной цели и в работе решались следующие **задачи**:

- 1) Изучить явление топологического связывания с точки зрения внутренней структуры квантовых атомов;
- 2) Определить метрику прочности топологического связывания для сравнения с определёнными в рамках подхода Взаимодействующие Квантовые Атомы компонентами энергии межатомного взаимодействия;
- 3) Разработать метод, позволяющий получать компоненты межатомной энергии взаимодействия, определенные в рамках подхода Взаимодействующие Квантовые Атомы, для межмолекулярных взаимодействий, минуя основные вычислительные проблемы оригинального подхода;
- 4) Разработать новый алгоритм генерации межатомных поверхностей, позволяющий корректно определить топологический атом даже в сложных случаях с пологим распределением электронной плотности.

Для решения первой задачи нами было решено воспользоваться выражением (1), которое позволяет оценить изменения во внутренней структуре квантового атома при образовании топологического связывания:

$$\{ \langle [\hat{H}, r_{\Omega}^2] \rangle_{\Omega} + \langle [\hat{H}, r_{\Omega}^2] \rangle_{\Omega}^* \} = 3 \int_{\Omega} d\tau(\mathbf{r}_{\Omega}) \rho(\mathbf{r}_{\Omega}) + \int_{\Omega} d\tau(\mathbf{r}_{\Omega}) \nabla \rho(\mathbf{r}_{\Omega}) \cdot \mathbf{r}_{\Omega} = \oint_{IAS} dS \rho(\mathbf{r}_{\Omega}) \mathbf{r}_{\Omega} \cdot \mathbf{n}(\mathbf{r}_{\Omega}) \quad (1)$$

По сути данный коммутатор измеряет искажение векторного поля градиента электронной плотности относительно радиального состояния, характерного для свободного атома. Для свободного атома интеграл по поверхности по понятным причинам равен нулю – его поверхность находится на бесконечности, где электронная плотность стремится к нулю. В таком случае первый член равен заселенности атомного бассейна, а второй должен быть равен минус трем заселенностям (рис. 5, слева). В случае же связанного атома поверхностный интеграл уже не будет равен нулю. Первый член все также будет давать нам заселенность атомного бассейна, а второй будет являться мерой искажения векторного поля градиента электронной плотности (рис. 5, справа).

Свободный атом: $3N(\Omega)$ $-3N(\Omega)$ $= 0$
 Связанный атом: $3N'(\Omega)$ *мера искажения* $\neq 0$

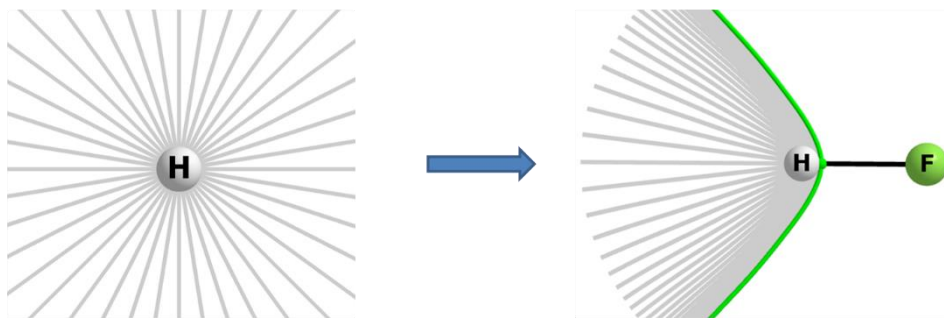


Рисунок 5. Иллюстрация к изменениям в системе при образовании топологического связывания, которые описывает выражение (1).

При этом мы, формально, рассматриваем происходящие внутри атомных бассейнов изменения при образовании топологического связывания, которые, очевидно, должны быть тесно связаны с явлением электронной делокализации. Однако предложенный в выражении (1) коммутатор усредняется по атомному бассейну путем интегрирования – т.е. мы получаем только интегральную метрику. С целью более тщательного изучения перераспределения логично попытаться перейти к скалярному полю, определенному во всех точках бассейна, что и было сделано путем введения функции (2), далее именуемой как «делокализационная плотность».

$$f_{\delta}(\mathbf{r}_{\Omega}) = \rho(\mathbf{r}) + \frac{1}{3} \nabla \rho(\mathbf{r}) \cdot \mathbf{r}, \forall \mathbf{r} \in \Omega \quad (2)$$

Изучение распределения данной функции для различных систем позволило установить, что она может использоваться как альтернатива других известных функций-метрик локализации электронов, позволяя анализировать радиальные особенности в атомах, выявлять неподелённые электронные пары и накопления в районе поверхностей. В качестве примера можно рассмотреть положения максимумов рассматриваемой функции и $-\Delta\rho(\mathbf{r})$ (рис. 6).

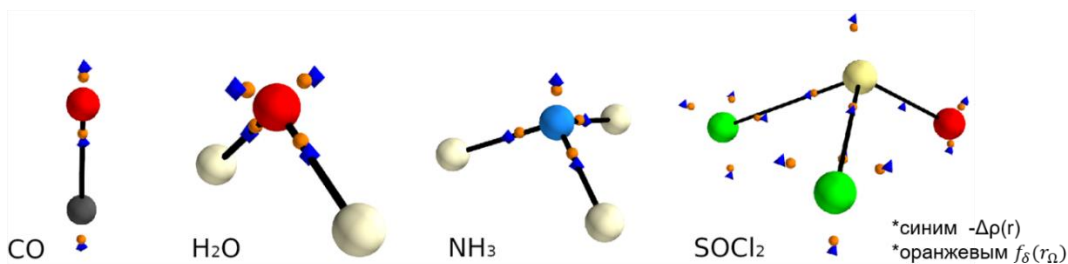


Рисунок 6. Положения максимумов функций $f_{\delta}(\mathbf{r}_{\Omega})$ и $-\Delta\rho(\mathbf{r})$.

Как вы можете заметить, положения практически совпадают. Итак, предложенный коммутатор в дифференциальной форме действительно может служить как функция делокализации электронов в бассейне.

Для решения второй задачи мы решили рассмотреть выражение (1) с точки зрения метрики топологического связывания. Логично, что для этого нам необходимо измерять искажение внутренней структуры не для одного, а для пары атомов, между которыми образуется эта самая топологическая связь. Поэтому нами было решено просто сложить два поверхностных интеграла и таким образом перейти выражению (2).

$$\oint_{IAS} dS \rho(r_{\Omega}) \mathbf{r}_{\Omega} \cdot \mathbf{n}(r_{\Omega}) + \oint_{IAS} dS \rho(r_{\Omega'}) \mathbf{r}_{\Omega'} \cdot \mathbf{n}(r_{\Omega'}) = \mathbf{R}_{\Omega\Omega'} \cdot \oint_{IAS} dS \rho(r_{\Omega}) \mathbf{n}(r_{\Omega}) \quad (2)$$

На основании предложенной метрики топологического связывания нами была разработана следующая параметрическая модель (3) с параметрами p_1 и p_2 .

$$\frac{k}{|\mathbf{R}_{\Omega\Omega'}|^{p_2}} \oint_{IAS} dS \rho(r_{\Omega})^{p_1} \mathbf{R}_{\Omega\Omega'} \cdot \mathbf{n}(r_{\Omega}) + b \quad (3)$$

Затем предложенная параметрическая модель (3) была использована в процедуре восстановления компонентов энергии связи, определенные в рамках подхода IQA. Мы решили начать с попытки восстановить именно межатомные обменно-корреляционный вклад ввиду того, что его получение является наиболее ресурсозатратным. Для первоначальной проверки мы воспользовались следующими наборами (см. рис. 7).

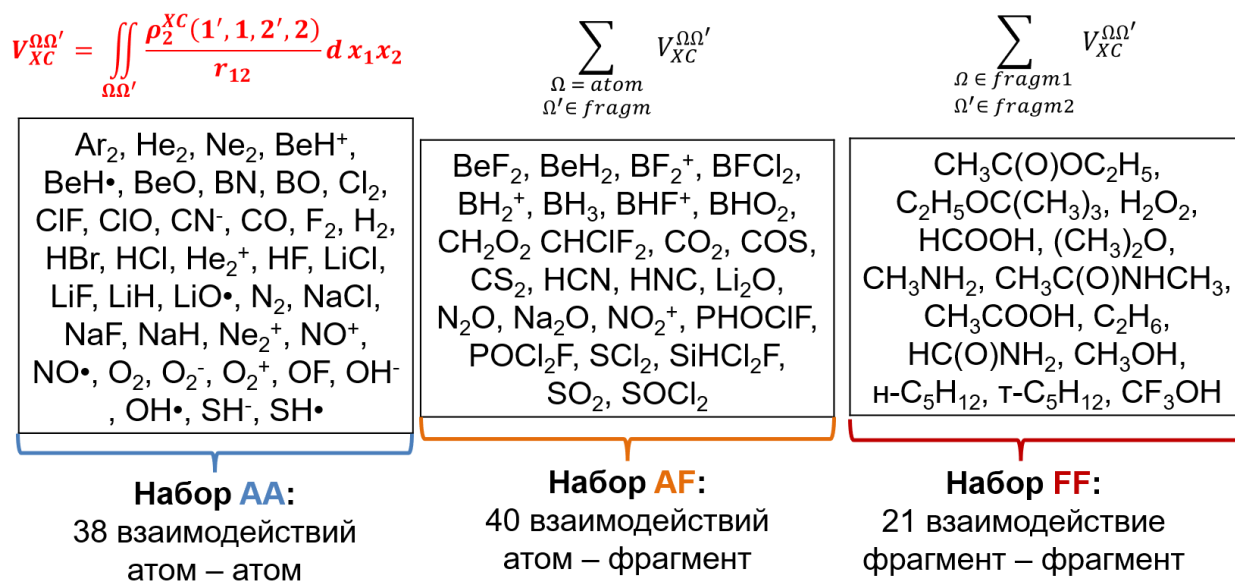


Рисунок 7. Рассматриваемые наборы AA, AF и FF, а также соответствующие обменно-корреляционные межатомные (или межфрагментные) энергии.

Набор **AA** состоит из 38 двухатомных систем, содержащем как невалентные взаимодействия, так и более прочные ковалентные связи и связи с ионным характером. Второй и третий наборы служат для проверки применимости модели к полноценным внутримолекулярным взаимодействиям между атомом и молекулярным фрагментом в случае набора **AF** и двумя молекулярными фрагментами в случае набора **FF**.

Максимизация коэффициента детерминации путем оптимизации параметров p_1 и p_2 методом BFGS для набора **AA** позволила получить отличные линейные регрессии с достаточно малыми средними абсолютными ошибками. Можно заметить, что межатомная обменно-корреляционная энергия очень неплохо согласуется с предложенным выражением для меры перестройки электронной плотности: R^2 выше 98 процентов во всех случаях (рис. 8). При этом известно, что увеличение уровня теории, от метода Хартри-Фока к методу связанных кластеров, приводит к более точному описанию обменно-корреляционных эффектов. В этом смысле и наша зависимость становится более статистически достоверной для более высокоуровневых методов.

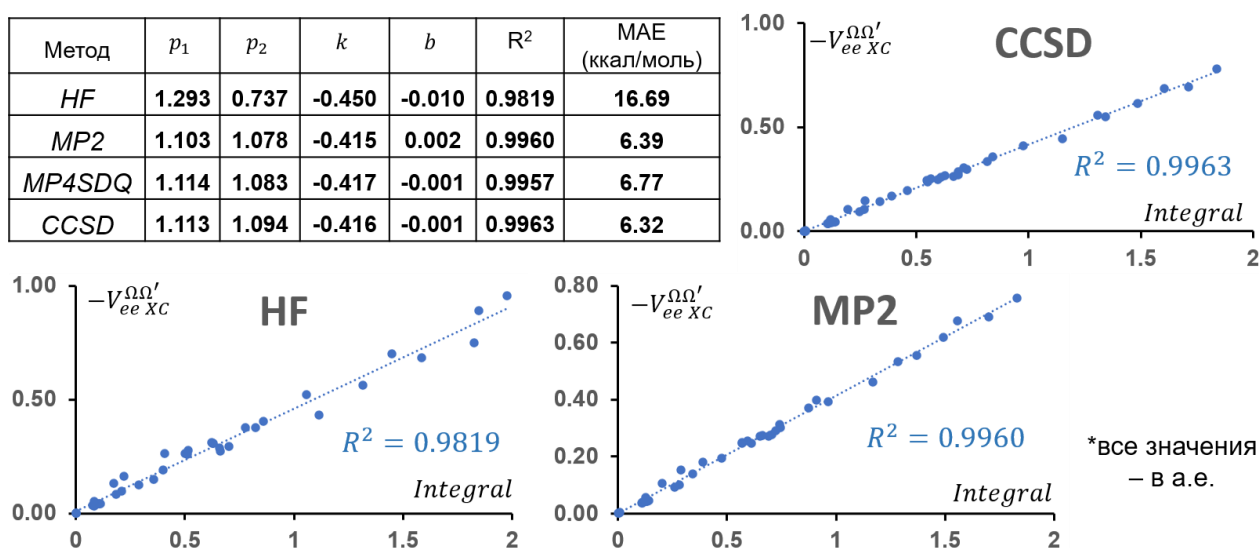


Рисунок 8. Полученные в результате уточнения параметров p_1 и p_2 тренды, а также итоговые коэффициенты линейных регрессий, коэффициенты детерминации и средние абсолютные ошибки для систем набора **AA**, рассчитанных различными методами.

Однако переходя к оптимизации модели для других двух наборов, полученные результаты достаточно посредственны, в особенности в случае восстановления взаимодействий между парой молекулярных фрагментов (рис. 9). Во-первых, коэффициенты детерминации достаточно низки, во-вторых, параметры и коэффициенты модели значительно отличаются от таковых для двухатомных систем.

Набор AF					
Метод	p_1	p_2	k	b	R^2
HF	1.318	0.942	-0.517	-0.003	0.9495
MP2	1.259	0.829	-0.418	-0.008	0.9878
CCSD	1.266	0.838	-0.418	-0.011	0.988

Набор FF					
Метод	p_1	p_2	k	b	R^2
HF	4.155	-6.552	1E-04	-0.244	0.3614
MP2	1.225	0.137	-0.242	0.078	0.9428
CCSD	1.223	0.134	-0.241	0.081	0.9117

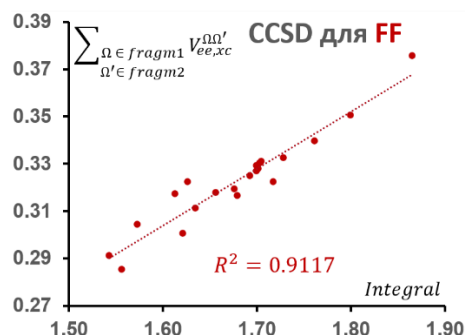


Рисунок 9. Полученные в результате уточнения параметров p_1 и p_2 тренды, а также итоговые коэффициенты линейных регрессий, коэффициенты детерминации и средние абсолютные ошибки для систем наборов **AF** и **FF**, рассчитанных различными методами.

В качестве объяснения подобного поведения модели нами было выдвинуто предположение, что она недоучитывает изменения в несвязанных атомных бассейнах взаимодействующих фрагментов, т.е. позволяет оценить только искажения для пары непосредственно связанных атомов. Несмотря на описанную проблему, можно заключить, что предложенная модель может с некоторыми условиями служить метрикой прочности топологического связывания.

По причине вышеуказанных ограничений предложенной модели было решено перейти к межмолекулярным невалентным взаимодействиям, которые по большей части не приводят к существенным искажениям электронной структуры взаимодействующих молекул. Наибольшее искажение в таком случае происходит как раз для атомов, между которыми образуется «топологическая связь», отвечающая за невалентное взаимодействие. Для анализа данного типа взаимодействий были выбраны два набора супрамолекулярных систем, в первый входят 53 бимолекулярных ассоциата, связанных различными по характеру межмолекулярными взаимодействиями, во второй – 13 вырезанных из конденсированной фазы кластеров (рис. 10).

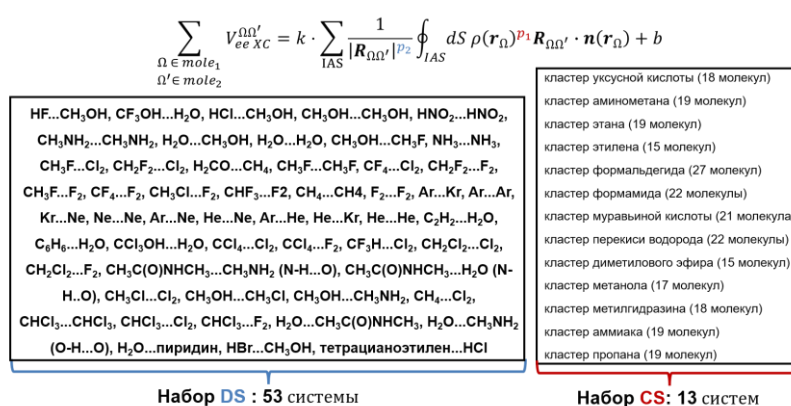


Рисунок 10. Системы, используемые для изучения взаимосвязи предложенной параметрической модели и межмолекулярной обменно-корреляционной энергии.

К сожалению, в случае межмолекулярных поверхностей возникли прикладные проблемы из-за несовершенства существующих подходов к описанию поверхностей нулевого потока. В качестве примера можно привести метод Рунге-Кутты второго порядка, используемый в программе Multiwfn с целью получения градиентных траекторий для описания поверхности. В случае невалентных взаимодействий нередко случается так, что эти траектории распределены по поверхности крайне неравномерно (рис. 11, слева), что крайне отрицательно сказывается на качестве ее описания и последующей генерации сетки интегрирования путем триангуляции. По этой причине нами был разработан и имплементирован в Multiwfn свой алгоритм, основанный на бисекции и позволяющий получать более равномерное распределение «скелетных лучей» (рис. 11, справа), что позволило повысить качество интегрирования.

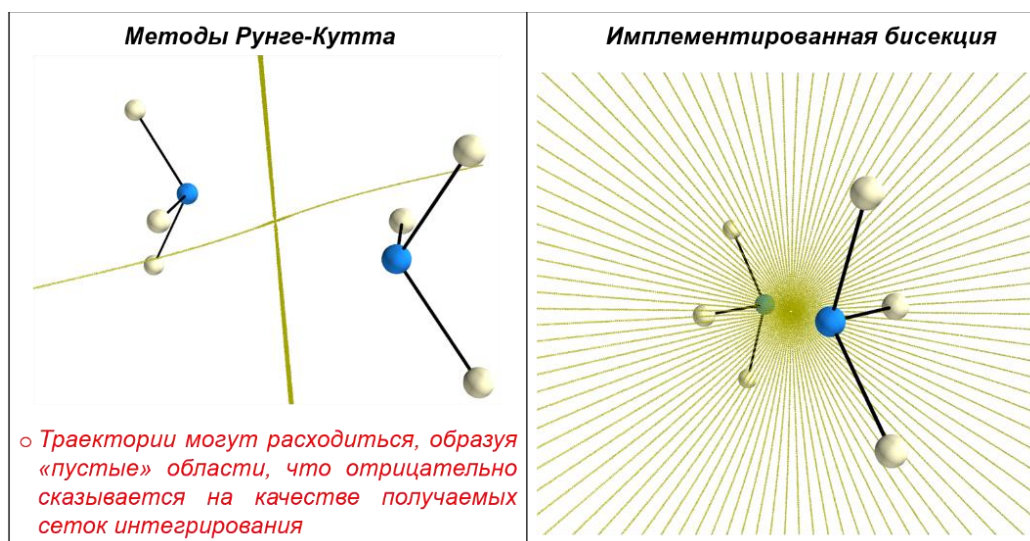


Рисунок 11. Иллюстрация к проблеме расхождения «скелетных лучей» для используемых методов описания поверхностей нулевого потока и лучи, продуцируемые разработанным в рамках данной работы алгоритмом.

В качестве наглядной иллюстрации хотелось бы привести сравнение интегралов, полученных с применением траекторий Рунге-Кутты и бисекции, с некоторым референсом. Для такого референса был выбран интеграл вида $\oint_{IAS}^{\square} dS\rho(\mathbf{r}_{\Omega})$, выдаваемый программой AIMAll. В случае двухатомных систем, что бисекция, что метод Рунге-Кутты фактически выдают идентичное AIMAll-у значение поверхностного интеграла (рис. 12, слева). Однако, когда мы переходим к кластерам, точность интегрирования с использованием «скелетных лучей» из метода Рунге-Кутты значительно ухудшается. Интегралы, полученные с применением бисекции же, в свою очередь, хорошо согласуются с выбранным референсом (рис. 12, справа). Итак, задачу разработать новый алгоритм генерации межатомных поверхностей можно считать успешно выполненной.

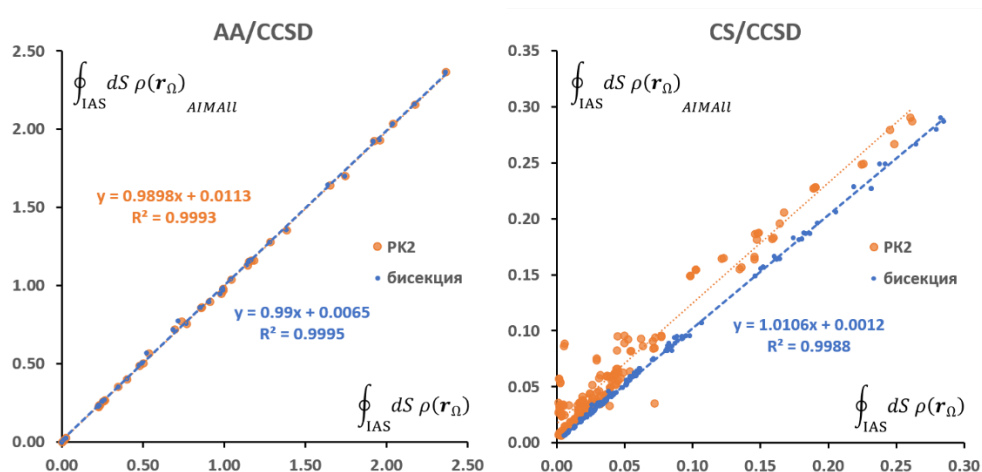


Рисунок 12. Сравнение интегралов, полученных с использованием метода Рунге-Кутты и бисекции, с интегралами из программы AIMAll.

Итак, перейдем к применению предложенной параметрической модели к восстановлению компонент энергии межмолекулярных взаимодействий. Энергия связывания состоит из энергии деформации и суммы кулоновского и обменно-корреляционного вкладов для образовавшегося взаимодействия (рис. 13).

$$E_{bind} = E_{in-situ} + E_{def}$$

$$\begin{matrix} \swarrow & \searrow \\ V_{cl}^{AB} & V_{xc}^{AB} \end{matrix}$$

Рисунок 13. Компоненты энергии связывания, определенные в рамках подхода Взаимодействующие Квантовые Атомы.

Для представленных наборов **DS** и **CS** в первую очередь была произведена оптимизация параметрической модели относительно обменно-корреляционной энергии $V_{xc}^{AB} \equiv V_{ee\ xc}^{AB}$ (рис 14А). Здесь необходимо обсудить переход от «тяжелых» пост-ХФ методов, достаточно редко применяющихся в анализе невалентных взаимодействий из-за стремительного роста ресурсозатратности при увеличении системы, к методам теории функционала плотности. С этой целью набор димеров (**DS**) был рассчитан посредством MP2 и PBE0. Результирующие регрессии (рис. 14Б) очень схожи, финальные параметры и коэффициенты крайне близки, что говорит о возможности использовать выбранный метод ТФП без потери качества восстановления V_{xc}^{AB} предлагаемой параметрической моделью. Оптимизации для набора с кластерами (**CS**) привела к линейной регрессии с отличным коэффициентом детерминации и средней абсолютной ошибкой, но несколько отличающимся параметрам и коэффициентами. С целью установить параметры, которые удовлетворительно работали бы и для бимолекулярных ассоциатов, и для кластеров молекул, была произведена оптимизация

для комбинированного набора **DS+CS**, в результате которого была получена линейная регрессия с низкой средней ошибкой и высоким коэффициентом детерминации (рис. 14B).

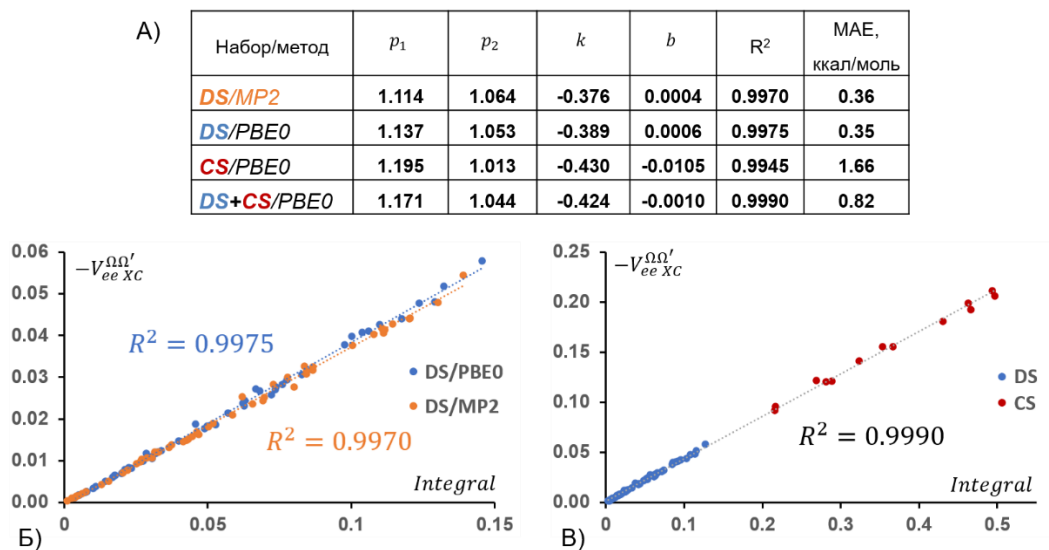


Рисунок 14. Результаты оптимизации параметрической модели для восстановления межмолекулярной обменно-корреляционной энергии.

Мы решили не ограничиваться только обменно-корреляционным членом и провели аналогичную оптимизацию модели уже относительно энергии деформации E_{def} , что привело к линейным регрессиям с высокими коэффициентами детерминации и малыми значениями средних абсолютных ошибок (рис. 15). Это позволяет использовать предлагаемую параметрическую модель для восстановления как обменно-корреляционной компоненты, так и для энергии деформации в случае невалентных межмолекулярных взаимодействий.

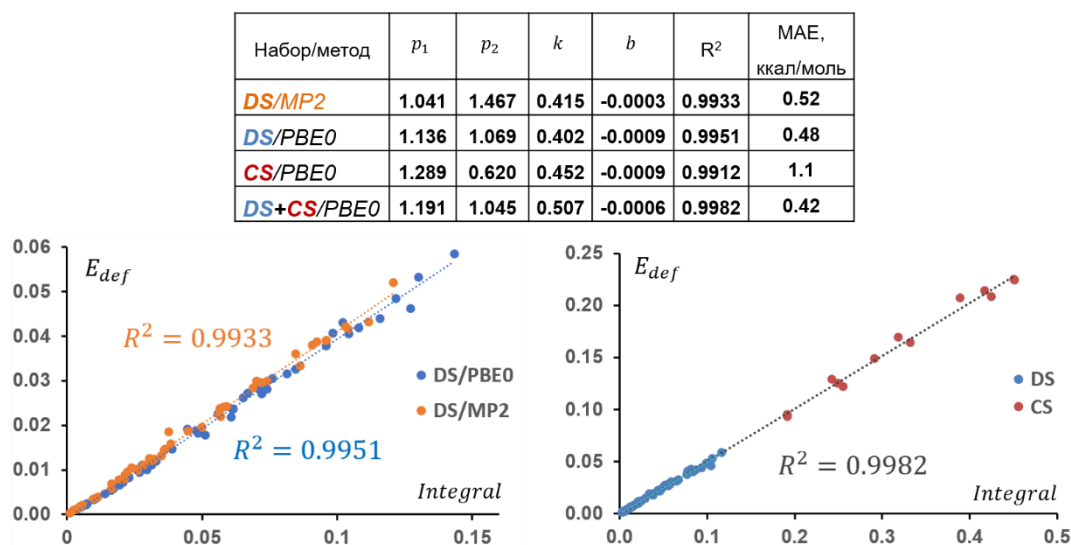


Рисунок 15. Результаты оптимизации параметрической модели для восстановления энергии деформации.

Единственной оставшейся компонентой межмолекулярных взаимодействий является кулоновская энергия, получение которой наименее ресурсозатратно, так как не требует использования матрицы плотности второго порядка или вычисления начального состояния системы до появления интересующего нас взаимодействия. Тем не менее, она все еще получается посредством 6 мерного интеграла. Однако, попытка оптимизации параметрической модели относительно энергии кулоновского взаимодействия привела к неудовлетворительным результатам. Как следствие, для получения данной энергетической компоненты без использования шестимерных интегралов, мы воспользовались усеченным мультипольным разложением вплоть до октуполь-монопольных и квадруполь-дипольных членов (R^{-7}) включительно. Результаты такого приближения оказались вполне удовлетворительными с достаточно низкими значениями средних ошибок и высокими коэффициентами детерминации (рис. 16).

Набор/метод	k	b	R^2	MAE, ккал/моль
DS/PBE0	0.930	-0.0003	0.9885	0.3
CS/PBE0	0.940	-0.0016	0.9787	2.1
DS+CS/PBE0	0.962	-0.0003	0.9895	0.6

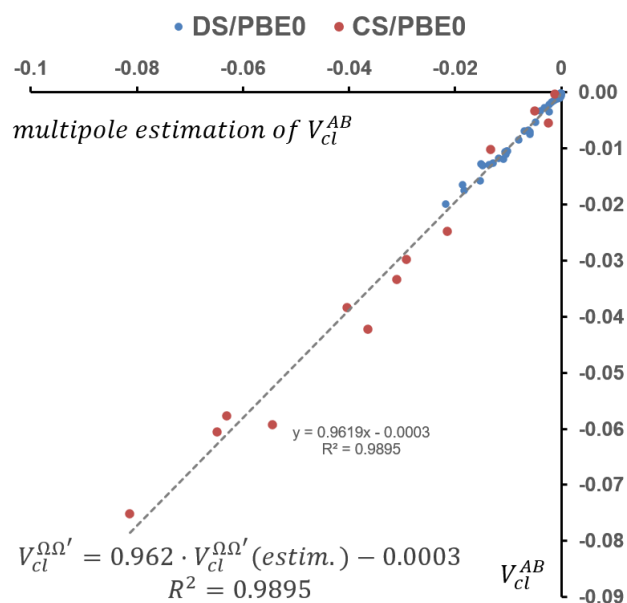


Рисунок 16. Применение усеченного мультипольного разложения для восстановления кулоновской компоненты энергии невалентных межмолекулярных взаимодействий.

Таким образом, суммируем преимущества предложенного подхода:

- В первую очередь, виду того, что предлагаемый подход основан на теории АМ, он анализирует явно выделяемые в пространстве межмолекулярные контакты, что является преимуществом над орбитальными методами анализа разложении энергии.
- Во-вторых, предлагаемый подход требует для своей работы только функцию электронной плотности, в отличие от стандартной процедуры из подхода Взаимодействующих Квантовых Атомов. Также, как следствие, подход применим к экспериментальной электронной плотности.

- Переход от объёмных интегралов к поверхностным приводит к значительному ускорению расчетов. Так, для набора **DS** предложенный метод восстановления компонент энергии межмолекулярных взаимодействий в среднем в 4 раза быстрее стандартного, а для набора **CS** ускорение уже в 8 раз.
- Сумма восстанавливаемых компонент позволяет оценить интегральную энергию связывания E_{bind} . Так, средняя абсолютная ошибка для набора **DS+CS**, рассчитанного посредством PBE0, составила 1.2 ккал/моль.