Numerical solution of a class of quasi-linear matrix equations

Valeria Simoncini

Dipartimento di Matematica
Alma Mater Studiorum - Università di Bologna
valeria.simoncini@unibo.it

Joint works with M. Porcelli (UniBo), Y. Hao (Inst. Applied Physics & Comput. Math., Beijing)

The quasi-linear matrix equation problem

Find $X \in \mathbb{R}^{n \times m}$ such that

$$AX + XB + f(X)C = D$$

- $f : \mathbb{R}^{n \times m} \to \mathbb{R}$ linear or nonlinear function
- $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, and $C, D \in \mathbb{R}^{n \times m}$

For certain f, it may occur that $m = n$.

General hypothesis:

A and $-B$ have no common eigenvalues, so that $\mathcal{L} : X \mapsto AX + XB$ is invertible

The quasi-linear matrix equation problem

Find $X \in \mathbb{R}^{n\times m}$ such that

$$AX + XB + f(X)C = D$$

- $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ linear or nonlinear function
- $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, and $C, D \in \mathbb{R}^{n\times m}$

For certain f, it may occur that $m = n$.

General hypothesis:

A and $-B$ have no common eigenvalues, so that $\mathcal{L} : X \mapsto AX + XB$ is invertible

Building up complexity in f

$f : \mathbb{R}^{n\times m} \to \mathbb{R}$ linear or nonlinear function

0. Exception. $f(X) = \sigma_j X$, $j = 1, \ldots, s$

1. f linear:

$$f(X) = \mathrm{trace}(HX), \qquad \text{for some} \quad H$$

   For instance:

   $\star \quad H = I \qquad f(X) = \mathrm{trace}(X)$

   $\star \quad H = uv^T \qquad f(X) = v^T X u$

2. f nonlinear. Composition of
   - Linear with nonlinear, e.g.

$$f(X) = \mathrm{trace}(\exp(-X))$$

   - Nonlinear with linear, e.g.

$$f(X) = \exp(-\mathrm{trace}(X))$$

Building up complexity in f

$f : \mathbb{R}^{n\times m} \to \mathbb{R}$ linear or nonlinear function

0. Exception. $f(X) = \sigma_j X$, $j = 1, \ldots, s$
1. f linear:
$$f(X) = \mathrm{trace}(HX), \qquad \text{for some} \quad H$$
   For instance:
   - $\star \quad H = I \qquad f(X) = \mathrm{trace}(X)$
   - $\star \quad H = uv^T \qquad f(X) = v^T X u$
2. f nonlinear. Composition of
   - Linear with nonlinear, e.g.
$$f(X) = \mathrm{trace}(\exp(-X))$$
   - Nonlinear with linear, e.g.
$$f(X) = \exp(-\mathrm{trace}(X))$$

Building up complexity in f

$f : \mathbb{R}^{n\times m} \to \mathbb{R}$ linear or nonlinear function

0. Exception. $f(X) = \sigma_j X$, $j = 1, \ldots, s$

1. f linear:

$$f(X) = \mathrm{trace}(HX), \qquad \text{for some} \quad H$$

   For instance:

   $\star \qquad H = I \qquad f(X) = \mathrm{trace}(X)$

   $\star \qquad H = uv^T \qquad f(X) = v^T X u$

2. f nonlinear. Composition of
   - Linear with nonlinear, e.g.

$$f(X) = \mathrm{trace}(\exp(-X))$$

   - Nonlinear with linear, e.g.

$$f(X) = \exp(-\mathrm{trace}(X))$$

The linear problem. 1

Let

$$AX + XB + f(X)C = D \qquad (\bullet)$$

with $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ a linear function

Closed form solution:

Let M, N be the solutions to the Sylvester equations $AM + MB = D$ and $AN + NB = C$, resp. Assume that $1 - f(N) \neq 0$. Then the solution to $(\bullet)$ is given by

$$X = M + \sigma N, \quad \sigma = \frac{f(M)}{1 - f(N)}$$

$1 - f(N) = 0$ leads to either infinite or no solutions.

Instead of $(\bullet)$ we can use the mathematically equivalent equation

$$X = M + f(X)N, \qquad N = -\mathcal{L}^{-1}(C), M = \mathcal{L}^{-1}(D)$$

(more appropriate for small rather than large scale problems)

The linear problem. 1

Let

$$AX + XB + f(X)C = D \qquad (\bullet)$$

with $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ a linear function

Closed form solution:

Let M, N be the solutions to the Sylvester equations $AM + MB = D$ and $AN + NB = C$, resp. Assume that $1 - f(N) \neq 0$. Then the solution to $(\bullet)$ is given by

$$X = M + \sigma N, \quad \sigma = \frac{f(M)}{1 - f(N)}$$

$1 - f(N) = 0$ leads to either infinite or no solutions.

Instead of $(\bullet)$ we can use the mathematically equivalent equation

$$X = M + f(X)N, \qquad N = -\mathcal{L}^{-1}(C), M = \mathcal{L}^{-1}(D)$$

(more appropriate for small rather than large scale problems)

The linear problem. 1

Let

$$AX + XB + f(X)C = D \qquad (\bullet)$$

with $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ a linear function

Closed form solution:

Let M, N be the solutions to the Sylvester equations $AM + MB = D$ and $AN + NB = C$, resp. Assume that $1 - f(N) \neq 0$. Then the solution to $(\bullet)$ is given by

$$X = M + \sigma N, \quad \sigma = \frac{f(M)}{1 - f(N)}$$

$1 - f(N) = 0$ leads to either infinite or no solutions.

Instead of $(\bullet)$ we can use the mathematically equivalent equation

$$X = M + f(X)N, \qquad N = -\mathcal{L}^{-1}(C), M = \mathcal{L}^{-1}(D)$$

(more appropriate for small rather than large scale problems)

The linear problem. 2

Some examples: $N = -\mathcal{L}^{-1}(C), M = \mathcal{L}^{-1}(D)$

1. $AX + XB + \mathrm{trace}(X)C = D$. Then

$$X = M + \sigma N, \quad \sigma = \frac{\mathrm{trace}(M)}{1 - \mathrm{trace}(N)}$$

2. $AX + XB + (v^T X u)M = C$. Then

$$X = M + \sigma N, \quad \sigma = \frac{v^T M u}{1 - v^T N u}$$

The linear problem. 2

Some examples: $N = -\mathcal{L}^{-1}(C), M = \mathcal{L}^{-1}(D)$

1. $AX + XB + \mathrm{trace}(X)C = D$. Then

$$X = M + \sigma N, \quad \sigma = \frac{\mathrm{trace}(M)}{1 - \mathrm{trace}(N)}$$

2. $AX + XB + (v^T X u)M = C$. Then

$$X = M + \sigma N, \quad \sigma = \frac{v^T M u}{1 - v^T N u}$$

The linear problem. 3

⇒ The approach also solves a seemingly unrelated problem

Let

$$AX + XB + C_1 X C_2 = D, \qquad C_1, C_2 \quad \text{rank-one} \quad \text{matrices}$$

Letting $C_i = u_i v_i^T$, $i = 1, 2$, then

$$C_1 X C_2 = u_1 v_1^T X u_2 v_2^T = (v_1^T X u_2) u_1 v_2^T \equiv f(X) C$$

♣ The closed form is just the (vector) Sherman-Morrison formula in disguise

(for general low-rank C_1, C_2, see Y. Hao, V.Simoncini, 2021)

Other linear generalizations

- Multiterm case

$$AX + XB + f_1(X)C_1 + \ldots + f_\ell(X)C_\ell = D$$

with $f_j : \mathbb{R}^{n\times m} \to \mathbb{R}$, $j = 1, \ldots, \ell$ linear functions

Closed form solution:

$$X = M + \sum_{i=1}^{\ell} \sigma_i N_i,$$

where $\sigma_j = f_j(X)$ are determined by solving the $\ell \times \ell$ linear system

$$\begin{bmatrix} 1 - f_1(N_1) & -f_1(N_2) & \cdots & -f_1(N_\ell) \\ -f_2(N_1) & 1 - f_2(N_2) & \cdots & -f_2(N_\ell) \\ \vdots & \vdots & \ddots & \vdots \\ -f_\ell(N_1) & \cdots & \cdots & 1 - f_\ell(N_\ell) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_\ell \end{bmatrix} = \begin{bmatrix} f_1(M) \\ \vdots \\ f_\ell(M) \end{bmatrix} \Leftrightarrow (I - F)\boldsymbol{\sigma} = \mathbf{f},$$

(M.Porcelli, V.S., LAA 2023), application to solid mechanics

Other linear generalizations

- Multiterm case

$$AX + XB + f_1(X)C_1 + \ldots + f_\ell(X)C_\ell = D$$

with $f_j : \mathbb{R}^{n\times m} \to \mathbb{R}$, $j = 1, \ldots, \ell$ linear functions

Closed form solution:

$$X = M + \sum_{i=1}^{\ell} \sigma_i N_i,$$

where $\sigma_j = f_j(X)$ are determined by solving the $\ell \times \ell$ linear system

$$\begin{bmatrix} 1 - f_1(N_1) & -f_1(N_2) & \cdots & -f_1(N_\ell) \\ -f_2(N_1) & 1 - f_2(N_2) & \cdots & -f_2(N_\ell) \\ \vdots & \vdots & \ddots & \vdots \\ -f_\ell(N_1) & \cdots & \cdots & 1 - f_\ell(N_\ell) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_\ell \end{bmatrix} = \begin{bmatrix} f_1(M) \\ \vdots \\ f_\ell(M) \end{bmatrix} \Leftrightarrow (I - F)\boldsymbol{\sigma} = \mathbf{f},$$

(M.Porcelli, V.S., LAA 2023), application to solid mechanics

First examples of nonlinear setting

$$f(X) = \mathrm{trace}(X^p), \quad \text{with} \quad p \in \mathbb{N},\; p > 1$$

The square power:

$$\begin{aligned} f(X) &= \mathrm{trace}(X^2) = \mathrm{trace}((M + f(X)N)(M + f(X)N)) \\ &= f(M) + 2\,\mathrm{trace}(MN)f(X) + f(X)^2 f(N). \end{aligned}$$

second order scalar equation in $f(X)$ with roots r_1, r_2.

Closed form:

$$X_{(1)} = M + r_1 N, \qquad X_{(2)} = M + r_2 N.$$

- Similar procedure for, e.g., $f(X) = \|X\|_F^2 = \mathrm{trace}(X^T X)$
- For $f(X) = \mathrm{trace}(X^{-1})$, $M = \boldsymbol{m}_1 \boldsymbol{m}_2^T$ rank-one and N invertible.

If the matrix equation $X = M + f(X)N$ admits nonsingular solutions, then these are $X_{(i)} = M + r_i N$, $i = 1, 2, 3$ where r_i are the roots of

$$r^3 + \eta_2 r^2 + \eta_1 r + \eta_0 = 0,$$

with $\eta_2 = \boldsymbol{m}_2^T N^{-1} \boldsymbol{m}_1$, $\eta_1 = -f(N)$ and $\eta_0 = \eta_1 \eta_2 + \boldsymbol{m}_2^T N^{-2} \boldsymbol{m}_1$

First examples of nonlinear setting

$$f(X) = \mathrm{trace}(X^p), \quad \text{with} \quad p \in \mathbb{N},\ p > 1$$

The square power:

$$\begin{aligned} f(X) &= \mathrm{trace}(X^2) = \mathrm{trace}((M + f(X)N)(M + f(X)N)) \\ &= f(M) + 2\,\mathrm{trace}(MN)f(X) + f(X)^2 f(N). \end{aligned}$$

second order scalar equation in $f(X)$ with roots r_1, r_2.

Closed form:

$$X_{(1)} = M + r_1 N, \qquad X_{(2)} = M + r_2 N.$$

- Similar procedure for, e.g., $f(X) = \|X\|_F^2 = \mathrm{trace}(X^T X)$
- For $f(X) = \mathrm{trace}(X^{-1})$, $M = \boldsymbol{m}_1 \boldsymbol{m}_2^T$ rank-one and N invertible.

If the matrix equation $X = M + f(X)N$ admits nonsingular solutions, then these are $X_{(i)} = M + r_i N$, $i = 1, 2, 3$ where r_i are the roots of

$$r^3 + \eta_2 r^2 + \eta_1 r + \eta_0 = 0,$$

with $\eta_2 = \boldsymbol{m}_2^T N^{-1} \boldsymbol{m}_1$, $\eta_1 = -f(N)$ and $\eta_0 = \eta_1 \eta_2 + \boldsymbol{m}_2^T N^{-2} \boldsymbol{m}_1$

First examples of nonlinear setting

$$f(X) = \mathrm{trace}(X^p), \quad \text{with} \quad p \in \mathbb{N},\ p > 1$$

The square power:

$$\begin{aligned} f(X) &= \mathrm{trace}(X^2) = \mathrm{trace}((M + f(X)N)(M + f(X)N)) \\ &= f(M) + 2\,\mathrm{trace}(MN)f(X) + f(X)^2 f(N). \end{aligned}$$

second order scalar equation in $f(X)$ with roots r_1, r_2.

Closed form:

$$X_{(1)} = M + r_1 N, \qquad X_{(2)} = M + r_2 N.$$

- Similar procedure for, e.g., $f(X) = \|X\|_F^2 = \mathrm{trace}(X^T X)$
- For $f(X) = \mathrm{trace}(X^{-1})$, $M = \boldsymbol{m}_1 \boldsymbol{m}_2^T$ rank-one and N invertible.

If the matrix equation $X = M + f(X)N$ admits nonsingular solutions, then these are $X_{(i)} = M + r_i N$, $i = 1, 2, 3$ where r_i are the roots of

$$r^3 + \eta_2 r^2 + \eta_1 r + \eta_0 = 0,$$

with $\eta_2 = \boldsymbol{m}_2^T N^{-1} \boldsymbol{m}_1$, $\eta_1 = -f(N)$ and $\eta_0 = \eta_1 \eta_2 + \boldsymbol{m}_2^T N^{-2} \boldsymbol{m}_1$

First examples of nonlinear setting

$$f(X) = \mathrm{trace}(X^p), \quad \text{with} \quad p \in \mathbb{N},\ p > 1$$

The square power:

$$\begin{aligned} f(X) &= \mathrm{trace}(X^2) = \mathrm{trace}((M + f(X)N)(M + f(X)N)) \\ &= f(M) + 2\,\mathrm{trace}(MN)f(X) + f(X)^2 f(N). \end{aligned}$$

second order scalar equation in $f(X)$ with roots r_1, r_2.

Closed form:

$$X_{(1)} = M + r_1 N, \qquad X_{(2)} = M + r_2 N.$$

- Similar procedure for, e.g., $f(X) = \|X\|_F^2 = \mathrm{trace}(X^T X)$
- For $f(X) = \mathrm{trace}(X^{-1})$, $M = \boldsymbol{m}_1 \boldsymbol{m}_2^T$ rank-one and N invertible.

If the matrix equation $X = M + f(X)N$ admits nonsingular solutions, then these are $X_{(i)} = M + r_i N$, $i = 1, 2, 3$ where r_i are the roots of

$$r^3 + \eta_2 r^2 + \eta_1 r + \eta_0 = 0,$$

with $\eta_2 = \boldsymbol{m}_2^T N^{-1} \boldsymbol{m}_1$, $\eta_1 = -f(N)$ and $\eta_0 = \eta_1 \eta_2 + \boldsymbol{m}_2^T N^{-2} \boldsymbol{m}_1$

The general linear-nonlinear

$$f(X) = \phi(\psi(X)), \quad \phi : \mathbb{R}^{n\times n} \to \mathbb{R}, \quad \psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},$$

where ϕ is linear, and ψ is a (nonlinear) matrix function

Note: in the following, $\phi(Y) = \mathrm{trace}(Y)$ E.g., $f(X) = \mathrm{trace}(\exp(-X))$

Use

$$X = M + f(X)N$$

and assume N diag.ble, $N = Q\Lambda Q^{-1}$. Then

$$Q^{-1}XQ = Q^{-1}MQ + f(X)\Lambda,$$

Note that (for trace invariance)

$$f(X) = \mathrm{trace}(\psi(X)) = \mathrm{trace}(\psi(Q^{-1}XQ)) = f(Q^{-1}XQ),$$

so that

$$X_1 = M_1 + f(X_1)\Lambda, \qquad X_1 \equiv Q^{-1}XQ, \; M_1 \equiv Q^{-1}MQ$$

$\Rightarrow$ Only the diagonal is updated!

The general linear-nonlinear

$$f(X) = \phi(\psi(X)), \quad \phi : \mathbb{R}^{n\times n} \to \mathbb{R}, \quad \psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},$$

where ϕ is linear, and ψ is a (nonlinear) matrix function

Note: in the following, $\phi(Y) = \mathrm{trace}(Y)$ E.g., $f(X) = \mathrm{trace}(\exp(-X))$

Use

$$X = M + f(X)N$$

and assume N diag.ble, $N = Q\Lambda Q^{-1}$. Then

$$Q^{-1}XQ = Q^{-1}MQ + f(X)\Lambda,$$

Note that (for trace invariance)

$$f(X) = \mathrm{trace}(\psi(X)) = \mathrm{trace}(\psi(Q^{-1}XQ)) = f(Q^{-1}XQ),$$

so that

$$X_1 = M_1 + f(X_1)\Lambda, \qquad X_1 \equiv Q^{-1}XQ, \; M_1 \equiv Q^{-1}MQ$$

$\Rightarrow$ Only the diagonal is updated!

Numerical solution

Fixed point iteration:

$$X_1^{(k+1)} = M_1 + f(X_1^{(k)})\Lambda, \qquad \text{for some } X_1^{(0)}$$

Definiteness properties:

Let $M_1 \succ 0$ and $\Lambda \succeq 0$, and let $X_1^{(0)} = M_1$.
i) If f is a nonnegative function satisfying $f(X) \le f(Y)$ for $Y \succeq X$, then $X_1^{(k+1)} \succeq X_1^{(k)}$ for all ks
ii) If f is a nonnegative function satisfying $f(X) \ge f(Y)$ for $Y \succeq X$, then the iterates $X_1^{(k+1)} - X_1^{(k)}$ alternate definiteness at each k

Numerical solution

Fixed point iteration:

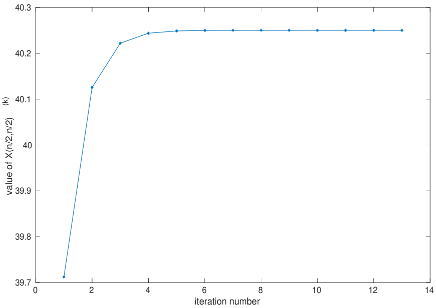
$$X_1^{(k+1)} = M_1 + f(X_1^{(k)})\Lambda, \qquad \text{for some } X_1^{(0)}$$

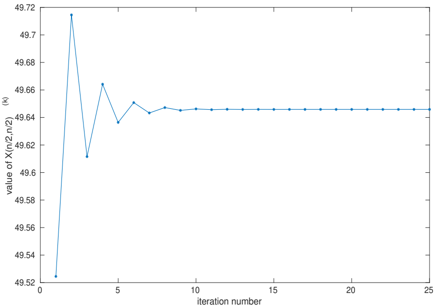
Definiteness properties:

Let $M_1 \succ 0$ and $\Lambda \succeq 0$, and let $X_1^{(0)} = M_1$.
i) If f is a nonnegative function satisfying $f(X) \leq f(Y)$ for $Y \succeq X$, then $X_1^{(k+1)} \succeq X_1^{(k)}$ for all ks
ii) If f is a nonnegative function satisfying $f(X) \geq f(Y)$ for $Y \succeq X$, then the iterates $X_1^{(k+1)} - X_1^{(k)}$ alternate definiteness at each k

An example

Convergence of the $(n/2, n/2)$ diagonal element of $X_1^{(k)}$





Left: $f(X) = \mathrm{trace}(X^{1/2})$

Right: $f(X) = \mathrm{trace}(\exp(-X))$

Convergence to exact solution $X_1^\star$

Consider $f(X) = \mathrm{trace}(\exp(-X))$

Let $E^{(k)} = X_1^{(k)} - X_1^\star$

Ostrowski-type theorem:

Assume $M_1 \succeq 0$ and $\Lambda \succeq 0$.
If $\mathrm{trace}(\Lambda \exp(-X_1^\star)) = \sigma < 1$ then there exist an $X_1^{(0)}$ and a $\sigma_1 \in [0, 1)$ such that

$$\|E^{(k+1)}\| \leq \sigma_1 \|E^{(k)}\|,$$

for $k \geq 0$, for any matrix norm $\|\cdot\|$.

Note: A corresponding result holds for $f(X) = \mathrm{trace}(X^{\frac{1}{2}})$

An example

Consider:

♣ $X^\star = \sqrt{\alpha} G$

♣ $G = (G_0^T G_0)^{\frac{1}{2}}$, with G_0 =`randn(n,n)` (Matlab seed `rng(1)`)

♣ N similar to G, and $M = X^\star - f(X^\star)N$

$\Rightarrow$ α influences the magnitude of the Frechet derivative

$$X_1^{(k+1)} = M_1 + f(X_1^{(k)})\Lambda, \qquad \text{for} \;\; X_1^{(0)} = M_1$$

$\mathrm{trace}(\Lambda \exp(-X_1^\star))$	α	k	$\frac{\lVert X^{(k+1)} - (M + f(X^{(k+1)})N) \rVert}{\lVert M \rVert}$
0.079	12.589	3	8.3190e-08
0.176	10.000	6	3.4123e-08
0.335	7.9433	11	3.7944e-08
0.570	6.3096	23	6.9902e-08
0.889	5.0119	117	9.6324e-08
1.296	3.9811	500	3.5943e-01
1.789	3.1623	500	1.2832e+00

Considerations on the large scale problem

- Linear problem. $f(X) = \mathrm{trace}(X)$,

$$\widetilde{X} \equiv \widetilde{M} + \sigma \widetilde{N}, \quad \sigma = \frac{f(\widetilde{M})}{1 - f(\widetilde{N})},$$

 where $\widetilde{M}, \widetilde{N}$ approximate M and N resp.
 (easy case)

- Linear-nonlinear problem. For fixed point iteration,

$$\widetilde{X}^{(k+1)} \equiv \widetilde{M} + f(\widetilde{X}^{(k)})\widetilde{N}.$$

 which requires approximating $f(\widetilde{X}^{(k)})$, e.g., $f(\widetilde{X}) = \mathrm{trace}(\psi(\widetilde{X}))$ - a problem in its own.

Considerations on the large scale problem

- Linear problem. $f(X) = \mathrm{trace}(X)$,

$$\widetilde{X} \equiv \widetilde{M} + \sigma \widetilde{N}, \quad \sigma = \frac{f(\widetilde{M})}{1 - f(\widetilde{N})},$$

 where $\widetilde{M}, \widetilde{N}$ approximate M and N resp.
(easy case)

- Linear-nonlinear problem. For fixed point iteration,

$$\widetilde{X}^{(k+1)} \equiv \widetilde{M} + f(\widetilde{X}^{(k)})\widetilde{N}.$$

which requires approximating $f(\widetilde{X}^{(k)})$, e.g., $f(\widetilde{X}) = \mathrm{trace}(\psi(\widetilde{X}))$ - a problem in its own.

Conclusions

- Quasi-linear matrix equations are a new source of open problems
- The large scale setting is a challenge
- Generalizations to tensor case is possible

REFERENCES

Margherita Porcelli, and V. Simoncini
Numerical solution of a class of quasi-linear matrix equations
Linear Algebra and Its Applications 664C, 2023

Yue Hao and V. Simoncini
The Sherman-Morrison-Woodbury formula for generalized linear matrix equations and applications
Numer. Linear Algebra w/Appl. 28(5), 2021

www.dm.unibo.it/~simoncin
valeria.simoncini@unibo.it

… See you there

METT X

10th Workshop on Matrix Equations and Tensor Techniques

September 13–15, 2023

RWTH Aachen University (main building)

https://www.igpm.rwth-aachen.de/workshop/mett2023

Special Issue of ETNA (Electronic Transactions on Numerical Analysis), open for participants only!

Fully (diamond) Open Access without OA charges!