Discovery of Conservation Laws via Matrix Search

Oliver Schulte and Mark S. Drew
School of Computing Science
Simon Fraser University
Vancouver, Canada
oschulte@cs.sfu.ca

SFU

Outline

- *Problem Definition*: Scientific Model Discovery as Matrix Search.
- Algorithm for discovering *maximally simple maximally strict* conservation laws.
- Comparison with:
  - Particle physics Standard Model (quark model).
  - Molecular structure model.

The Matrix Search Model

- (Valdes, Zytkow, Simon AAAI 1993) Modelling **reactions** in chemistry, physics, engineering.
- *n* entities participating in *m* reactions.
- Input: **Reaction integer matrix** R_{mxn}.
- Output: **Hidden feature integer matrix** Q_{nxq} s.t. $RQ = 0$.
- *Q* classifies reactions as "possible" or "impossible".

Example: Particle Physics

Reactions and Quantities represented as Vectors (Aris 69; Valdés-Pérez 94, 96)

- *i* = 1,…*n* entities
- *r(i)* = # of entity *i* among reagents - # of entity *i* among products.

Particle / Process	1 p	2 π^0	3 μ^-	4 e^+	5 e^-	6 ν_μ	7 $\overline{\nu}_e$
$\mu^- \to e^- + \nu_\mu + \overline{\nu}_e$	0	0	1	0	-1	-1	-1
$p \to e^+ + \pi^0$	1	-1	0	-1	0	0	0
$p + p \to p + p + \pi^0$	0	-1	0	0	0	0	0

Particle / Quantity	1 p	2 π^0	3 μ^-	4 e^+	5 e^-	6 ν_μ	7 $\overline{\nu}_e$
Baryon Number	1	0	0	0	0	0	0
Electric Charge	1	0	-1	1	-1	0	0

A quantity is **conserved** in a reaction if and only if the corresponding vectors are **orthogonal.**

Conserved Quantities in the Standard Model

- Standard Model based on Gell-Mann's quark model (1964).
- Full set of particles: $n = 193$.
- Quantity $\longleftrightarrow$ **Particle Family** (Cluster).

	Particle	Charge	Baryon#	Tau#	Electron#	Muon#
1	Σ^-	-1	1	0	0	0
2	$\overline{\Sigma}^+$	1	-1	0	0	0
3	n	0	1	0	0	0
4	$\overline{n}$	0	-1	0	0	0
5	p	1	1	0	0	0
6	$\overline{p}$	-1	-1	0	0	0
7	π^+	1	0	0	0	0
8	π^-	-1	0	0	0	0
9	π^0	0	0	0	0	0
10	γ	0	0	0	0	0
11	τ^-	-1	0	1	0	0
12	τ^+	1	0	-1	0	0
13	ν_τ	0	0	1	0	0
14	$\overline{\nu}_\tau$	0	0	-1	0	0
15	μ^-	-1	0	0	0	1
16	μ^+	1	0	0	0	-1
17	ν_μ	0	0	0	0	1
18	$\overline{\nu}_\mu$	0	0	0	0	-1
19	e^-	-1	0	0	1	0
20	e^+	1	0	0	-1	0
21	ν_e	0	0	0	1	0
22	$\overline{\nu}_e$	0	0	0	-1	0

The Learning Task (Toy Example)

Given:
1. fixed list of known detectable particles.
2. Input reactions

Not Given:
1. # of quantities
2. Interpretation of quantities.

Reactions

$$\mu^- \to e^-$$

$$p + p \to p + p + \pi^0$$

Reaction Matrix *R*

$$\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

Learning →

Output

Quantity Matrix *Q*

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & -1 \\ 0 & -1 \end{pmatrix}$$

Cols in *Q* are conserved, so *RQ = 0*.

Chemistry Example

- Langley et al. 1987.
- Reactions among Chemical Substances

Substance / Reaction	1 Hydrogen	2 Nitrogen	3 Oxygen	4 Ammonia	5 Water
2 Hydrogen + 1 Oxygen $\rightarrow$ 2 Water $= 2s_1 + s_2 \rightarrow 2s_5$	2	0	1	0	-2
3 Hydrogen + 1 Nitrogen $\rightarrow$ 2 Ammonia $= 3s_1 + s_2 \rightarrow 2s_4$	3	1	0	-2	0

- Interpretation: #element atoms in each substance molecule.
- #element atoms *conserved* in each reaction!

	Element / Substance	H	N	O
1	Hydrogen	2	0	0
2	Nitrogen	0	2	0
3	Oxygen	0	0	2
4	Ammonia	3	1	0
5	Water	2	0	1

The Totalitarian Principle

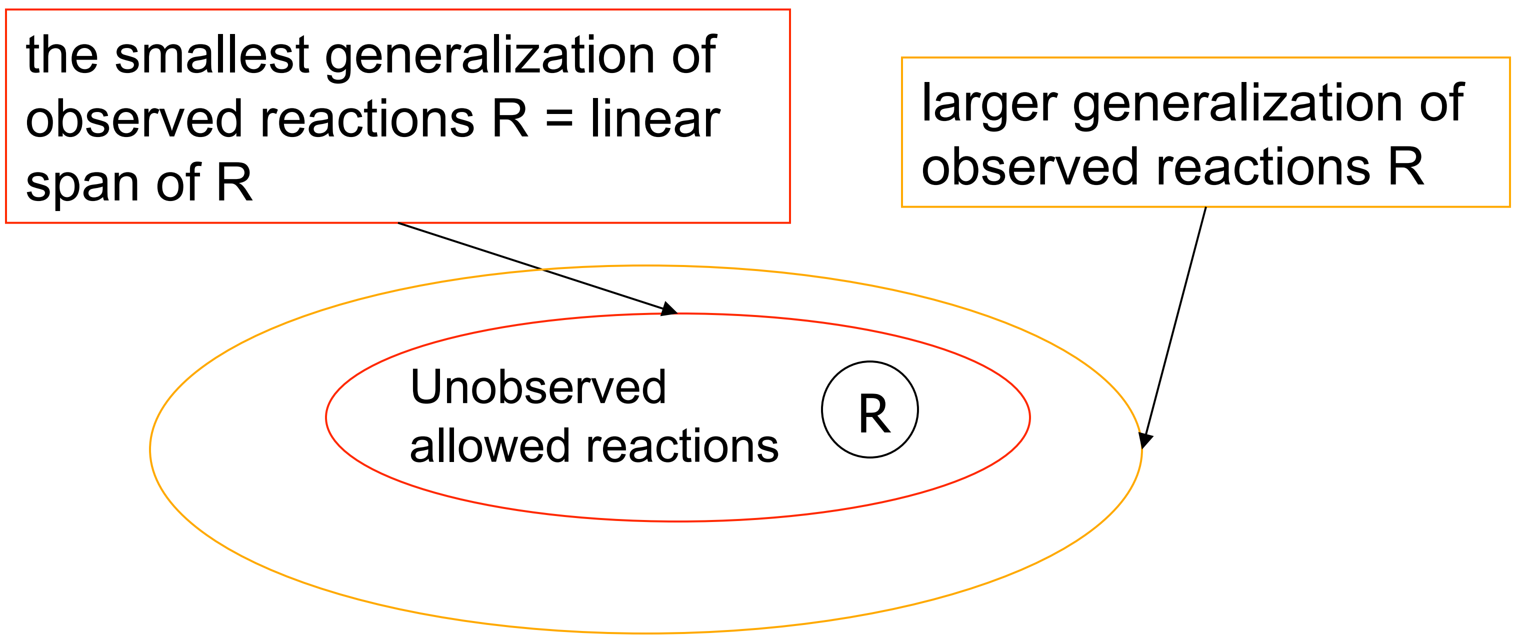
- There are many matrices Q satisfying *RQ = 0*---how to select?
- Apply classic "maximally specific criterion" from version spaces (Mitchell 1990).
- Same general intuition used by physicists:
  - "everything that *can* happen without violating a conservation law *does* happen." Ford 1963.
  - "anything which is not prohibited is compulsory". Gell-Mann 1960s.
- Learning-theoretically optimal (Schulte and Luo COLT 2005)

Maximal Strictness (Schulte IJCAI 09)

Definition. *Q* **is maximally strict** for *R* if *Q* allows a minimal superset of *R*.

Proposition. *Q* is maximally strict for *R* iff the columns of *Q* are a basis for the nullspace of *R*.

nullspace of $R = null(R) = \{v: Rv = 0\}$



Maximally Simple Maximally Strict Matrices (MSMS)

- **L1-norm** $|M|$ of matrix M = sum of absolute values of entries.
- **Definition.** Conservation matrix Q is an **MSMS matrix** for reaction matrix R iff Q minimizes $|Q|$ among maximally strict matrices.

Minimization Algorithm

- **Problem.** Minimize L1-norm $|Q|$, subject to nonlinear constraint: Q columns are basis for nullspace of R.
- **Key Ideas.**

1. *Preprocess* to find a basis V of *null(R)*.
   Search space = $\{X \text{ s.t. } Q = VX\}$.
   X is small continuous change-of-basis matrix.
2. *Discretize* after convergence.
   1. Set small values to 0.
   2. Multiply by lcd to obtain integers.

Example: Chemistry

Input Data

Substance / Reaction	1 Hydrogen	2 Nitrogen	3 Oxygen	4 Ammonia	5 Water
2 Hydrogen + 1 Oxygen → 2 Water $= 2s_1 + s_2 \rightarrow 2s_5$	2	0	1	0	-2
3 Hydrogen + 1 Nitrogen → 2 Ammonia $= 3s_1 + s_2 \rightarrow 2s_4$	3	1	0	-2	0

Minimization Program

$$\begin{pmatrix} 2/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1/2 & 0 \\ 2/3 & 0 & 1/2 \end{pmatrix}$$

MSMS Matrix

multiply by lcd

	Element / Substance	H	N	O
1	Hydrogen	2	0	0
2	Nitrogen	0	2	0
3	Oxygen	0	0	2
4	Ammonia	3	1	0
5	Water	2	0	1

Pseudo Code

Algorithm 1 Minimization Scheme for Finding a Maximally Simple Maximally Strict Conservation Law Matrix

1. Given a set of input reactions R find an orthonormal basis V for the nullspace of R. The basis V is an $n \times q$ matrix.
2. Let any linear combination of V be given by $Q = VX$, with X an $q \times q$ set of coefficients.
   Initialize X to $X_0 = I$, where I is the identity matrix of dimension q.
   Define $\mathcal{I}_1(X) = |VX|$, the L1-norm of the matrix VX.
   Define $\mathcal{I}_2(X) = \sum(X^T X - I)^2$.
3. Minimize $\mathcal{I}_1 + \alpha\mathcal{I}_2$ over X, with α constant, subject to the following constraint:
   (a) To derive an integer version $\widetilde{Q}$, we assign $Q = VX$; $\widehat{\mathbf{q}}_k = \mathbf{q}_k / max(\mathbf{q}_k)$, $k = 1..q$; $\widehat{Q}\left(\widehat{Q} < \varepsilon\right) = 0$; $\widetilde{Q} = sgn(\widehat{Q})$.
   (b) $\widetilde{Q}$ must have full rank: $rank(\widetilde{Q}) = q$.

Comparison with Standard Model

- Implementation in Matlab, use built-in *null* function. Code available on-line.
- Dataset
  - complete set of 193 particles (antiparticles listed separately).
  - included most probable decay for each unstable particle $\Rightarrow$ 182 reactions.
  - Some others from textbooks for **total of 205** reactions.

Results

- S = Standard Model Laws.
- Q = output of minimization.

α	Families Recovered	Runtime (sec)	$\mathcal{I}(Q)$	$\mathcal{I}(S)$	$L1(Q)$	$L1(S)$	difference Q vs. S
20	4/4	16.44	22.67	22.31	22.21	21.96	**C** replaced by linear combination
10	4/4	15.74	22.20	22.31	21.96	21.96	**C** replaced by linear combination
0	n/a	6.95	15.92	22.31	15.92	21.96	invalid local minimum

Ex2: charge given as input.

α	Families Recovered	Runtime (sec)	$\mathcal{I}(Q)$	$\mathcal{I}(S)$	$L1(Q)$	$L1(S)$	difference Q vs. S
20	2/4	7.68	16.65	15.55	16.63	15.52	**E**, **M** replaced by linear combination
10	4/4	8.40	15.55	15.55	15.52	15.52	exact match
0	n/a	10.68	11.52	15.55	11.52	15.52	invalid local minimum

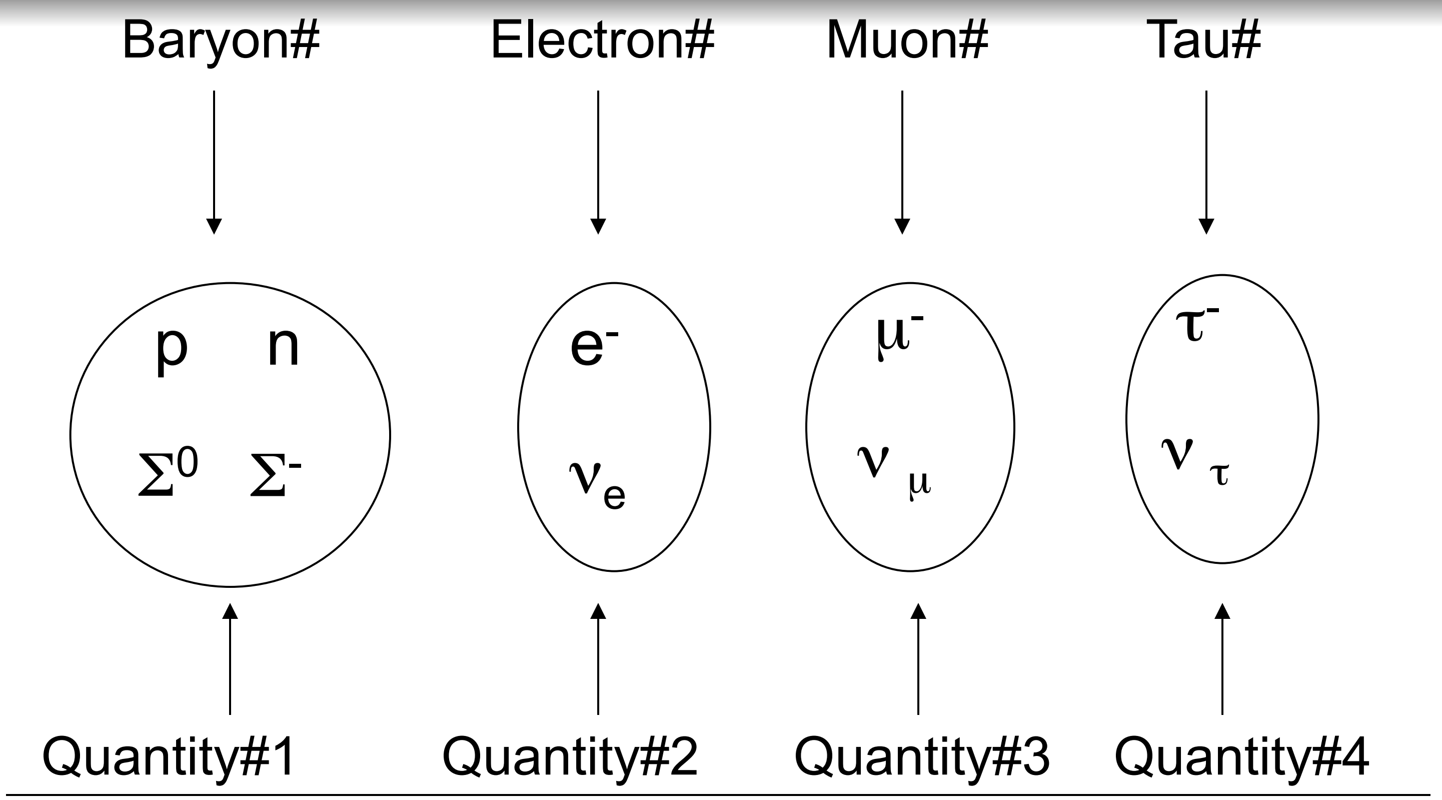
Conclusion

- Conservation Matrix Search problem: for input reactions *R*, solve *RQ*=0.
- **New model selection criterion**: choose maximally simple maximally strict Q.
- Efficient local search optimization.
- Comparison with Standard quark Model
  - Predictively equivalent.
  - (Re)discovers particle families–predicted by theorem.
- MSMS criterion formalizes scientists' objective.

Thank You

- Any questions?

The Family Determination Theorem: Illustration

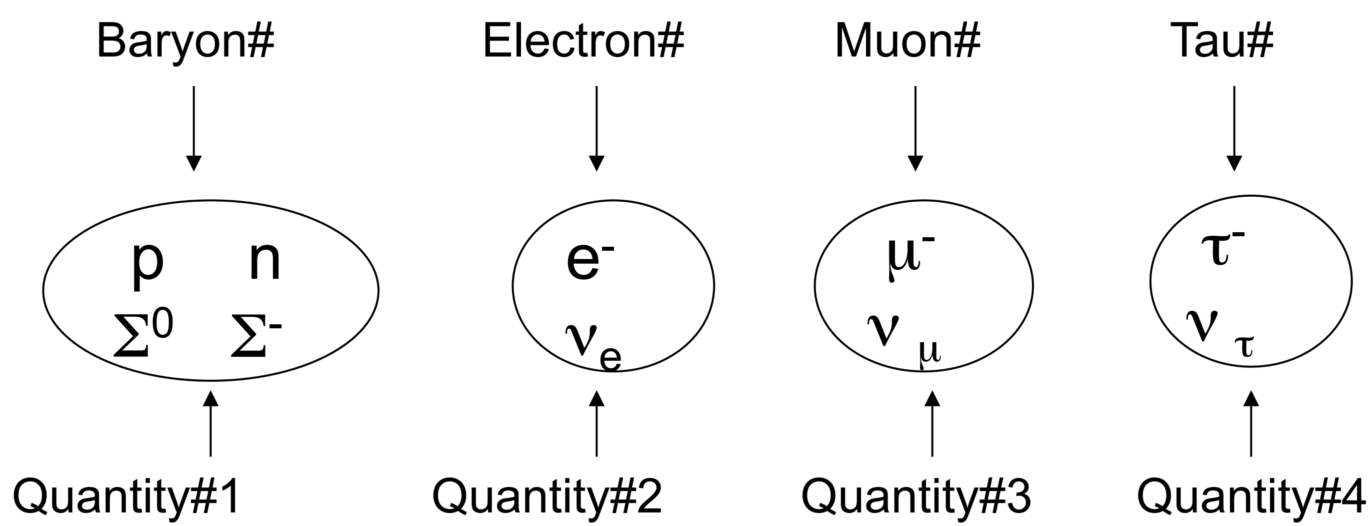


Any alternative set of 4 Q#s with disjoint carriers

Simplicity and Particle Families

Theorem (Schulte and Drew 2006). Let R be a reaction data matrix. If there is a maximally strict conservation matrix Q with disjoint entity clusters, then

- The clusters (families) are *uniquely determined.*
- There is a unique MSMS matrix Q^*.



Any alternative set of 4 Q#s with disjoint carriers