Transposition Method for BSDEs/BSEEs, and Applications

Xu Zhang
School of Mathematics
Sichuan University, China
zhang_xu@scu.edu.cn
Outline:

1. The classical transposition method in PDEs
2. Transposition solution to BSDEs
3. Numerical schemes: a finite transposition method
4. Stochastic optimal control problems in infinite dimensions
5. Well-posedness of vector-valued BSEEs
6. Well-posedness of operator-valued BSEEs
7. Pontryagin-Type Stochastic Maximum Principle
1. The classical transposition method in PDEs

In this talk, we shall present a new method to solve BSDEs/BSEEs. The idea comes from the transposition method for (deterministic) non-homogeneous boundary value problems by the book of J.-L. Lions and E. Magenes (1972), see also the book of J.-L. Lions (1988).

We now recall the main idea in the classical transposition method to solve the following wave equation with non-homogeneous Dirichelt boundary conditions:

\[
\begin{cases}
  y_{tt} - \Delta y = 0 & \text{in } Q \equiv (0, T) \times G, \\
  y = u & \text{on } \Sigma \equiv (0, T) \times \Gamma, \\
  y(0) = y_0, \quad y_t(0) = y_1 & \text{in } G,
\end{cases}
\]

where $T > 0$, $G$ is a nonempty open bounded domain in $\mathbb{R}^d$ ($d \in \mathbb{N}$) with $C^2$ boundary $\Gamma$, $(y_0, y_1) \in L^2(G) \times H^{-1}(G)$ and $u \in L^2((0, T) \times \Gamma)$. 

When \( u \equiv 0 \), one can use the standard Semigroup Theory to show the well-posedness of (1).

When \( u \not\equiv 0 \), one needs to use the transposition method. For this purpose, for any \( f \in L^1(0, T; L^2(\Omega)) \) and \( g \in L^1(0, T; H^1_0(\Omega)) \), consider the following adjoint problem of (1):

\[
\begin{cases}
\zeta_{tt} - \Delta \zeta = f + g_t, & \text{in } Q, \\
\zeta = 0, & \text{on } \Sigma, \\
\zeta(T) = \zeta_t(T) = 0, & \text{in } G.
\end{cases}
\]

(2)

It is easy to show that the equation (2) admits a unique solution \( \zeta \in C([0, T]; H^1_0(G)) \cap C^1([0, T]; L^2(G)) \), which enjoys the regularity \( \frac{\partial \zeta}{\partial \nu} \in L^2(\Sigma) \).
In order to give a reasonable definition for the solution to the non-homogenous boundary problem (1) by the transposition method, we consider first the case when $y$ is sufficiently smooth. The following result holds:

Assume $g \in C_0^\infty(0, T; H^1_0(G))$ and that $y \in H^2(Q)$ satisfies (1). Then

$$\int_Q fydxdt - \int_Q gy_tdxdt$$

$$= \int_G \zeta(0)y_1dx - \int_G \zeta_t(0)y_0dx - \int_\Sigma \frac{\partial \zeta}{\partial \nu}ud\Sigma. \quad (3)$$
Note that (3) still makes sense even if the regularity of \( y \) is relaxed as \( y \in C([0, T]; L^2(G)) \cap C^1([0, T]; H^{-1}(G)). \) Because of this, one introduces the following:

**Definition 1.** We call \( y \in C([0, T]; L^2(G)) \cap C^1([0, T]; H^{-1}(G)) \) a solution to (1), in the sense of transposition, if \( y(0) = y_0, \ y_t(0) = y_1, \) and for any \( f \in L^1(0, T; L^2(G)) \) and any \( g \in L^1(0, T; H^1_0(G)) \), it holds that

\[
\begin{align*}
\int_Q fy dx dt - \int_0^T \langle g, y_t \rangle_{H^1_0(G), H^{-1}(G)} dt &= \langle \zeta(0), y_1 \rangle_{H^1_0(G), H^{-1}(G)} + \int_\Omega \zeta_t(0)y_0 dx - \int_\Sigma \frac{\partial \zeta}{\partial \nu} ud\Sigma,
\end{align*}
\]

where \( \zeta \) is the unique solution to (2).
One can show the well-posedness of (1) in the sense of transposition. The principle idea of this method is to interpret the solution to one forward wave equation with non-homogeneous Dirichlet boundary conditions in terms of another backward wave equation with non-homogeneous source terms.

We shall use this idea to interpret BSDEs/BSEEs in terms of SDEs/SEEs. This enables us

- To provide a new method for solving BSDEs/BSEEs with general filtration;
- To give a new numerical schemes for solving BSDEs (even with the natural filtration);
- To establish a general Pontryagin-type stochastic maximum principle in infinite dimensions.

The transposition method is a variant of duality method. Like a mirror, it provides a way to see something which is not easy to be detected directly.
2. Transposition solution to BSDEs

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space with $\mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$, on which a 1-dimensional standard Brownian motion $\{w(t)\}_{t \in [0, T]}$ is defined.

Denote by $\mathbb{W}$ the natural filtration generated by $\{w(t)\}$ and augmented by all the $\mathbb{P}$-null sets.

Consider the following semilinear BSDE:

$$
\begin{align*}
  dy(t) &= f(t, y(t), Y(t))dt + Y(t)dw(t) \quad \text{in } [0, T], \\
  y(T) &= y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n),
\end{align*}
$$

(4)

where $f(\cdot, \cdot, \cdot)$ satisfies the usual Lipschitz condition and $f(\cdot, 0, 0) \in L^2_{\mathbb{F}}(\Omega; L^1(0, T; \mathbb{R}^n))$. 
• The case $\mathcal{F} = \mathcal{W}$ is well-understood, see the classical works J.-M. Bismut (1973), E. Pardoux and S. Peng (1990). In this case, one uses the Martingale Representation Theorem (MRT) to guarantee the existence of $Y(\cdot)$, which is however not constructive, and therefore the corresponding computation becomes a difficult problem.

• When $\mathcal{W} \subsetneq \mathcal{F}$, since MRT does not work any more, the work by N. El Karoui and S.-J. Huang (1997) shows that one needs to introduce an extra corrected term to (4), and therefore, it is even more difficult to “compute” the above $Y(\cdot)$.
Recently, by replacing $Y(t)dw(t)$ in (4) by $dM(t)$ (with $M(\cdot)$ being a square-integrable martingale), G. Liang T. Lyons and Z. Qian (2008) developed another approach for the well-posedness of BSDEs with the general filtration.

The advantage of this approach is that MRT is not required, either. But the cost is that the corrected term $Y(\cdot)$ in (4) is suppressed. Note that this term plays a crucial role in some problems, say the Pontryagin-type maximum principle for general stochastic optimal control problems. Also, the comparison theorem is not clear in this setting because the usual duality analysis is not available.
Similar to the transposition method for non-homogeneous boundary value problems, for fixed $t \in [0, T]$, we start from the following linear (forward) stochastic differential equation

$$\left\{ \begin{array}{l}
dz(\tau) = u(\tau)d\tau + v(\tau)dw(\tau), \tau \in (t, T], \\
z(t) = \eta.
\end{array} \right. \tag{5}$$

It is clear that, for given $u(\cdot) \in L^2_F(\Omega; L^1(t, T; \mathbb{R}^n))$, $v(\cdot) \in L^2_F(\Omega; L^2(t, T; \mathbb{R}^n))$ and $\eta \in L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^n)$, equation (5) admits a unique strong solution $z(\cdot) \in L^2_F(\Omega; C([t, T]; \mathbb{R}^n))$. 
Now, if equation (4) admits a strong solution \((y(\cdot), Y(\cdot)) \in L^2_F(\Omega; C([0, T]; \mathbb{R}^n)) \times L^2_F(0, T; L^2(\Omega; \mathbb{R}^n))\) (say, when \(F = \mathcal{W}\)), then, applying Itô’s formula to \(\langle z(t), y(t) \rangle\), it follows

\[
\mathbb{E} \langle z(T), y_T \rangle - \mathbb{E} \langle \eta, y(t) \rangle = \mathbb{E} \int_t^T \langle z(\tau), f(\tau, y(\tau), Y(\tau)) \rangle \, d\tau
\]

\[
+ \mathbb{E} \int_t^T \langle u(\tau), y(\tau) \rangle \, d\tau
\]

\[
+ \mathbb{E} \int_t^T \langle v(\tau), Y(\tau) \rangle \, d\tau.
\]
This inspires us to introduce the following new notion for the solution to the equation (4).

**Definition 2.** We call \((y(\cdot), Y(\cdot)) \in L^2_F(\Omega; D([0, T]; \mathbb{R}^n)) \times L^2_F(\Omega; L^2(0, T; \mathbb{R}^n))\) a transposition solution to the equation (4) if for any \(t \in [0, T]\), \(u(\cdot) \in L^2_F(\Omega; L^1(t, T; \mathbb{R}^n))\), \(v(\cdot) \in L^2_F(\Omega; L^2(t, T; \mathbb{R}^n))\) and \(\eta \in L^2_{F_t}(\Omega; \mathbb{R}^n)\), the identity (6) holds.

Clearly, any transposition solution to the equation (4) coincides with its strong solution whenever the filtration \(\mathbb{F}\) is natural.
The well-posedness of BSDEs in the sense of transposition method is as follows

**Theorem 1.** (Q. Lü and X. Zhang, 2010) For any given \( y_T \in L^2_{\mathcal{F}_T}(\Omega) \), the equation (4) admits a unique transposition solution \( (y(\cdot), Y(\cdot)) \in L^2_{\mathcal{F}}(\Omega; D([0, T]; \mathbb{R}^n)) \times L^2_{\mathcal{F}}(\Omega; L^2(0, T; \mathbb{R}^n)) \). Furthermore, there is a constant \( C > 0 \), depending only on \( K \) and \( T \), such that

\[
\left| (y(\cdot), Y(\cdot)) \right|_{L^2_{\mathcal{F}}(\Omega; D([0, T]; \mathbb{R}^n)) \times L^2_{\mathcal{F}}(\Omega; L^2(0, T; \mathbb{R}^n))} \leq C \left[ \left| f(\cdot, 0, 0) \right|_{L^2_{\mathcal{F}}(\Omega; L^1(0, T; \mathbb{R}^n))} + \left| y_T \right|_{L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)} \right].
\]

(7)
Our method needs neither the MRT nor the Itô formula. One can expect this method could be adopted to more general situation.

The main advantage of our method consists in the fact that the duality analysis is contained in the definition of solutions, and therefore, one can easily deduce a similar comparison theorem for transposition solutions of (4) by using almost the same approach as in N. El Karoui, S. Peng and M. C. Quenez (1997).

Also, it is even easier to establish the Pontryagin-type maximum principle for general stochastic optimal control problems than to solve the same problem with the natural filtration because, again, the desired duality analysis is contained in our definition of solutions.
More importantly, our method provides a variational formulation for BSDEs. To see this, we recall that, for any $f \in L^2(G)$, in order to find the strong solution $y \in H^1_0(G) \cap H^2(G)$ of the classical elliptic PDE:

$$\begin{cases}
-\Delta y = f, \text{ in } G, \\
y = 0, \text{ on } \Gamma,
\end{cases} \quad (8)$$

one considers first its variational formulation

$$\int_G (\nabla y) \cdot (\nabla \phi) \, dx = \int_G f \phi \, dx, \quad \forall \phi \in H^1_0(G). \quad (9)$$

Clearly, (6) is the variational formulation of the BSDE (4). We shall see later that this variational formulation is the basis of our numerical method to solve BSDEs.
3. Numerical schemes: a finite transposition method

Our numerical method is, in spirit, very close to the classical finite element method solving deterministic PDE (8), to be recalled below.

The main idea of the classical finite element method is quite simple. Instead of looking for solution $y \in H^1_0(G) \cap H^2(G)$ to (8) directly, one chooses a suitable finite dimensional space $H_N \subset (H^1_0(G))$, called a finite element space, $N \in \mathbb{N}$, and solves $y_N \in H_N$ to the following discrete version of the variational equation (9), i.e.

$$
\int_G (\nabla y_N) \cdot (\nabla \phi) dx = \int_G f \phi dx, \quad \forall \phi \in H_N. \quad (10)
$$

This is the so-called finite element equation, which is an algebraic equation. One then shows the convergence of the sequence $\{y_N\}$ in suitable spaces.
Clearly, our method to solve BSDEs, to be presented later, can be viewed as a stochastic version of the finite element-type method.

Nevertheless, the notion of “stochastic finite element method” has already been used for other purpose, say R. Ghanem and P. Spanos (1991), M. Kleiber and T. D. Hien (1992), A. Nouy (2009) and references therein for solving random PDEs. Note also that our method is quite different from that in these references, and therefore instead we call it a finite transposition method.

There are at least two reasons for us to develop this new numerical approach for BSDEs. The first one is that, we can solve the BSDE with general filtration. The second is that, even in the case of natural filtration, our method is quite different from the existing ones for solving BSDEs.
We consider only the linear problem with a nonhomonomous term $f(\cdot) \in L^2_F(\Omega; L^1(0, T; \mathbb{R}^n))$:

$$
\begin{align*}
\begin{cases}
dy(t) = f(t)dt + Y(t)dw(t), & t \in [0, T), \\
y(T) = y_T
\end{cases}
\end{align*}
$$

and assume that $L^2_F(\Omega; \mathbb{R}^n)$ is separable.

Write $\mathcal{R}_N = \{t_\ell \mid t_\ell = \frac{\ell}{2^N}T, \ \ell = 0, \cdots, 2^N\}$. For any $k \in \{0, \cdots, 2^N - 1\}$, define a sequence of simple processes $\{e_{ki}(\cdot, \cdot)\}_{i=1}^{M_{k,N}}$ by

$$
e_{ki}(t, \omega) = \begin{cases} 
\chi_{[t_k,t_{k+1}]}(t)h_{ki}(\omega), & 0 \leq k < 2^N - 1, \\
\chi_{[t_k,T]}(t)h_{ki}(\omega), & k = 2^N - 1,
\end{cases}
$$

where $\{M_{0,N}, M_{1,N}, \cdots, M_{2^N-1,N}\}$ is an increasing sequence of integers, and $\{h_{ki}\}$ satisfy the following:
1) For any fixed \( k \in \{0, \cdots, 2^N - 1\} \), \( \{h_{ki}\}_{i=0}^{M_{k,N}} \) is an orthogonal set in \( L^2_{\mathcal{F}_{tk}}(\Omega; \mathbb{R}^n) \), and the norm \( |h_{ki}|_{L^2_{\mathcal{F}_{tk}}(\Omega; \mathbb{R}^n)} = \sqrt{\frac{2^N}{T}} \), and hence \( |e_{ki}|_{L^2_{\mathcal{F}}(\Omega; L^2(0,T; \mathbb{R}^n))} = 1; \)

2) If \( 0 \leq k < \ell \leq 2^N - 1 \), then \( \{h_{ki}\}_{i=0}^{M_{k,N}} \subset \{h_{\ell i}\}_{i=0}^{M_{\ell,N}} \); and

3) If \( s_0 = \frac{k_0}{2^{N_0}}T \) for some \( N_0 \in \mathbb{N} \) and \( k_0 \in \{0, \cdots, 2^{N_0} - 1\} \), then \( s_0 = 2^{\ell-N_0}k_0T/2^\ell \in \mathcal{R}_\ell \) for any \( \ell \geq N_0 \). For \( \ell \geq N_0 \), write \( k_\ell = 2^{\ell-N_0}k_0 \). Then, \( \{h_{k_{\ell i}}\}_{i=0}^{M_{k_{\ell},\ell}} \subset \{h_{k_{\ell i}}\}_{i=0}^{M_{k_{\ell},\ell}} \) for \( N_0 \leq j < \ell \), and \( \bigcup_{\ell=N_0}^{\infty} \{h_{k_{\ell i}}\}_{i=1}^{M_{k_{\ell},\ell}} \) is an orthogonal basis of \( L^2_{\mathcal{F}_{s_0}}(\Omega; \mathbb{R}^n) \).
Denote by $H_N$ the subspace of $L^2_{\mathcal{F}}(\Omega; L^2(0, T; \mathbb{R}^n))$ spanned by $\{e_{0i}\}_{i=1}^{M_{0,N}}, \cdots, \{e_{2N-1,i}\}_{i=1}^{M_{2N-1,N}}$. This is the finite element space that we will employ below for our BSDE problems. Replace $L^2_{\mathcal{F}}(\Omega; L^2(0, T; \mathbb{R}^n))$ by $H_N$, then the desired numerical scheme follows by trying to find $y_N, Y_N \in H_N$ such that the following discrete variational equality holds for all $u, v \in H_N$, and $\eta = 0$.

\[
\mathbb{E} \langle z(T), y_T \rangle = \mathbb{E} \int_0^T \langle z(\tau), f(\tau) \rangle \, d\tau \\
+ \mathbb{E} \int_0^T \langle u(\tau), y_N(\tau) \rangle \, d\tau \\
+ \mathbb{E} \int_0^T \langle v(\tau), Y_N(\tau) \rangle \, d\tau
\]  

(12)
To find the $y_N$ in $H_N$, suppose $y_N = \sum_{k=0}^{2^N-1} \sum_{i=0}^{M_{k,N}} \alpha_{ki} e_{ki}$.

Choosing $u = e_{ki}$, $v = 0$ and $\eta = 0$, we get $z_{ki}(t) = \int_0^t e_{ki}(\tau) d\tau$, and hence

$$
\mathbb{E}\langle z_{ki}(T), y_T \rangle
= \mathbb{E} \int_0^T \langle z_{ki}(\tau), f(\tau) \rangle d\tau + \sum_{\ell,j} \alpha_{\ell j} \mathbb{E} \int_0^T \langle e_{ki}(\tau), e_{\ell j}(\tau) \rangle d\tau.
$$

Since $\{e_{ki}\}$ is an orthogonal basis of $H_N$, it follows that $\mathbb{E} \int_0^T \langle e_{\ell j}(\tau), e_{ki}(\tau) \rangle d\tau = \delta_{k\ell} \delta_{ij}$. Therefore,

$$
\alpha_{ki} = \frac{T}{2^N} \mathbb{E}\langle h_{ki}, y_T \rangle
- \mathbb{E} \int_0^T \langle (\tau \land t_{k+1} - \tau \land t_k) h_{ki}, f(\tau) \rangle d\tau.
$$
Similarly, suppose $Y_N = \sum_{k=0}^{2^N-1} \sum_{i=0}^{M_{k,N}} \beta_{ki} e_{ki}$. By taking $u = 0$, $\eta = 0$ and $v = e_{ki}$ to get a corresponding $z_{ki}(t) = \int_{0}^{t} e_{ki}(\tau) d\omega(\tau)$, we find that

$$\beta_{ki} = \mathbb{E}\langle (w(t_{k+1}) - w(t_k)) h_{ki}, y_T \rangle$$

$$- \mathbb{E} \int_{0}^{T} \langle (w(\tau \wedge t_{k+1}) - w(\tau \wedge t_k)) h_{ki}, f(\tau) \rangle d\tau.$$

We can show the convergence of the sequence $\{(y_N, Y_N)\}$ of numerical solutions, i.e.

**Theorem 2.** (P. Wang and X. Zhang, 2011)

The sequence $\{(y_N, Y_N)\}$ tends to $(y, Y)$ in $L^2_{\mathbb{F}}(\Omega; D([0, T], \mathbb{R}^n)) \times L^2_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^n))$, as $N \to \infty$. 
4. Stochastic optimal control problems in infinite dimensions

Consider the following controlled forward stochastic evolution equation

\[
\begin{align*}
\frac{dx(t)}{dt} &= \left[ Ax(t) + a(t, x(t), u(t)) \right] dt \\
&\quad \quad + b(t, x(t), u(t)) dw(t), \quad t \in (0, T], \quad (13)
\end{align*}
\]

where \( A \) is an unbounded linear operator (on a Hilbert space \( H \)), generating a \( C_0 \)-semigroup. Let \( U \) be a metric space. Put

\[
\mathcal{U}[0, T] \triangleq \left\{ u(\cdot) : [0, T] \to U \mid u(\cdot) \text{ is } \mathcal{F}\text{-adapted} \right\}.
\]
Define a cost functional \( J(\cdot) \) as follows:

\[
J(u(\cdot)) \triangleq \mathbb{E} \left[ \int_0^T g(t, x(t), u(t))dt + h(x(T)) \right].
\]

We consider the following optimal control problem:

**Problem (P).** Find \( \bar{u}(\cdot) \in \mathcal{U}[0, T] \) such that

\[
J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0, T]} J(u(\cdot)).
\]

Any \( \bar{u}(\cdot) \in \mathcal{U}[0, T] \) satisfying

\[
J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0, T]} J(u(\cdot)),
\]

is called an *optimal control*, the corresponding \( \bar{x}(\cdot) \equiv x(\cdot; \bar{u}(\cdot)) \) and \( (\bar{x}(\cdot), \bar{u}(\cdot)) \) are called an *optimal state process/trajectory* and *optimal pair*, respectively.
Our next goal is to give a Pontryagin-type maximum principle for the above stochastic optimal control problem.

- The case when $\dim H < \infty$ is now well-understood, see the classical work S. Peng (1990).


- The case when the control appears in the diffusion term and the control set is nonconvex: Only two previous references (i.e., X.Y. Zhou (1993) addressing the linear problem, and S. Tang and X. Li (1994) for the problem with special data).
Main difficulty: How to define the solution to operator-valued BSEE?

- When $H = \mathbb{R}^n$, an $\mathbb{R}^{n \times n}$ (matrix)-BSDE can be regarded as an $\mathbb{R}^{n^2}$ (vector)-valued BSDE.

- When $\dim H = \infty$, $\mathcal{L}(H)$ (with the uniform operator topology) is still a Banach space. Nevertheless, it is neither reflexive nor separable even if $H$ itself is separable.

- There exist no satisfactory stochastic integration/evolution equation theories in general Banach spaces, say how to define the Itô integral $\int_0^T Q(s)dw(s)$ (for an operator-valued process $Q(\cdot)$)? The existing result on stochastic integration/evolution equation in UMD Banach spaces does not fit the present case because, if a Banach space is UMD, then it is reflexive.
Our work in this respect is available in the following paper:


This paper was posted at arXiv on April 15, 2012 but it is now an “old” work. Indeed, just about 2 months later, three very interesting papers (which generalized/improved part of our results) were posted at arXiv:


- Both [FHT] and [DM1]-[DM2] considered only the case of natural filtration, i.e. $F = W$, while [LZ] is addressed to the general filtration.

- Also, [LZ] provided a more detailed analysis on the second order adjoint equations.
5. Well-posedness of vector-valued BSEEs

Consider the following vector-valued backward stochastic differential equation:

\[
\begin{align*}
\begin{cases}
   dy = -A^* y dt + f(t, y, Y) dt + Y dw & \text{in } [0, T), \\
y(T) = y_T.
\end{cases}
\end{align*}
\]

In order to give the definition of the transposition solution to (14), we introduce the following forward stochastic differential equation:

\[
\begin{align*}
\begin{cases}
   dz = (Az + v_1) dt + v_2 dw & \text{in } (t, T], \\
z(t) = \eta.
\end{cases}
\end{align*}
\]

Here \( v_1(\cdot) \in L^1_{\mathcal{F}}(t, T; L^q(\Omega; H)) \), \( v_2(\cdot) \in L^2_{\mathcal{F}}(t, T; L^q(\Omega; H)) \), \( \eta \in L^q_{\mathcal{F}_t}(\Omega; H) \), and \( \frac{1}{p} + \frac{1}{q} = 1 \).
Definition 3. We call \((y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0, T]; L^p(\Omega; H)) \times L^2_{\mathbb{F}}(0, T; L^p(\Omega; H))\) a transposition solution to (14) if for any \(t \in [0, T]\), \(v_1(\cdot) \in L^1_{\mathbb{F}}(t, T; L^q(\Omega; H))\), \(v_2(\cdot) \in L^2_{\mathbb{F}}(t, T; L^q(\Omega; H))\) and \(\eta \in L^q_{\mathbb{F}_t}(\Omega; H)\), it holds that

\[
\mathbb{E}\langle z(T), y_T \rangle_H - \mathbb{E} \int_t^T \langle z(s), f(s, y(s), Y(s)) \rangle_H \, ds
\]

\[
= \mathbb{E}\langle \eta, y(t) \rangle_H + \mathbb{E} \int_t^T \langle v_1(s), y(s) \rangle_H \, ds
\]

\[+ \mathbb{E} \int_t^T \langle v_2(s), Y(s) \rangle_H \, ds.\]
Theorem 3. (Q. Lü and X. Zhang, 2012) Let $H$ be any Hilbert space. For any $y_T \in L^p_{\mathcal{F}_T}(\Omega; H)$, and any $f(\cdot, \cdot, \cdot) : [0, T] \times H \times H \to H$, the equation (14) admits one and only one unique transposition solution $(y(\cdot), Y(\cdot)) \in D_F([0, T]; L^p(\Omega; H)) \times L^2_F(0, T; L^p(\Omega; H))$. Furthermore, there is a constant $C$ such that

$$
| (y(\cdot), Y(\cdot)) |_{D_F([t,T]; L^p(\Omega; H)) \times L^2_F(t,T; L^p(\Omega; H))} 
\leq C \left[ | f(\cdot, 0, 0) |_{L^1_F(t,T; L^p(\Omega; H))} + | y_T |_{L^p_{\mathcal{F}_T}(\Omega; H)} \right],
$$

$\forall t \in [0, T]$. 
6. Well-posedness of operator-valued BSEEs

Further, we consider the following operator-valued backward stochastic evolution equation:

\[
\begin{cases}
    dP = -(A^* + J^*(t))Pdt - P(A + J(t))dt - K^*PKdt \\
    -(K^*Q + QK)dt + Fdt + Qdw \quad \text{in } [0, T), \\
    P(T) = P_T.
\end{cases}
\]

Here \( F \in L^1_{\mathcal{F}}(0, T; L^2(\Omega; \mathcal{L}(H))), \quad P_T \in L^2_{\mathcal{F}_T}(\Omega; \mathcal{L}(H)), \) and \( J, K \in L^4_{\mathcal{F}}(0, T; L^\infty(\Omega; \mathcal{L}(H))). \)
In order to define the transposition solution to the equation (17), we introduce the following two stochastic differential equation:

\[
\begin{align*}
\begin{cases}
  dx_1 &= (A + J)x_1 ds + u_1 ds + Kx_1 dw + v_1 dw \quad \text{in } (t, T], \\
  x_1(t) &= \xi_1,
\end{cases}
\end{align*}
\tag{18}
\]

\[
\begin{align*}
\begin{cases}
  dx_2 &= (A + J)x_2 ds + u_2 ds + Kx_2 dw + v_2 dw \quad \text{in } (t, T], \\
  x_2(t) &= \xi_2.
\end{cases}
\end{align*}
\tag{19}
\]

Here $\xi_1, \xi_2 \in L^4_{\mathcal{F}_t}(\Omega; H)$, $u_1, u_2 \in L^2_{\mathcal{F}}(t, T; L^4(\Omega; H))$ and $v_1, v_2 \in L^4_{\mathcal{F}}(t, T; L^4(\Omega; H))$. 
Definition 4. We call \((P(\cdot), Q(\cdot)) \in D_{\mathbb{F},w}(0, T; L^2(\Omega; \mathcal{L}(H))) \times L^2_{\mathbb{F},w}(0, T; L^2(\Omega; \mathcal{L}(H)))\) a transposition solution to (17) if for any \(t \in [0, T], \xi_1, \xi_2 \in L^4_{\mathcal{F}_t}(\Omega; H), u_1(\cdot), u_2(\cdot) \in L^2_{\mathbb{F}}(t, T; L^4(\Omega; H))\) and \(v_1(\cdot), v_2(\cdot) \in L^4_{\mathbb{F}}(t, T; L^4(\Omega; H)),\) it holds that

\[
\mathbb{E}\langle P_T x_1(T), x_2(T) \rangle_H - \mathbb{E} \int_t^T \langle F(s)x_1(s), x_2(s) \rangle_H ds
\]
\[
= \mathbb{E}\langle P(t)\xi_1, \xi_2 \rangle_H + \mathbb{E} \int_t^T \langle P(s)u_1(s), x_2(s) \rangle_H ds
\]
\[
+ \mathbb{E} \int_t^T \langle P(s)x_1(s), u_2(s) \rangle_H ds + \mathbb{E} \int_t^T \langle P(s)K(s)x_1(s), v_2(s) \rangle_H ds
\]
\[
+ \mathbb{E} \int_t^T \langle P(s)v_1(s), Kx_2(s) \rangle_H ds + \mathbb{E} \int_t^T \langle P(s)v_1(s), v_2(s) \rangle_H ds
\]
\[
+ \mathbb{E} \int_t^T \langle Q(s)v_1(s), x_2(s) \rangle_H ds + \mathbb{E} \int_t^T \langle Q(s)x_1(s), v_2(s) \rangle_H ds.
\]
Denote by $\mathcal{L}_2(H)$ the set of the Hilbert-Schmidt operators on $H$.

**Theorem 4.** (Q. Lü and X. Zhang, 2012) Assume that $H$ is a separable Hilbert space and $L^p_{\mathcal{F}_T}(\Omega)$ ($1 \leq p < \infty$) is a separable Banach space. Then, for any $P_T \in L^2_{\mathcal{F}_T}(\Omega; \mathcal{L}_2(H))$, $F \in L^1_F(0, T; L^2(\Omega; \mathcal{L}_2(H)))$ and $J, K \in L^4_F(0, T; L^\infty(\Omega; \mathcal{L}(H)))$, the equation (17) admits one and only one transposition solution $(P, Q)$ with the regularity $(P(\cdot), Q(\cdot)) \in D_F([0, T]; L^2(\Omega; \mathcal{L}_2(H))) \times L^2_F(0, T; \mathcal{L}_2(H))$. Furthermore,

$$\left| (P, Q) \right|_{D_F([0, T]; L^2(\Omega; \mathcal{L}_2(H))) \times L^2_F(0, T; \mathcal{L}_2(H))} \leq C \left[ |F|_{L^1_F(0, T; L^2(\Omega; \mathcal{L}_2(H)))} + |P_T|_{L^2_{\mathcal{F}_T}(\Omega; \mathcal{L}_2(H))} \right].$$

(20)
Theorems 4 indicates that, in some sense, the transposition solution introduced in Definition 4 is a reasonable notion for the solution to (17).

Unfortunately, we are unable to prove the existence of transposition solution to (17) in the general case.

We shall introduce below a weaker version of solution, i.e., the relaxed transposition solution (to (17)), which looks awkward but it suffices to establish the desired Pontryagin-type stochastic maximum principle for Problem (P) in the general setting.
Definition 5. We call \((P(\cdot), Q(\cdot), \hat{Q}(\cdot)) \in D_{F,w}([0, T]; L^4_3(\Omega; \mathcal{L}(H))) \times \mathcal{Q}[0, T]\) a relaxed transposition solution to (17) if for any \(t \in [0, T]\), \(\xi_1, \xi_2 \in L^4_{F_t}(\Omega; H)\), \(u_1(\cdot), u_2(\cdot) \in L^2_F(t, T; L^4(\Omega; H))\) and \(v_1(\cdot), v_2(\cdot) \in L^4_F(t, T; L^4(\Omega; H))\), it holds that

\[
\mathbb{E}\langle P_T x_1(T), x_2(T) \rangle_H - \mathbb{E} \int_t^T \langle F(s) x_1(s), x_2(s) \rangle_H ds
\]

\[
= \mathbb{E}\langle P(t) \xi_1, \xi_2 \rangle_H + \mathbb{E} \int_t^T \langle P(s) u_1(s), x_2(s) \rangle_H ds
\]

\[
+ \mathbb{E} \int_t^T \langle P(s) x_1(s), u_2(s) \rangle_H ds + \mathbb{E} \int_t^T \langle P(s) K(s) x_1(s), v_2(s) \rangle_H ds
\]

\[
+ \mathbb{E} \int_t^T \langle P(s) v_1(s), K x_2(s) \rangle_H ds + \mathbb{E} \int_t^T \langle P(s) v_1(s), v_2(s) \rangle_H ds
\]

\[
+ \mathbb{E} \int_t^T \langle v_1(s), \hat{Q}(t)(\xi_2, u_2, v_2)(s) \rangle_H ds + \mathbb{E} \int_t^T \langle Q(t)(\xi_1, u_1, v_1)(s), v_2(s) \rangle_H ds.
\]
• It is easy to see that, if \((P(\cdot), Q(\cdot))\) is a transposition solution to (17), then one can find a relaxed transposition solution \((P(\cdot), Q(\cdot), \hat{Q}(\cdot))\) to the same equation (from \((P(\cdot), Q(\cdot))\)). Indeed, they are related by

\[
Q(s)x_1(s) = Q^t(\xi_1, u_1, v_1)(s),
\]

\[
Q(s)^*x_2(s) = \hat{Q}^t(\xi_2, u_2, v_2)(s).
\]

This means that, we know only the action of \(Q(s)\) (or \(Q(s)^*\)) on the solution processes \(x_1(s)\) (or \(x_2(s)\)).

• However, it is unclear how to obtain a transposition solution \((P(\cdot), Q(\cdot))\) to (17) by means of its relaxed transposition solution \((P(\cdot), Q(\cdot), \hat{Q}(\cdot))\). It seems that this is possible but we cannot do it at this moment.
Well-posedness result for the equation (17) in the general case

**Theorem 5.** (Q. Lü and X. Zhang, 2012) Assume that $H$ is a separable Hilbert space, and $L^p_{\mathcal{F}_T}(\Omega; \mathbb{C})$ ($1 \leq p < \infty$) is a separable Banach space. Then, for any $P_T \in L^2_{\mathcal{F}_T}(\Omega; \mathcal{L}(H))$, $F \in L^1_{\mathcal{F}}(0, T; L^2(\Omega; \mathcal{L}(H)))$ and $J, K \in L^4_{\mathcal{F}}(0, T; L^\infty(\Omega; \mathcal{L}(H)))$, the equation (17) admits one and only one relaxed transposition solution $(P(\cdot), Q(\cdot), \hat{Q}(\cdot))$. Furthermore,

$$
\|P\|_{\mathcal{L}(L^2_{\mathcal{F}}(0, T; L^4(\Omega; H)), L^2_{\mathcal{F}}(0, T; L^4(\Omega; H)))} + \sup_{t \in [0, T]} \left\| \left( Q(t), \hat{Q}(t) \right) \right\|_{\left( \mathcal{L}(L^4_{\mathcal{F}_t}(\Omega; H) \times L^2_{\mathcal{F}_t}(t, T; L^4(\Omega; H)) \times L^2_{\mathcal{F}_t}(t, T; L^4(\Omega; H)), L^2_{\mathcal{F}_t}(t, T; L^3(\Omega; H)) \right)^2} 
\leq C \left[ |F|_{L^1_{\mathcal{F}}(0, T; L^2(\Omega; \mathcal{L}(H)))} + |P_T|_{L^2_{\mathcal{F}_T}(\Omega; \mathcal{L}(H))} \right].
$$

(21)
• Both [FHT] and [DM1, DM2] used a quite different strategy to analyze the solvability of the operator-valued BSEE (17).

• The adjust component $Q(\cdot)$ (in the operator-valued BSEE (17)) was suppressed in both [FHT] and [DM1, DM2]. This does not affects the proof of Pontryagin-type stochastic maximum principle but may be inconvenient for other purposes. For example, using our method, it is easy to solve the following BSEE:

$$
\begin{aligned}
    dP &= -(A^* + J^*(t))P dt - P(A + J(t))dt - K^*PK dt \\
    &\quad - (K_1^*Q + QK_2)dt + F dt + Q dw \\
    P(T) &= P_T.
\end{aligned}
$$

Here both $K_1$ and $K_2$ may be different from $K$. The cost is that our analysis is much more complicated than that in [FHT] and [DM1, DM2].
7. Pontryagin-Type Stochastic Maximum Principle

For \((t, x, u, k_1, k_2) \in [0, T] \times H \times U \times H \times H\), write
\[
\mathbb{H}(t, x, u, k_1, k_2) = \langle k_1, a(t, x, u) \rangle_H + \langle k_2, b(t, x, u) \rangle_H - g(t, x, u).
\]

**Theorem 6.** (Q. Lü and X. Zhang, 2012) Let \((\bar{x}(\cdot), \bar{u}(\cdot))\) be an optimal pair of Problem (P), and let \((y(\cdot), Y(\cdot))\) be the transposition solution to (14) with \(p = 2\), and \(y_T\) and \(f(\cdot, \cdot, \cdot)\) given by

\[
\begin{align*}
y_T &= -h_x(\bar{x}(T)), \\
f(t, y_1, y_2) &= -a_x(t, \bar{x}(t), \bar{u}(t))^* y_1 - b_x(t, \bar{x}(t), \bar{u}(t))^* y_2 \\
&\quad + g_x(t, \bar{x}(t), \bar{u}(t)).
\end{align*}
\]

(22)
Assume that \( b_x(\cdot, \bar{x}(\cdot), \bar{u}(\cdot)) \in L^4_{\text{IF}}(0, T; L^\infty(\Omega; \mathcal{L}(D(A)))) \), and \((P(\cdot), Q(\cdot), \hat{Q}(\cdot))\) is the relaxed transposition solution to the equation (17) in which \( P_T, J(\cdot), K(\cdot) \) and \( F(\cdot) \) are given

\[
\begin{align*}
P_T &= -h_{xx}(\bar{x}(T)), \quad J(t) = a_x(t, \bar{x}(t), \bar{u}(t)), \\
K(t) &= b_x(t, \bar{x}(t), \bar{u}(t)), \quad F(t) = -H_{xx}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)).
\end{align*}
\]

Then

\[
\text{Re } H(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)) - \text{Re } H(t, \bar{x}(t), u, y(t), Y(t)) - \frac{1}{2} \left\langle P(t) \left[ b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u) \right], b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u) \right\rangle_H \\
\geq 0, \quad \forall u \in U, \text{ a.e. } [0, T] \times \Omega.
\]
• Our assumption $b_x(\cdot, \bar{x}(\cdot), \bar{u}(\cdot)) \in L^4_{\mathbb{P}}(0,T;L^{\infty}(\Omega;\mathcal{L}(D(A))))$ means extra regularities on the data (say $x_0$ and so on). Nevertheless, some of these regularities are also needed, say to check the differentiability of the nonlinearities $a(\cdot,\cdot,\cdot)$ etc. Note that the function $\sin(y)$ from $L^2(0,1)$ to $L^2(0,1)$ is NOT Fréchet differentiable (but it is Fréchet differentiable, say from $H^1_0(0,1)$ to $L^2(0,1)$).

• Considerably different from the works [FHT] and [DM1]-[DM2], the continuity of $P(t)$ is NOT used in the proof of our maximum principle. Therefore, one may extend the optimal control problem to the setting of unbounded control, say boundary/point control problems. In this case, the continuity of $P(t)$ is NOT always guaranteed.
Thank You